Quarterly Journal of Experimental Psychology (1975) 27, 405-410

PICTORIAL ENCODING OF SENTENCES IN SENTENCE-PICTURE COMPARISON

BARBARA TVERSKY

Department of Psychology, Hebrew University, Jerusalem, Israel

Subjects are shown to verbally encode sentences for simultaneous comparison with pictures but pictorially encode sentences for later comparison with pictures. This is taken as further evidence that subjects adapt their encodings or representations of stimuli to demands of the task.

Introduction

How do we decide whether a simple sentence, such as *star isn't above plus,* is true of a simple picture, such as $\overset{*}{+}$? Such a task reveals how sentences and pictures are represented and compared. In a series of studies examining verification times in sentence-picture comparison, Clark and Chase (Chase and Clark, 1972; Clark and Chase, 1972, 1974) have presented evidence that both sentence and picture are encoded in terms of the same elementary propositions composed of names of objects along with names of their properties. The elements of the propositions are then compared one by one to produce a response.

According to Clark and Chase (1974), both linguistic and perceptual considerations determine how pictures are represented. If one object is perceived as the point of reference, the other object will be described in relation to it. For instance, if we replace the + with —, which may be perceived as the horizon, the star will be coded as above or below the line. In the absence of such considerations, linguistic evidence leads to expectation of a preference for encoding pictures in terms of *above,* which is unmarked, rather than *below.* Several experiments supported these claims, that characteristics of the picture as well as demands of the task determine how the pictures are encoded.

There is evidence from tasks requiring same-different judgments between verbal descriptions and pictures that task variables should also affect representations of sentences. When subjects expect to compare a verbal stimulus to a picture, and when they have sufficient time, they are able to construct a pictorial code of the verbal stimulus (Posner, Boies, Eichelman and Taylor, 1969; Seymour, 1974; Tversky, 1969). Comparison of a pictorially-coded stimulus with a picture is faster than comparison of a verbally-encoded stimulus to a picture (Tversky, 1969). Moreover, the subject can reduce his memory load by representing a sentence in a pictorial code to compare with a picture rather than retaining a verbal code of the sentence, deriving a verbal code of the picture, and then comparing the two verbal codes, all in memory. In the present experiment, sentences, such as *stick isn't above ball,* and pictures such as o, are presented either simultaneously, or successively for timed true–false judgments. It was predicted that ia the simultaneous condition, subjects would encode sentences verbally, as in the previous Clark and Chase (1974) studies, but that in the successive condition, subjects would encode the sentences pictorially, to expedite comparison with the picture.

How can a pictorial representation of the present stimuli be distinguished from a verbal representation? Shepard and Chipman (1970) proposed that the type of internal representation can be characterized by discovering what they termed a second-order isomorphism between the relations among internal representations and the relations among the corresponding external objects. In that paper, the similarity space obtained by judgments on outline drawings of the states was shown to be essentially isomorphic to the space obtained by shape judgments on the names of the states, indicating that the latter judgments were based on pictorial representations of the states. In a similar vein, Tversky (1969) found that under conditions encouraging verbal coding, comparison reaction times followed a pattern conforming to verbal similarities among the stimuli but that under conditions favouring pictorial encoding, reaction times were isomorphic to pictorial properties of the actual stimuli. For the present stimuli, a pictorial representation would simply reflect the spatial positions of the two objects. Neither the preposition of the sentence, nor the positivity of the sentence should affect reaction times to judge "true" or "false". Comparison times to a pictorial code should reflect, with these stimuli, only match or mismatch, since it is well documented that mismatches, or "no" or "false" responses are longer than matches (Bamber, 1969; Bindra, Donderi and Nishisato, 1968; Tversky, 1969). Comparison times to verbal codes, on the other hand, reflect linguistic properties of the sentences. For instance, codes containing a negative are responded to considerably slower than codes without, and this effect typically interacts with the truth of the comparison such that false negatives are faster than true negatives. Finally, codes containing *above* are typically processed faster than codes containing *below* as the preposition. Thus, it was expected that sentences would be verbally encoded when presented simultaneously with the comparison pictures, but pictorially coded when presented prior to picture, in order to expedite the comparison.

Method

Subjects

The subjects were 16 Hebrew-speaking students at the Hebrew University, who were either paid or received course credit for their participation in the experiment.

Stimuli

The stimuli were eight sentences, translated from Hebrew, *stick is above ball, stick is below ball, stick isn't above ball, stick isn't below ball,* plus the four sentences obtained by interchanging *stick* and *ball* and two pictures, one of a black circle on top of an elongated thick black rectangle and the other of the rectangle above the circle. Both pictures and sentences were constructed from Letraset characters, photographed, and presented on slides.

Procedure

There were 16 trials per block, formed from combining each of the eight sentences with each of the two pictures. Each subject participated in 11 blocks, each a different random order of the 16 trials; the first block was discarded as practice.

In the simultaneous condition, the sentence was presented simultaneously with and to the right of the picture (Hebrew is read from right to left) until the subject responded. The subjects were instructed to look first at the sentence, then the picture, and to respond as quickly as possible without making errors. In the successive condition, the sentence was displayed for 5 s followed by a 5-s delay, followed by the picture, which remained in view until the subject responded. These subjects were also instructed to respond rapidly without making errors. In both cases, the procedure and conditions of the task were carefully explained to the subjects, but no suggestions as to how to encode the stimuli were given.

Subjects initiated each trial by pressing a foot lever, and responded by pressing one of two keys. Half the subjects in each condition responded "true" with their right index finger, and half with their left index finger. The experiment lasted about 25–40 min.

Results

Separate analyses of variance were performed on the latencies of correct responses from the simultaneous condition and from the successive condition. The factors were true–false, positive-negative, above–below, and stick-ball (as subject of the sentence). Differences in the pattern of results between the two conditions are indicative of differential encoding. Errors, which constituted 7·6% of the simultaneous condition data and 5·6% of the successive condition data, were eliminated from all analyses.

In the simultaneous condition, positive sentences (sentences without a *not*) were responded to significantly faster than negative ($F = 19{\cdot}39$, $df = 1, 7$, $P < 0{\cdot}01$), and positively interacted significantly with truth of sentence ($F = 20{\cdot}66$, $df = 1, 7$, $P < 0{\cdot}01$), such that while false positive sentences were slower than true positive sentences, false negatives were faster than true negatives. The effect of positivity was 938 ms and of the interaction, 230 ms. These are to be contrasted to effects of at most 50 ms for the remaining factors, none of which was significant. Mean reaction times for the various sentence types are displayed in Table I.

TABLE I

Mean latencies in ms for the four true and four fake sentence types for simultaneous presentation of sentence and picture

Positive	True	above	1537	True	
		below	1590	1564	Positive
	False	above	1762	False	1699
		below	1904	1833	
Negative	True	above	2772	True	
		below	2693	2733	Negative
	False	above	2501	False	2637
		below	2582	2542	

All of these results were also obtained by Clark and Chase (1974) in their third experiment *in* which sentence and picture were presented simultaneously and subjects were instructed to read the sentence before looking at the picture. In addition, Clark and Chase's subjects took 80 ms longer to encode the line of the picture prior to the star, a perceptual effect on picture encoding. In the present experiment, the two stimuli were equally salient perceptually, and, according to the reaction time data, equally easy to encode. Clark and Chase also found that sentences containing *below* took longer to encode than sentences containing *above*, an effect on verbal encoding which was not replicated in Hebrew.

In the successive condition, the only significant finding was that "true" responses were 137 *ms* faster than "false" responses $(F = 10{\cdot}91, df = 1, 7, F < 0{\cdot}05)$. The effects of positivity, preposition and sentence subject were each less than 20 ms in magnitude and far from significant. The overall average reaction time was 808 ms, a large improvement from the 2168 ms obtained in the simultaneous condition, Mean reaction times for the various sentence types are displayed in Table II. Most of the subjects ia the successive condition were informally questioned after the experiment as to how they performed the task. Most reported that they imagined what figures would appear in each position. Two subjects reported integrating the two figures into a single one, e.g. "o" into a triangle and "o" into an upside down triangle.

TABLE II

Mean latencies in ms for the four true and four false sentence types for successive presentation of sentence and picture

True	Positive	above	727	Positive	True
		below	726	727	
	Negative	above	778	Negative	740
		below	726	753	
False	Positive	above	875	Positive	
		below	893	884	False
	Negative	above	893	Negative	877
		below	845	869	

Discussion

The hypothesis that sentences would be verbally encoded for simultaneous comparison to a picture, but pictorially encoded for comparison to a picture presented later was clearly supported by the data. When given sufficient time, subjects construct or generate a pictorial code from a verbal one in accordance with demands of the task. The evidence supporting this conclusion is the finding of entirely different patterns of reaction times under simultaneous and successive conditions. Under the simultaneous condition verbal properties of the stimuli affected the reaction times; sentences containing *not* took longer to process than positive sentences and this factor interacted with the truth of the sentence. Under successive presentation, only the match between sentence and picture affected reaction times; mismatches took longer to process. The former result is expected under verbal encoding, while the latter is expected under pictorial encoding.

It is particularly interesting that pictorial encoding appeared only when sufficient time was allotted, and not otherwise. Time is commonly presumed to allow forgetting or replacement of visual codes by verbal ones (e.g. Sperling, 1963). Here, it seems that the very existence of the memory constraint encouraged construction of a pictorial code, for at least two, complementary reasons. First, picture-picture matches are faster than name-picture matches (Seymour, 1974; Tversky, 1969, 1974). Second, there is evidence that both pictures and pictorial codes are more resistant to forgetting than words or verbal codes. Pictures are better recognized than sentences and words (Shepard, 1967) and words and sentences that have been imaged or that are easy to image are better retained than verbal material that has not been imaged or that is abstract (Bower, 1972; Jorgensen and Kintsch, 1973; Paivio, 1971). Thus, the interval between the sentence and the picture could be utilized in order to speed comparison time and to reduce memory load by the same operation, conversion of a verbal code to a pictorial one.

The present experiment is consistent with experiments, cited in the introduction, that demonstrate construction or generation of pictorial codes from words or simple descriptions. The results are also consistent with experiments demonstrating differential encoding of stimuli in anticipation of retrieval task (Carey and Lockhart, 1974; Frost, 1972; Tversky, 1973). In a now-classic article, Craik and Lockhart (1972), point out that encoding in memory can be carried out to various depths, from sensory or perceptual levels to meaningful levels. The depth to which processing is continued, in turn, determines the the memorability of the items. The present results, along with others demonstrating flexibility of encoding, can be used to modify Craik and Lockhart's analysis. Not only can encoding be carried on to various levels, but it can be carried on in alternative modalities. Demands of the task, as well as characteristics of the stimuli, affect choice of encoding modality; and the modality selected is revealed in performance of the task.

This research was supported by a grant from the Human Development Institute of the Hebrew University, by funds provided by the School for Overseas Students of the Hebrew University, and by the Batsheva de Rothschild Fund. The author thanks Varda Sacks for helpful suggestions at every phase of the study and for faithfully executing the experiments and to Tracy Sherman for writing the computer program. The author is grateful to Herbert H. Clark and an anonymous reviewer for helpful comments on a previous version of this paper.

References

BAMBER, D. (1969). Reaction times and error rates for "same"–"different" judgments of multidimensional stimuli. *Perception and Psychophysics*, **6**, 169–74.

BINDRA, D., DONDERI, D. C. and NISHISATO, S. (1968). Decision latencies of "same" and "different" judgments. *Perception and Psychophysics*, **3**, 121–30.

BOWER, G. H. (1972). Mental imagery and associative learning. In Lee W. GREGG (Ed.), *Cognition in Learning and Memory*. New York: Wiley.

CAREY, S. T. and LOCKHART, R. S. (1973). Encoding differences in recognition and recall. *Memory and Cognition*, **1**, 297–300.

CHASE, W. G. and CLARK, H. H. (1972). Mental operations in the comparison of sentences and pictures. In Lee W. GREGG (Ed.), *Cognition in Learning and Memory*. New York: Wiley.

CLARK, H. H. and CHASE, W. G. (1972). On the process of comparing sentences against pictures. *Cognitive Psychology,* 3, 472–517.

CLARK, H. H. and CHASE, W. G. (1974). Perceptual coding strategies in the formation and verification of descriptions. *Memory and Cognition, 2,* 101–111.

CRAIK, F. I. M. and LOCKHART, R. S. (1972). Levels of processing: a framework for memory research. *Journal of Verbal Learning and Verbal Behavior,* 11, 671–84.

FROST, N. (1972). Encoding and retrieval in visual memory tasks. *Journal of Experimental Psychology,* 95, 317–26.

JORGENSEN, C. C. and KINTSCH, W. (1973). The role of imagery in the evaluation of sentences. *Cognitive Psychology,* 4, 110–16.

PAIVIO, A. (1971). *Imagery and Verbal Processes.* New York: Holt, Rinehart and Winston.

POSNER, M. I., BOIES, S. J., EICHELMAN, W. H. and TAYLOR, R. L. (1969). Retention of visual and name codes of single letters. *Journal of Experimental Psychology,* 79, (1, Pt. 2).

SHEPARD, R. N. (1967). Recognition memory for words, sentences and pictures. *Journal of Verbal Learning and Verbal Behavior,* 6, 156–63.

SHEPARD, R. N. and CHIPMAN, S. (1970). Second-order isomorphism of internal representations: shapes of states. *Cognitive Psychology,* 1, 1–17.

SPERLING, G. A. (1963). A model for visual memory tasks. *Human Factors,* 5, 19–31.

SEYMOUR, P. H. K. (1974). Pictorial codings of verbal descriptions. *Quarterly Journal of Experimental Psychology,* 26, 39–51.

TVERSKY, B. (1969). Pictorial and verbal encoding in a short-term memory task. *Perception and Psychophysics, 6,* 225–33.

TVERSKY, B. (1973). Encoding processes in recognition and recall. *Cognitive Psychology, S*, 275-87.

TVERSKY, B. (1974). Breadth of pictorial and verbal codes in memory. *Bulletin of the Psychonomic Society,* 4, 65–8.

Received 5 *August* 1974