The process of verifying affirmative and negative sentences against pictures

CHAO-MING CHENG and HUEI-JANE HUANG
National Taiwan University, Taipei, Taiwan, Republic of China

Experiments 1 and 2 investigated sentence context effects on picture encoding in a sentence-picture verification and a picture naming task, respectively. The results showed that a picture following a sentence was encoded faster when the two were congruent than when they were incongruent. Experiment 3 compared two conditions: Under one condition, true affirmatives, false affirmatives, true negatives, and false negatives were mixed in each block of presentation. Under the other condition, different sentence types were presented in different blocks. The results showed (1) that errors committed in verification were largely negation errors, but seldom falsification errors, and (2) that there was a decrease of falsification time, but a resistance to change in negation time from the mixed to the blocked presentation. These results were interpreted to mean that falsification time results from a longer time to encode the picture and to confirm or disconfirm the truth index value by the polarity of the sentence in false affirmatives and true negatives than in true affirmatives and false negatives, whereas negation time and negation errors result from a response-suppression operation.

In 1972, Clark and Chase proposed a model that accounts for the results of experiments on sentence verification. According to this model, when a subject is required to determine whether a sentence is true relative to what is shown in a picture, he/she first encodes both the sentence and the picture in a propositional format, so the two can be properly compared. A propositional format is said to be an abstract subject-predicate relation embedded in a polarity marker that is either affirmative or negative for the sentence and is always affirmative for the picture. Verifications then start from a comparison of the embedded part of the sentence representation with that of the picture representation, followed by a comparison of the two embedding parts (the polarity markers). The model assumes that the comparison operations also keep track of a truth index whose value is initially set at "true" at the start of the comparisons. The value of the truth index remains unchanged as long as a comparison results in a match, but it is shifted to its contrary whenever a mismatch occurs. This mismatch and index shifting together will consume an extra amount of time. The model also assumes that negative sentences are more difficult, and hence take longer, to encode than affirmatives.

Using the above notions, Clark and Chase (1972) were able to predict an increase of verification latency in the order of true affirmatives, false affirmatives, false negatives, and true negatives. For example, true

negatives take longer to verify than true affirmatives because sentence encoding time is longer in the former and because both the inner and outer parts of the sentence and picture representations do not match in the former but do match in the latter.

Subsequently, Carpenter and Just (1975) developed a more parsimonious model based on a single underlying iterative operation that predicts verification latencies as well as the Clark and Chase (1972) model does. This model differs from the Clark and Chase model in two significant aspects: It assumes (1) that whenever two corresponding constituents mismatch, the entire comparison sequence is reinitialized from the beginning, with the mismatched parts tagged and then treated as a match on subsequent comparisons, and (2) that each comparison itself consumes a fixed amount of time, and therefore the latency of verifying a task is completely determined by the total number of comparisons involved in the task (the time taken to shift a truth index value was treated in this model as practically negligible).

Current investigations (Carpenter & Just, 1975; Catlin & Jones, 1976; Shoben, 1978) on these two models have centered on their ability to account for the systematic difference in performance between the condition in which sentence precedes picture and the condition in which picture precedes sentence. This difference is best illustrated by the ratio of negation time to falsification time being about 4:1 in the sentence-first condition and about 2:1 in the picture-first condition (see Carpenter & Just, 1975). Carpenter and Just attributed this ratio difference to negatives' being represented with a large scope of negation in the sentence-first condition and represented with a small scope of negation in the picture-first condition. Thus, a mismatch between

the two polarity markers in the sentence-first condition, according to the Carpenter and Just model, would result in more recomparisons, thereby increasing verification time for negatives. In other words, Carpenter and Just attributed the change in the ratio of negation time to falsification time across the two presentation conditions to the change in negation time. However, as Catlin and Jones (1976) noted, the change in the ratio did not result from a change in negation time, but rather from a change in falsification time, with falsification time found to be longer in the picture-first than in the sentence-first condition.

To allow the Carpenter and Just (1975) model to be able to overcome this difficulty, Shoben (1978) recently modified the model by suggesting that a negative sentence is always represented with a large scope of negation, no matter whether it precedes or follows a picture to be compared. In addition, Shoben suggested that the Carpenter and Just model should adopt the notion of sentence recoding, for the same reason that Clark and Chase (1972) used it, to explain the process differences between the picture-first and sentence-first conditions. The resulting revised version of the Carpenter and Just model is then able to account for the change in the ratio of negation time to falsification time in terms of a change in falsification time (see Shoben, 1978). According to this revised version, the number of operations required for true affirmatives, false affirmatives, false negatives, and true negatives is k, $k + 1$, $k + 4$, and $k + 5$, respectively, in the sentence-first condition, and k, $k + 2$, $k + 4$, and $k + 6$, respectively, in the picture-first condition. It would follow, then, from this prediction, as well as from the Clark and Chase (1972) model, that the change in falsification time is due to an increase of verification time for false affirmatives and true negatives, as the presentation is moved from the sentence-first to the picture-first condition, whereas that for true affirmatives and false negatives remains unchanged through the two presentation conditions. Unfortunately, the latter part of this prediction is not supported by existing data. For example, Clark and Chase (1972, Experiment 2) also demonstrated a latency increase from the sentence-first to the picture-first condition in verifying true affirmatives and false negatives.

The present study assessed these two models through a direct evaluation of basic assumptions on which they are based. Except that the Clark and Chase (1972) model assumes a potential contribution of differential sentence encoding to negation time, both models agree with each other in the notion that falsification and negation time are interpreted by the two models to be the extra times consumed by the same mental operation, which is triggered by a comparison of two corresponding constituents that mismatch. In the Clark and Chase model, the operation is that of finding a mismatch between two constituents and shifting a truth index value to its contrary due to a mismatch of constituents. In the Carpenter and Just (1975) model, it is that of recomparisons of constituents. In other words, these two models have commonly regarded the process at the comparison stage, rather than at other stages, of the information processing sequence as crucial and responsible for verification latency differential across the four tasks.

It should be noted that the present interpretation has regarded the Clark and Chase (1972) model as a comparison model without neglecting the fact that the model also incorporates the sentence-encoding parameter. In the present evaluation of the model, the contribution of this parameter will also be considered. However, it is worth mentioning that the predictive power of this model is rather insensitive to the inclusion-exclusion of this parameter in the model (see Clark & Chase, 1972). Furthermore, it is the comparison rather than the encoding part of the model that is important and can explain the fact that verification latency still increases in the order of true affirmatives, false affirmatives, false negatives, and true negatives, even when sentence encoding (or comprehension) time is either experimentally matched across the two sentence types (Gough, 1966) or excluded from response measurement (MacLeod, Hunt, & Mathews, 1978).

EXPERIMENT 1

In Experiment 1, we tested these two comparison models by comparing two verification conditions differing only in the form in which the picture to be compared with the sentence was presented. Under one condition, the picture was printed in a clear form. Under the other, the picture was printed in a visually degraded form. It was reasoned that, under the visual degradation, the subject should take time to differentiate the stimulus picture from the visual noise and therefore verification latency should be longer under this condition than under the visually clear condition. However, the variation of picture form would not affect the comparison process that is supposed to occur after the picture has been encoded. Nor would it affect sentence encoding time in a sentence-first presentation. Thus, the two comparison models will predict that there will be no interaction of picture form with truth value within the same sentence type and no interaction of picture form with sentence type. On the other hand, if there is such a significant interaction, it would imply a differential picture encoding effect on sentence-picture verification.

Method

Subjects. The subjects were 30 undergraduates of the National Taiwan University. They participated in the experiment as part of their requirements for an introductory psychology course.

Stimulus materials. The sentence type used in the experiment was similar to "The figure on the right is (isn't) a triangle,"
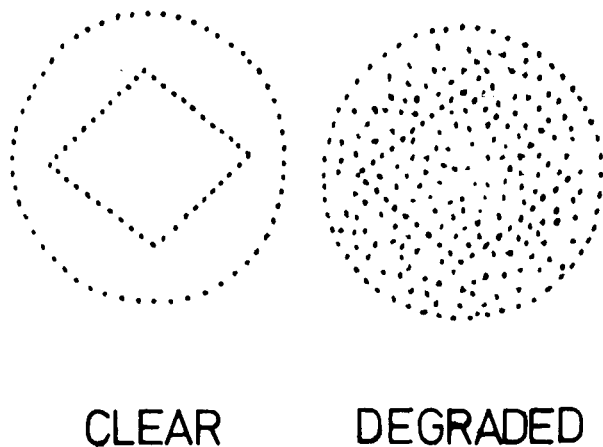
CLEAR   DEGRADED

Figure 1. Examples of the two picture forms used in Experiment 1.

written in Chinese. There were eight different sentences, with four affirmatives and four negatives, used to describe four geometrical figures: diamond, square, trapezoid, and triangle. The four figures were prepared in two forms. One form was visually clear, in which a figure was drawn with dotted lines surrounded by a large circle also drawn with dotted lines. The circle served as an attention scope for locating the figure. The other form was visually degraded; it was produced in the same way as the clear one, except that there were random dots filled in the scope circle. Examples of these two visual forms of picture are shown in Figure 1. A pretest confirmed that the figure in the degraded form was more difficult to identify than the figure in the clear one.

For the purpose of experimental trials, each sentence was paired twice with the picture that it described and paired once with two other pictures. The result was 32 displays, with 8 displays for each of the four tasks. Each display was printed on a card for tachistoscopic presentation, with the sentence on the left and the picture on the right. The characters of each sentence were equally spaced and were adjusted such that the sentence length was kept constant across the two sentence types. The space between the last character of the sentence and the picture was also kept constant (4.5 cm) across all displays. Since there were two picture forms, there were 64 displays used for the experiment.

Apparatus. A tachistoscope was used to display the stimulus cards. An electric circuit was arranged such that it started a Lafayette stopclock when a stimulus card appeared on the window and held the clock running until the subject pushed one of the two response buttons. There were two light bulbs connected to the two response buttons, which lit as the subject pushed the corresponding button, by means of which the experimenter could detect the button that was pressed for a given trial.

Design. The experiment was carried out in a 2 by 2 by 2 completely within-subjects design. The three variables were (1) picture form (clear vs. degraded), (2) sentence type (affirmative vs. negative), and (3) truth value (true vs. false). There were eight trials for each of the eight conditions formed by these three variables. The 64 trials for each subject were presented with block randomization, with each block unit consisting of a trial for each of the eight conditions. The presentation sequence of the eight blocks was randomized for each subject.

Procedure. The subjects were tested individually. The subject was instructed that a stimulus display would appear on the window after a warning signal and he/she was required to attend to the sentence first, then to the picture, and then to push either a "true" or a "false" button as quickly as possible. The subject

was first given eight practice trials, with one trial for each condition, and was given feedback about the correctness of his response as required by the task. This practice session was then followed by the experimental session of eight blocks of eight stimulus presentations and tests without feedback. For half the subjects, the "true" button was located on the right and was to be pushed with the right thumb. For the other half, the "true" button was on the left, to be pushed with the left thumb. The latencies of correct responses, which were timed from the stimulus onset to the response onset, were recorded. The experiment lasted about 20 min for each subject.

Results

Figure 2 presents the mean latencies for the eight conditions. Each point is the average of the latencies of all correct responses made by the 30 subjects in each condition. Figure 2 shows that overall, true affirmatives were verified fastest and true negatives slowest, with false affirmatives and false negatives in between. Thus, the usual pattern of verification latency across the four tasks was also obtained in the present experiment with stimulus sentences written in Chinese and verified by Chinese subjects. Figure 2 also shows that the latencies were longer in the degraded condition than in the corresponding clear conditions.

The main purpose of the present experiment was to investigate whether picture form would interact with the variables (sentence type and truth value) in determining verification latency. An analysis of variance of the
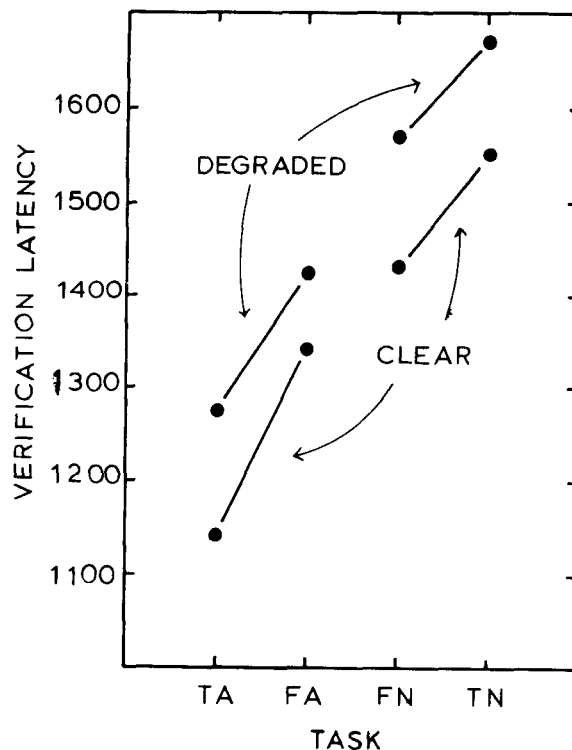

Figure 2. Verification latency (in milliseconds) as a function of sentence type, truth value, and picture form (TA = true affirmative; FA = false affirmative; FN = false negative; TN = true negative).

latencies across true and false affirmatives showed that the verification latencies were longer under the degraded than under the clear condition [$F(1,29) = 44.5$,[1] MSe = 7,851], that false affirmatives took longer to verify than true affirmatives across the two picture forms [$F(1,29) = 141.9$, MSe = 6,472], and that there existed a significant interaction of Picture Form by Truth Value [$F(1,29) = 4.3$, MSe = 4,286]. The significant interaction resulted primarily from the fact that false affirmatives took 199 msec longer to verify than true affirmatives under the clear condition, but false affirmatives took only 153 msec longer under the degraded condition. These two latency differences were significant, as tested by a Wilcoxon test ($z = 4.76$, $N = 30$, $T = 1$, for the former, and $z = 4.68$, $N = 30$, $T = 5$, for the latter). The finding of this significant interaction cannot be explained by either of the two comparison models.

An analysis of variance of the latencies across the negatives also revealed a longer verification time under the degraded than under the clear condition [$F(1,29) = 16.11$, MSe = 22,734]. True negatives took longer to verify than false negatives across the two picture forms [$F(1,29) = 57.34$, MSe = 8,499.9]. The difference in latency between true negatives and false negatives was 120 msec under the clear condition and 101 msec under the degradation. These two latency differences were also significant as tested by a Wilcoxon test ($z = 3.24$, $N = 30$, $T = 75$, for the former, and $z = 2.75$, $N = 29$, $T = 90.13$, for the latter). While these results may indicate a pattern of interaction of picture form with truth value similar to the one found for the affirmatives, the interaction effect failed to reach an acceptable level of significance.

The amount of negation time was 253.5 msec under the clear condition and 272 msec under the degradation condition. These two values were not significantly different from each other, suggesting that the process responsible for negation time is not influenced by the variation of picture form.

Discussion

The results of the present experiment revealed a significant interaction of picture form with truth value in verifying affirmative sentences. This finding is not consistent with the prediction of the two comparison models. A way to account for this finding is to assume that, in sentence-picture verification, encoding the sentence first produces different effects on encoding picture in different tasks, depending upon the information congruity between the sentence and the picture. It is likely that, when a sentence like "The figure on the right is a triangle" is paired with a picture of a triangle, encoding the sentence may produce a facilitative effect on the encoding of the picture, because the activation processes of the picture may share the same pathway with those of the sentence and the pathway has already been activated by the presentation of the sentence. This facilitative effect would result in a saving of an amount of time, a, for encoding the picture, compared with the time for encoding a picture presented alone. On the other hand, when the same sentence is paired with an incongruent picture of a diamond, the encoding of the sentence may create an interference effect on the encoding of the picture, because the activation of lexical information about the picture will not occur until a necessary suppression of the currently activated information about the sentence has been completed. This interference effect would result in an extra consumption of time, b, for encoding the picture, compared with the time for encoding a picture presented alone. Thus, in the sentence-first condition, it will take Time a+b longer for encoding the picture to verify false affirmatives than to verify true affirmatives.

When under the visual degradation, the geometric figures could not easily be discriminated until most of their features had been extracted from the visual noise, a task of picture identification that might force the subject to name the picture by relying more on stimulus information than on context. Thus, both the positive and negative priming effect of encoding the sentence on encoding the picture would be attenuated under the visual degradation condition, resulting in a smaller difference in verification time between false and true affirmatives.

The above encoding explanation of falsification time significantly contrasts the two comparison models that emphasize the importance of the comparison rather than the encoding process to falsification time. According to this encoding explanation, similar priming effects on encoding the picture should also occur in verifying negatives. Unfortunately, this prediction failed to gain support from the results of verifying negatives, which showed a lack of significant interaction of picture form with truth value. While this lack of significant interaction may be taken as evidence supporting the comparison models, the result is rather difficult to assess. The result may not be a direct consequence of a simple differential picture encoding operation, should it be there, as there may be some other mechanism (e.g., response shift) also involved in verifying the negatives. Of more importance is the fact that we do not know how these mental operations work together. One possibility is that they work in an additive manner; each operation begins only when the preceding one has ended. This additivity principle, however, may not be completely appropriate in the present discussion of sentence-picture verification. One immediate consideration against it is that the differences in response latency between false and true affirmatives and between true and false negatives, which should index the same process according to this additivity principle, are in fact not equal, as the previous studies and the present experiment consistently show. Thus it leaves the possibility that mental operations responsible for verifying negative sentences overlap in time, which would certainly compli-

cate the present evaluation of the contribution of picture encoding to verification of negative sentences.

EXPERIMENT 2

Experiment 2 provided a more direct test on the encoding hypothesis, which was suggested by the results of Experiment 1. The method of investigation was to measure and compare the time taken to name a picture following different context sentences in different tasks. The assumption for the present investigation is that a picture cannot be represented in a propositional format until it is categorized or named. The encoding hypothesis would therefore predict that reading a sentence like "These dots are (aren't) red" would result in a faster naming of a congruent picture of red dots than of an incongruent picture of dots of a color other than red.

Method

Subjects. Twenty-five undergraduates of the National Taiwan University served as the subjects. They participated in the experiment as part of their requirements for an introductory psychology course.

Stimulus materials. The stimulus sentences were eight different sentences of the form "These dots are (aren't) red," written in Chinese. Half of the sentences were affirmatives and the other half were negatives, whose complements were red, yellow, blue, and green. The picture was three rows of three dots of one color, red, yellow, blue, or green. Pairings of the sentences and pictures were the same as those in Experiment 1; there were eight displays for each of the following conditions: (1) an affirmative followed by a congruent picture (CA), (2) an affirmative followed by an incongruent picture (IA), (3) a negative sentence followed by a congruent picture (CN), and (4) a negative followed by an incongruent picture (IN). The sentence and the picture of each pairing were separately printed on a card for sequential presentation. The characters of each sentence were kept equally spaced, and the spaces were adjusted such that both the affirmatives and negatives subtended the same visual angle.

Apparatus. A two-field tachistoscope was used to display the sentence and picture sequentially. The sequence was arranged such that a test trial started with a sentence presented first on the left inspection field for 1.5 sec. The sentence was then terminated by the onset of a picture on the right field. A Lafayette stopclock started at the same time the picture was presented. A mild vocal sound through a throat microphone coupled to a voice key stopped the clock and also removed the picture from the presentation.

Design and Procedure. In order to be sure that the sentence presented was not ignored by the subject in this picture naming task and that the subject kept information about the sentence until he named the subsequently presented picture, the subject was required on each trial to perform two tasks. One was to vocalize as rapidly as possible the name of the color shown in the picture. This task was followed by the second task of writing on an answer sheet the two color names from the sentence and the picture. These task requirements were announced to the subject before the experiment.

The experiment used a within-subjects design; each subject was tested under each of the four conditions. The four conditions were presented to the subject with block randomization, with each block consisting of a trial for each condition.

The subjects were tested individually. Each subject was first given 8 practice trials, with 2 trials for each condition, and then was given 32 test trials across the four conditions. The latencies

of correct naming responses, which were timed from the onset of the picture to the onset of the vocalization, were recorded for analysis. The experiment lasted about 15 min for each subject.

Results and Discussion

Since the memory task was used for the purpose of forcing the subject to pay attention to the sentence as it was presented and to keep information about the sentence until the presentation of the picture, the present picture naming experiment may be regarded as an experiment to investigate the encoding process involved in sentence-picture verification. The memory data were not the present concern and were discarded. The overall error rate was quite low; only 6 out of a total of 800 responses, across the 25 subjects and the four conditions, were incorrect.

The mean latencies for CA, IA, CN, and IN were 637, 716, 638, and 721 msec, respectively. Each score was the average of the latencies of the correct naming responses made by the 25 subjects in each condition. As predicted, the data showed faster naming under the two congruent conditions than under the two incongruent conditions. There appeared to be no difference between the affirmatives and the negatives at either level of congruity. An analysis of variance of the latencies showed a significant effect of sentence-picture congruity $[F(1,24) = 30.44, MSe = 2,022]$. The difference in naming speed between the congruent and incongruent conditions was 79 msec when the context sentences were affirmatives and 83 msec when the context sentences were negatives. Although these two differences appear to be small, they were significant by a Wilcoxon test $(z = 4.37, N = 25, T = 0$, for the former, and $z = 4.18, N = 25, T = 7)$. The effect of sentence type was not significant, nor was the interaction of Sentence-Picture Congruity by Sentence Type.

Thus, the results of the present experiment showed a clear sentence context effect on picture naming. This finding is in accord with the Dyer and Severance (Note 1) study on Stroop interference with successive presentations of separate words and colors. In their study, one of the words "red," "blue," "green," and "yellow," printed in black, was presented to the subject before the presentation of a color patch of red, blue, green, or yellow. The results of this study showed that incongruent words preceding color patches delayed color naming relative to the condition with congruent words. In another study by Dyer (1971), in which a word naming a color was seen in black first and then was seen in one of four colors, delayed color naming was also found under the incongruent relative to the congruent combinations of words and colors. When compared with a control condition, responses under the congruent condition showed facilitation and those under the incongruent condition showed interference.

The results of the Dyer (1971; Dyer & Severance, Note 1) studies and of the present experiment can be explained by the cost-benefit theory proposed by

Posner and Snyder (1975a, 1975b). According to this theory, an input item will automatically activate a specific pathway in the nervous system. Another item that shares the same pathway will be processed faster and thus will be facilitated. Once a subject invests his attention to the input pathway, the benefit is increased and is accompanied by a widespread cost or inhibition in the ability of any other activated pathway to reach the mechanism of attention. Thus, according to this explanation, a picture following a congruent sentence will be encoded faster because the two share the same pathway activation. On the other hand, the processing of a picture following an incongruent sentence will be inhibited from rising to active attention, because the sentence input pathway is relatively attention catching. However, Posner and Snyder did not make explicit how an input, once activated but attention-inhibited, finally becomes available for readout. Provisionally, it may be accomplished through a restimulation of the inhibited input, as Klein (1964) suggested, and is probably accompanied by an effortful suppression of the activated and attention-catching pathway. Thus, we interpreted the present sentence context effect on picture naming to be an effect of encoding rather than an effect of response facilitation competition. In summary, the results of the present experiment and Experiment 1 provided two lines of evidence converging on an explanation that the process responsible for the occurrence of falsification time occurs prior to the comparison stage.

EXPERIMENT 3

In Experiment 3, we further tested the two comparison models by comparing two presentation conditions. Under one condition, the four tasks of true affirmatives, false affirmatives, false negatives, and true negatives were mixed in each block of presentation, so that the subject under this condition experienced uncertainty about the sentence type that was to be presented on each trial. This is the usual way adopted by previous studies to present the four tasks. Under the other condition, different tasks with respect to sentence type were presented in different blocks of presentation, so that the subject under this condition was able to know in advance the sentence type that was to be compared with the picture for a given trial.

Given the above two presentation conditions, it may be reasoned that the subject will take different strategies to deal with the tasks under different presentation conditions. The subject under the mixed presentation has to compare the polarity markers after comparing the inner constituents and before an appropriate response can be initiated and executed. On the other hand, under the blocked presentation, the subject may not have to compare the polarity markers because they are kept constant across all trials of a block. Thus, without any

negatives, the subject can simply study and compare the constituents and initiate an appropriate response based on the result of the comparison (e.g., to make a "true" response if the comparison results in a match and a "false" response if the comparison results in a mismatch), without consulting the polarity markers. Similarly, without any affirmatives, the subject in the "negative" blocked presentation can simply compare the inner constituents and then initiate a response corresponding to the contrary of the truth index value resulting from the comparison of inner constituents.

Thus, based on the above reasoning, the process of comparing the polarity markers will be bypassed in sentence-picture verification under the blocked presentation, and the two comparison models therefore predict a decrease of negation time and no change in falsification time as the condition is shifted from the mixed to the blocked presentation. This prediction will be clearer if Table 1 is consulted. In Table 1, the two comparison models are summarized. The symbols S, P, Ci, Co, and R represent the time taken to encode a sentence, to encode a picture, to compare the inner constituents, to compare the polarity markers, and to initiate an appropriate response, respectively. The time components within the parentheses are assumed to be saved, based on the above reasoning, under the blocked presentation. Note that the amount of time to be saved in true affirmatives is the same as that in false affirmatives and that in false negatives is the same as that in true negatives, but the latter two are longer than the former two.

Alternatively, it may be reasoned that the process of comparing the polarity markers is still indispensable even under the blocked presentation; the only advantage of the blocked over the mixed presentation is a saving of time taken to encode sentences. If this time saving is equally effective across the two sentence types, then both models would predict a general decrease of verification time across the four tasks, but there is no interaction of the variable of presentation condition with the two variables of sentence type and truth value. On the other hand, if the blocked presentation results in equal encoding time for both affirmatives and negatives, then the Clark and Chase (1972) model would still predict a change in negation time and no change in falsi-

Table 1
Summary of the Clark and Chase (1972) and the
Carpenter and Just (1975) Models

Task	Clark and Chase	Carpenter and Just
TA	$S + P + Ci + (Co) + R$	$S + P + Ci + (CO) + R$
FA	$S + P + Ci'* + (Co) + R'$	$S + P + 2Ci + (CO) + R$
FN	$S' + P + Ci + (Co') + R'$	$S + P + 2Ci + (4Co) + R$
TN	$S' + P + Ci' + (Co') + R''$	$S + P + 3Ci + (4Co) + R$

Note—TA = true affirmative; FA = false affirmative; FN = false negative; TN = true negative. See text for explanation of the other abbreviations. *$S' > S$, $Ci' > Ci$, $Co' > Co$, $R'' > R' > R$.

fication time across the two presentations, a prediction similar to the first prediction.

Method

The stimulus materials, design, and procedure were the same as those used in Experiment 2, except that the subject in the present experiment was required to perform a sentence-picture verification task instead of picture naming and the sentence and picture of each pairing were simultaneously presented, with the sentence on the left and the picture on the right. The equipment was the same as that used in Experiment 1.

Each subject in the experiment received four blocks of 16 stimulus displays, with two blocks of mixed presentation and two blocks of blocked presentation. The former two blocks each consisted of four displays for each of the four tasks. These displays were mixed in the block. One of the two blocked presentations consisted of the eight displays for true affirmatives and the eight for false affirmatives used in the mixed presentations, which were mixed in the block with the restriction that stimulus displays for a task do not occur on more than three consecutive trials. The other blocked presentation consisted of the eight displays for false negatives and the eight for true negatives used in the mixed presentations.

Before each block, the subject was informed about the stimulus materials, especially the sentence type(s), used in the block. The sequence of presenting the four blocks was randomly decided for each subject. The subjects were 30 undergraduates of the National Taiwan University. The experiment lasted about 25 min for each subject.

Results and Discussion

Errors. Table 2 shows the percent errors for each task under the mixed and blocked presentations. The maximum number of errors for each task was 240. The error pattern, like the latency pattern usually found in this area, showed an increase of errors in the order of true affirmatives, false affirmatives, false negatives, and true negatives. However, unlike the latency pattern, it revealed a relatively large difference between the two "negative" and the two "affirmative" tasks, and a small difference between the two truth values within the same sentence type. Table 2 also shows that the error rates tended to decrease from the mixed to the blocked presentation. This was especially true for negative sentences.

In the present analysis of errors, we defined falsification errors, as the counterpart of falsification time, as errors committed more due to a mismatch between two inner constituents. For the same reason, we defined negation errors as errors committed more in verifying negatives than in verifying affirmatives, namely, errors committed more due to a mismatch between the two polarity markers. Following these definitions, falsifica-

tion and negation errors were 2.55% and 9.55%, respectively, under the mixed presentation and 1.25% and 5.45%, respectively, under the blocked presentation. A chi-square test showed that the two negation error rates were highly significant $[\chi^2(1) = 25.2$ for the former, and $\chi^2(1) = 14.1$ for the latter]. On the other hand, the same test showed that neither of the two falsification error rates was significant.

If we consider response errors and latency as two different measures of the same process, the above finding of significant negation errors and nonsignificant falsification errors cannot be explained easily by the Clark and Chase (1972) model. This model postulates that the falsification and negation processes share the mental operations of finding mismatch between two corresponding constituents and of shifting a truth index value to its opposite due to this mismatch. It would follow that when this mental operation works either incorrectly or unsuccessfully, it will result in response errors. However, in order to account for the present error pattern, it is necessary for this model to explain why this mental operation does not result in significant errors following a mismatch of inner constituents and yet does result in significant errors following a mismatch of outer constituents.

On the other hand, the above error pattern is consistent with the prediction of the Carpenter and Just (1975) model. The ratio of negation errors to falsification errors was about 4:1 in both the mixed and the blocked presentations, which is the same ratio found in the latency data, and can therefore be explained by the iterative process of constituent comparisons. However, the finding of significant negation errors and nonsignificant falsification errors is also consistent with a view regarding the falsification process as different from the negation process; the former is an error-free mechanism and the latter is the error-prone mechanism.

The difference in percent errors between the mixed and blocked presentations was significant for the negatives $[\chi^2(1) = 4.39]$ but was not significant for the affirmatives. These results may suggest that, as the condition is shifted from the mixed to the blocked presentation, the step taken to compare the two polarity markers is saved, and therefore potential errors resulting from incorrect comparisons of the two polarity markers are reduced.

Correct latencies. The correct verification latencies are represented by the four solid lines of Figure 3. Each point represents the mean of the correct latencies made from a total of 240 trials, across the 30 subjects. An overall analysis of variance of the latencies showed that the mixed presentation resulted in longer verification time than did the blocked presentation, summed across the four tasks $[F(1,29) = 64.76, MSe = 4,135.4]$. The negatives took longer to verify than the affirmatives, across the two presentation conditions $[F(1,29) = 181.75, MSe = 2,619.7]$. The "false" reaction was also

Table 2
Percent Errors as a Function of Sentence Type, Truth Value, and Presentation Condition in Experiment 3

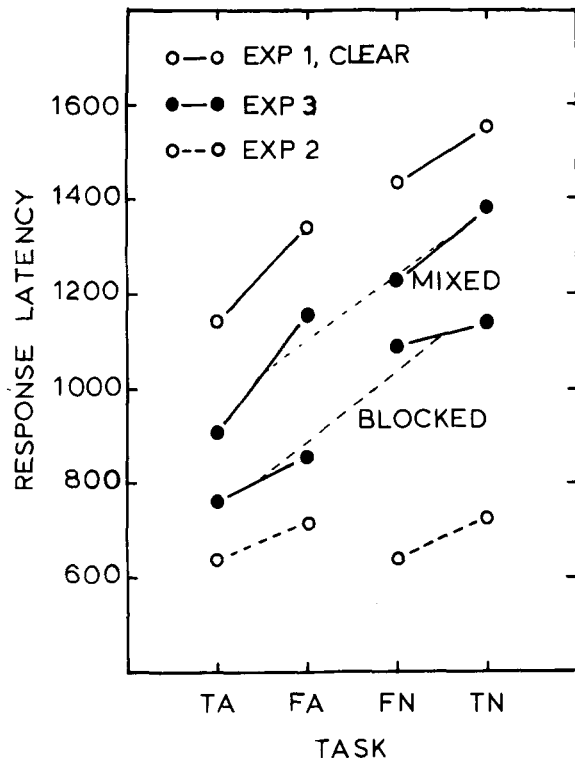| Presentation Condition	Affirmatives		Negatives	
	True	False	False	True
Mixed	2.9	4.5	11.0	14.5
Blocked	2.5	2.9	7.1	9.2

Figure 3. Response latency (in milliseconds) as a function of sentence type and truth value (TA = true affirmative; FA = false affirmative; FN = false negative; TN = true negative).

longer than the "true" reaction [F(1,29) = 9.08, MSe = 974.7].

The only two-way interaction to reach an acceptable level of significance was that of Sentence Type by Truth Value [F(1,29) = 134.6, MSe = 794.3]. This significant interaction deserves no special explanation, because it replicated the usual finding that false affirmatives take longer to verify than true affirmatives and true negatives take longer to verify than false negatives. The nonsignificant interaction of Presentation Condition by Sentence Type indicates that the amount of negation time is not sensitive to the shift from the mixed to the blocked presentation. The three-way interaction of Presentation Condition by Sentence Type by Truth Value was significant [F(1,29) = 50.98, MSe = 524.1]. This interaction was due to the degree of the interaction of Sentence Type by Truth Value being attenuated as the condition was shifted from the mixed to the blocked presentation. In other words falsification time was longer under the mixed than under the blocked presentation. A Wilcoxon test showed that the longer time taken to verify false affirmatives than to verify true affirmatives was longer under the mixed than under the blocked presentation (z = 4.6, N = 30, T = 9), and so was the time taken to verify true negatives longer than that taken to verify false negatives (z = 3.32, N = 30, T = 71).

In summary, the present experiment clearly demon-

strated a decrease of falsification time and a resistance to change in negation time from the mixed to the blocked presentation. Neither of these results can be successfully accounted for by either model. As mentioned earlier, these two models predict either (1) a change in negation time and no change in falsification time or (2) a resistance to change in both falsification time and negation time across the two presentation conditions. These predictions are simply contradictory to the present results.

It might be argued that the present interpretation of the process difference between the two presentations is invalid; the time to be saved under the blocked presentation may not be what was mentioned earlier. However, we see no other reasonable interpretation that would make the two models predict an outcome compatible with the present results. For example, it might be suggested that the time to be saved under the blocked presentation is the one to compare the inner constituents, rather than the time to compare the polarity markers. This suggestion is simply not acceptable, because one still has to compare the inner constituents even under the blocked presentation in order to make correct judgments. Or it might be suggested that under the blocked presentation, in which it is not necessary to take time to compare the polarity markers, the time taken to shift a truth index value to its contrary as a result of mismatch of two inner constituents (following the Clark & Chase, 1972, model) would also be shortened. This suggestion is not acceptable either, because it would demand that the Clark and Chase model predict a change in falsification time, as well as a change in negation time.

GENERAL DISCUSSION

Several findings of the present study suggest a need to postulate that falsification time has a source different from that of negation time. First, using a sentence-first paradigm, Experiment 1 showed a change in falsification time but no change in negation time following a change in the form in which the picture was presented: When the picture was shifted from a clear to a degraded form, the longer time taken to verify false affirmatives than to verify true affirmatives was shortened. This result was interpreted to mean that, in sentence-picture verification, encoding the sentence first produces a positive priming effect on encoding a picture that is congruent with the sentence (such as in true affirmatives) and produces a negative priming effect on encoding an incongruent picture (such as in false affirmatives). With the visual degradation of pictures, such a differential priming effect was attenuated because the identification of the picture under this condition relies more on stimulus information than on context. This interpretation not only successfully accounts for the degradation effect of Experiment 1, but also suggests an involvement

of differential picture encoding in the falsification process of sentence-picture verification. Although this interpretation failed to gain support from the results of verifying negative sentences, the failure has been attributed to a potential difficulty in directly assessing the priming effect on picture encoding in verifying negatives, because evidence has shown that mental operations involved in verifying negatives tend to work in a multiplicative rather than in an additive manner, and therefore the contribution of each operation to the verification process cannot be estimated by using the method of subtraction. On the other hand, negation time was found to be free from the variation of picture form, indicating that the negation process of sentence-picture verification does not occur at the picture encoding stage.

Second, Experiment 2 showed that a picture following a sentence was named faster when the two were incongruent. This result further supports the priming explanation of the visual degradation effect of Experiment 1, and the encoding explanation of the falsification process, as well. Third, Experiment 3 showed that errors committed in verifying sentences were largely negation errors, with few falsification errors. While this result is consistent with the iterative process of constituent comparisons proposed by Carpenter and Just (1975), the result is also consistent with the view regarding the falsification process to be independent of the negation process. And, finally, Experiment 3 also showed a decrease of falsification time and a resistance to change in negation time from the mixed to the blocked presentation of the four tasks. This result will be discussed in the next section in favor of the notion that falsification and negation are separate processes. In the next section, we propose an alternative view of sentence-picture verification, based on the present findings.

An Alternative View

Following Clark and Chase (1972), we propose that, in sentence-picture verification, both the sentence and the picture are encoded in a propositional format before they can be compared with each other. A propositional format is an abstract subject-predicate relation embedded in a polarity marker that is either positive or negative for sentences and is always positive for pictures. In the sentence-first paradigm, the process of verification starts from encoding the sentence, followed by encoding the picture. As the results of Experiments 1 and 2 suggested, encoding the picture in false affirmatives and true negatives will take Time a longer than that in true affirmatives and false negatives. This amount of time, a, is called negative priming time.

After the sentence and picture are encoded, a comparison is made between the embedded part of the sentence representation and that of the picture representation; this will result in either a match or mismatch.

It is important to note here that we do not assume that the truth index is preset at "true" at the start of the comparison of the inner constituents and then is subject to shift to its opposite whenever the two constituents mismatch. Rather, we assume that the truth index is set at a specific value only after the comparison of the inner constituents has been made: If the comparison results in a match, the truth index is set at the value of "true"; otherwise, it is set at "false." We adopt this assumption based on the reasoning that when one is required to judge whether or not a sentence is true of a picture, one is clever enough not to make any decision about the "truthness" of the sentence before comparing the two. It is also assumed that there is no difference in the time taken to compare the inner constituents among the four tasks and that the time taken to set the truth index at the value of "true" is not different from that taken to set it at "false." Once the comparison of inner constituents has been made, a response corresponding to the value of the truth index is ready to be made. However, before the response is executed, the polarity marker of the sentence is used to confirm or disconfirm the value of the truth index, and hence, the response. A "positive" marker always confirms the value of the index and allows the response to be executed. On the other hand, a "negative" marker always disconfirms the value and therefore suppresses the response to be made for the purpose of initiating its opposite. This response suppression will thus occur in verifying negatives only and will take an extra amount of time, b. The amount of Time b is called response-suppression time. It may well be argued that it will take longer to disconfirm than to confirm a value of the truth index, but, for the present purpose, it is not necessary to make such an argument. However, it is argued that it takes Time c longer to confirm or disconfirm the value "false" than to confirm or disconfirm the value "true." This amount of time is called falsification-confirmation time.

Table 3 presents the components of verification latency for the four tasks proposed by the present model, incorporating the above three parameters. The latency components not accounted for by these three parameters are assumed to be fixed across the four tasks and are collapsed into a single parameter, t_0. These latency components are the time taken to encode a sentence (either an affirmative or a negative), S, to encode a congruent picture, P, to compare the inner constituents, Ci, to confirm or disconfirm the truth index value "true," T, and to execute a response, R, among others. Thus Time t_0 is also the time taken to verify true affirmatives, which serves as base time, to which the three time parameters may be added to predict the verification latencies for the other three tasks. For example, false affirmatives take longer to verify than true affirmatives, because it takes longer to encode the picture and because it takes longer to confirm the value of the truth index by the polarity

Table 3
Process of Sentence-Picture Verification
as Proposed by the Present Model

Task	Latency Components
True Affirmatives	$t_0 = S + P + Ci + T + R$
False Affirmatives	$t_0 + a + c$
False Negatives	$t_0 + b$
True Negatives	$t_0 + a + c + b$

Note—See text for explanation of the parameters.

marker of the sentence in the former than in the latter. For another example, false negatives take longer to verify than true affirmatives, because the former involves an extra amount of time to suppress the response tendency resulting from the comparison of inner constituents. According to the present model, falsification time consists of two components: (1) negative-priming Time a and (2) falsification-confirmation Time c. On the other hand, negation time is exclusively contributed by the single parameter of response-suppression Time b. Thus, falsification and negation are regarded as separate processes, not only because they are represented with different parameters in the present model, but also because they have different sources.

Several immediate predictions by the present model are possible. First, the model will predict an increase of verification latency in the order of true affirmatives, false affirmatives, false negatives, and true negatives; when the three parameters, a, b, and c, are estimated from typical verification data, it will show the inequality $b > (a + c)$. It should be noted that, although the present model proposes a verification process different from the Clark and Chase (1972) model, as shown in Table 1, the two models involve the same number of parameters and yield the same predictive equation. Thus, the percentage of variance of data that can be described by the Clark and Chase model can also be described by the present model. Second, according to the present model, the mechanisms responsible for falsification time have to do with the delay of picture encoding and of the confirmation-disconfirmation of the truth index value; the mechanisms appear to be latency-producing, but error-free, mechanisms. On the other hand, the mechanism of response suppression, which is the sole mechanism responsible for negation time, is thought to be an error-prone mechanism, because it is not always possible to suppress successfully a response that is to be taken readily; when the suppression is not successful, it will result in an incorrect response. This accounts for the result of Experiment 3 that more negation errors than falsification errors occurred in verification.

Third, the present model can successfully account for the finding in Experiment 3 that there was a decrease of falsification time and a resistance to change in negation time from the mixed to the blocked presentation. When under the condition in which positive and negative

sentences are mixed, the subject may tend to respond "true" or "false," corresponding to the truth index value set by the result of comparing inner constituents, but such a tendency to respond should be subject to a final check by the polarity marker of the sentence. On the other hand, when under the blocked presentation, the subject bypasses this step of final check by the polarity marker of the sentence. Without any negatives, the subject can simply compare the inner constituents and initiate an appropriate response according to the result of the comparison, without consulting the polarity of the sentence. Similarly, without any positives, the subject can respond correctly, completely on the basis of the result of comparing inner constituents, but, in this case, the subject always has to suppress the response tendency suggested by the index value, in order to initiate a response opposite to this tendency. Thus, under the blocked presentation, in addition to the advantage of a saving of time taken to read sentences, the time taken to confirm or disconfirm a truth index value, which is longer for false affirmatives and true negatives than for true affirmatives and false negatives, is saved. This accounts for the results of a general decrease of verification time across the four tasks and of a decrease of falsification time from the mixed to the blocked presentation. On the other hand, since the mechanism of response suppression is still indispensable even under the blocked presentation, there is no difference in negation time between the two presentation conditions. In terms of the above interpretations, the present model also predicts that falsification time under the blocked presentation should reveal only the effect of differential picture encoding, and this seems true from a comparison of the curves representing the blocked presentation with the curves representing the results of Experiment 2, shown in Figure 3. That the former curves are much higher than the latter curves is quite reasonable, because the former were timed from the stimulus onset to the onset of a verification response, whereas the latter were timed from the picture onset, in a sequential presentation of sentence followed by picture, to the onset of naming the picture. An inspection by eye shows that these two sets of curves tend to be parallel, suggesting that the amount of falsification time under the blocked presentation is completely predictable by the differential picture naming time.

A final remark is that the present model did not consider a possible contribution of differential sentence encoding to verification; the time taken to comprehend a negative sentence was assumed in the present model to be equal to that taken to comprehend its affirmative counterpart. It may be wrong always to make this counterintuitive assumption. That a negative sentence is more difficult, and hence takes longer, to comprehend than an affirmative may be made to seem even more dramatic by considering a sentence with multiple

negatives, such as "It is not true that these dots are not red," as opposed to a simple affirmative, such as "These dots are red." It is certainly true that the former sentence is much more difficult to comprehend than the latter, and sometimes one even has to use some strategy to deal with a multinegative sentence, such as using the stragegy that two "nos" make a "yes," three "nos" make a "no," and so on. However, maintaining an argument that a simple negative is not more difficult to comprehend than a simple affirmative may not be invalid, because these two sentence types frequently appear in language, and probably with equal frequency.

REFERENCE NOTE

1. Dyer, F. N., & Severance, L. J. *Stroop interference with successive presentations of separate incongruent words and colors* (Report No. 989). Fort Knox, Ken: U.S. Army Medical Research Laboratory, June 1972.

REFERENCES

CARPENTER, P. A., & JUST, M. A. Sentence comprehension: A psycholinguistic processing model of verification. *Psychological Review*, 1975, **82**, 45-73.

CATLIN, J., & JONES, N. K. Verifying affirmative and negative sentences. *Psychological Review*, 1976, **83**, 497-501.

CLARK, H. H., & CHASE, W. G. On the process of comparing sentences against pictures. *Cognitive Psychology*, 1972, **3**, 472-517.

DYER, F. N. The duration of word naming responses: Stroop interference for different preexposures of the word. *Psychonomic Science*, 1971, **25**, 229-231.

GOUGH, P. B. The verification of sentences: The effects of delay of evidence and sentence length. *Journal of Verbal Learning and Verbal Behavior*, 1966, **5**, 492-496.

KLEIN, G. S. Semantic power measured through the interference of words with color-naming. *American Journal of Psychology*, 1964, **77**, 576-588.

MACLEOD, C. M., HUNT, E. B., & MATHEWS, N. N. Individual differences in the verification of sentence-picture relationships. *Journal of Verbal Learning and Verbal Behavior*, 1978, **17**, 493-507.

POSNER, M. I., & SNYDER, C. R. R. Attention and cognitive control. In R. L. Solso (Ed.), *Information processing and cognition: The Loyola symposium.* Hillsdale, N.J: Erlbaum, 1975. (a)

POSNER, M. I., & SNYDER, C. R. R. Facilitation and inhibition in the processing of signals. In P. M. A. Rabbitt & S. Dornic (Eds.), *Attention and performance VI.* London: Academic Press, 1975. (b)

SHOBEN, E. J. Choosing a model of sentence-picture comparison: A reply to Catlin and Jones. *Psychological Review*, 1978, **85**, 131-137.

NOTE

1. All statistical tests throughout this article were evaluated at the .05 level of significance.