DOCUMENT RESUME

ED 121 119 FL 007 637

AUTHOR	Just, Marcel Adam; Carpenter, Patricia A.
TITLE	Eye Fixations and Cognitive Processes.
INSTITUTION	Carnegie-Mellon Univ., Pittsburgh, Pa. Dept. of Psychology.
SPONS AGENCY	National Inst. of Education (DHEW), Washington, D.C.; National Inst. of Mental Health (DHEW), Bethesda, Md.
PUB DATE	Aug 75
GRANT	NIE-G-74-0016
NOTE	71p.; For related documents, see FL 007 636, FL 007 632 and EJ 112 015
EDRS PRICE	MF-$0.83 HC-$3.50 Plus Postage
DESCRIPTORS	*Cognitive Measurement; *Cognitive Processes; *Eye Fixations; Eye Movements; Eyes; Memory; Neurolinguistics; Psycholinguistics; *Psychological Studies; Reading Processes; Task Performance; Verbal Learning; Visual Perception

ABSTRACT

This paper reports on a study concerned with rapid mental operations of the central processor as it performs tasks such as the comparison of rotated figures, mental arithmetic, sentence verification, and memory scanning. The central processor is the site of most of the symbol manipulation that takes place in the human information processing system. One of the goals of the study is to demonstrate that the locus, duration and sequence of eye fixations are all closely tied to the activity of the central processor. The research is intended to discover how eye fixations are related to cognitive processes, the aim being to construct a theory of that relationship that generalizes across a number of task environments. The second goal of the research is to examine how the eye fixation data reveals the fine structure of the processor's activity in performing a number of cognitive tasks. (Author/CLK)

DEPARTMENT

OF

PSYCHOLOGY

Carnegie-Mellon University

EYE FIXATIONS AND
COGNITIVE PROCESSES[1]

Marcel Adam Just

and

Patricia A. Carpenter

Psychology Department
Carnegie-Mellon University
Pittsburgh, Pennsylvania 15213

Complex Information Processing

(CIP) Paper #296

August, 1975

Abstract

This paper presents a theoretical account of the sequence and duration of eye fixation during a number of simple cognitive tasks, such as mental rotation, sentence verification, and quantitative comparison. In each case, the eye fixation behavior is linked to a processing model for the task by assuming that the eye fixates the referent of the symbol being operated on.

A widely accepted view of the human information processing system is that most of the symbol manipulation takes place in a central processor, sometimes referred to as the active memory (Neisser, 1967), working memory (Newell & Simon, 1963), operational memory (Posner, 1967), or the immediate processor (Newell, 1973). This paper is concerned with the rapid mental operations of the central processor, and how they are reflected by the pattern and duration of eye fixations during a task involving visual input. We will examine the basic operators, parameters, and control structure of the central processor as it performs such tasks as the comparison of rotated figures (Shepard & Metzler, 1971), mental arithmetic (Parkman, 1971), sentence verification (Carpenter & Just, 1975), and memory scanning (Sternberg, 1969). These tasks generally take less than 5 or 10 sec to complete, and can be decomposed into very rapid mental operations, often estimated to consume between 50 to 800 msec each. The goals of this paper are (1) to demonstrate that the locus, duration, and sequence of eye fixations are all closely tied to the activity of the central processor, and (2) to examine how the eye fixation data reveals the fine structure of the processor's activity in performing a number of cognitive tasks.

Our proposal is that the eye fixates the referent of the symbol currently being processed if the referent is in view. That is, the fixation may reflect what is at the "top of the stack." If a number of symbols are processed in a particular sequence, then their referents should be fixated in the same sequence, and the duration of fixation on each referent may indicate how long the corresponding symbol is operated on. The obvious advantage of monitoring eye fixations is that the behavior within any particular trial can potentially be decomposed into various stages whose durations can be directly measured. By contrast, a single response latency cannot be interpreted or decomposed without reference to latencies in other conditions.

Another reason that eye fixations provide an appropriate measure in cognitive tasks is that fixation behavior can be sampled at high densities per unit time. A high sampling rate (say once every 200 msec) is necessary and appropriate when studying mental operations whose durations range from tens to hundreds of msec, because the durations of individual stages of processing (and hence changes in the duration) can be measured directly. The relation between duration of processes and sampling rate can be elucidated with an analogy to time-lapse photographs of slow or rapid processes. If one is studying the behavior of glaciers, it is sufficient to take a photograph once every few weeks; but to study the blossoming of a flower, one might want hourly photographs. Similarly, to study the rapid mental operations of the central processor, it is desireable to monitor its behavior many times per trial, so as to separate the behavior into stages. The trace of the stages may provide a specification of their respective durations and the sequence in which they occur.

Eye fixation studies have their historical roots in cognitive research dealing with reading. Almost one hundred years ago in 1878, Javal (cited by Mackworth, 1974) observed young children's eyes during reading, and contrary to the then popular conception of a continuous sweep across a line of print, he discovered that the eye made a series of discrete pauses separated by jumps. While some research pursued the role of eye fixations in reading (cf. Buswell, 1922; Dearborn, 1906; Huey, 1908; Woodworth, 1938), much of the subsequent psychological research focused on the jumps (saccades) rather than the pauses (cf., Alpern, 1962; Ditchburn, 1973; Yarbus, 1967, for overviews), and the behaviors that were investigated were oculo-motor rather than cognitive. Recently, there has been renewed research interest in the pauses themselves, and how they relate to underlying cognitive processes (cf., Tichomirov & Posnyanskaya, 1966; Winikoff, 1967). The current paper will examine eye fixations in several situations and account for the locus, sequence, and duration of eye fixations in terms of their relationship to underlying cognitive processes.

We will examine only tasks in which the subject must encode some information from a visual display, do some mental computations, and then produce a response that is contingent on the outcome of the computations. These tasks are well structured in that the subjects' goals are clear to them and to the experimenter. We will not concern ourselves with tasks that require subjects simply to read or scan a display without any specified purpose or response. We will analyze speeded tasks, in which the subject is asked to work accurately but quickly. Then, we will show how the total response latencies can be divided into processing stages on the basis of the locus and sequence of fixations.

The purpose of analyzing several tasks is to abstract the general characteristics of the central processor as they are revealed by eye fixation behavior. Generally, research programs and resulting papers revolve around a particular task, such as mental rotation, sentence comprehension, or memory scanning, attempting to discover or characterize the operations used in that particular task. Our goals here are slightly different. While we surely want to learn about the fine structure of the processes used in each task, we also wish to discover how eye fixations are related to cognitive processes. The aim is to construct a theory of that relationship that generalizes across a number of task environments.

PROCESSING ROTATED FIGURES

Eye fixations are intimately involved with our ability to visually encode spatially distributed information. We were interested in whether eye fixations would also indicate how visual information is internally manipulated. We chose for our task domain the "mental rotation" studies, in which people compare two figures in order to determine whether or not they depict the same three-dimensional object (Shepard & Metzler, 1971). In these studies, subjects were timed while they decided whether two figures were views of the same object (Fig. 1a or 1b), or views of different objects. The two objects in the Different trials (Fig. Ic) differed by a reflection (as well as by a rotation). The main independent variable was the angular disparity between the two views of the <u>same</u> object, that is, the amount of physical rotation necessary to align the two figures into congruence. The response latencies for the Same trials increased linearly with the degree of angular disparity. Shepard and Metzler attributed this increase in response time to a process of mental rotation. The slope of the response times as a function of the angular disparity was postulated to reflect the rate of mental rotation.

Insert Figure I about here

There are several key questions about the processes underlying performance in this task that are not easily answered by the response latency studies. We proposed that the following questions about the microstructure of the processes could be addressed by an eye fixation study.

1. How does the subject know which parts of the figure are to be rotated into each other? Before rotating one figure into another, the subject must decide which parts potentially correspond to each other. Eye fixations may indicate how this initial decision about correspondence is made.

2. How does the subject know how far to rotate one of the objects? One possibility would be that the subject makes some estimate of the angular disparity, and then performs a ballistic rotation (i.e., with the target orientation predetermined). Alternatively, the rotation

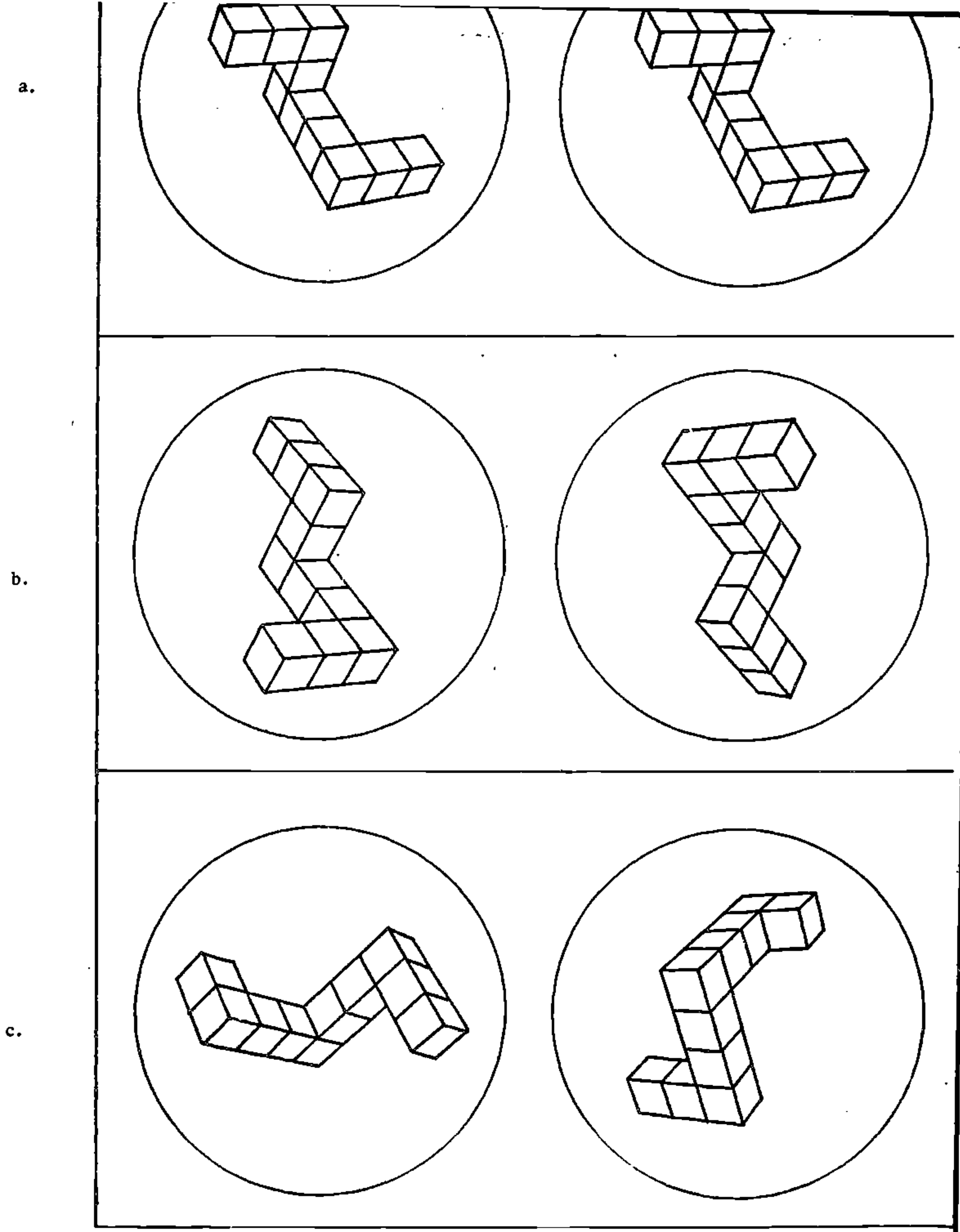

Figure 1(a) A pair of Same figures with 0° disparity;
(b) a pair of Same figures with 180° disparity;
(c) a pair of Different figures with 120° "disparity."

process may be monitored at various points along the way. The eye fixations may show whether the process is monitored.

3. Once the required rotation has been performed, how does the subject know whether the two figures represent the same object or not? The eye fixation behavior may reveal the comparison process that determines whether the two figures match or not after rotation.

We can examine these issues in some detail, by assuming that subjects fixate the external referent of the internal symbol being processed. Our objective was to identify component processes in this task by analyzing the scan paths and by observing how they changed with angular disparity. In a pilot eye-fixation experiment subjects compared two figures with different orientations in the picture plane, the plane perpendicular to the subjects' line-of-sight. The results suggested that there were three stages in the processing that we will call (1) search, (2)transformation and comparison, and (3)confirmation.

In the first stage, there is a search for segments of the two figures that superficially correspond to each other, for example, two segments at the end of the figures that both have three visible faces. The function of the search process is to select segments of the two figures that can potentially be transformed one into the other. During the next stage, transformation and comparison, the two corresponding segments are rotated into each other. A transform-and-compare operation is applied stepwise to the representations of the two segments. Each step of the transformation may correspond to a rotation, such that at the end of the transformation, the segment is represented at a new orientation. Each step of the transformation is followed by a comparison to determine whether the two orientations are now congruent. This stepwise transform-and-compare process continues until the necessary number of transformations have been made to make the internal representations of the two segments sufficiently congruent in orientation. The third stage, confirmation, involves a check of whether the rotation that brought the two segments into congruence will also bring other portions of the two figures into congruence. Processes roughly similar to search, transformation, and confirmation have been suggested by Metzler and Shepard, and their subjects' introspective reports supported the suggestions (1974, p. 169 and 178). The eye fixation data allow us to separate the performance into the three stages, and specify the nature of the processing within each stage.

Method. The experiment was a Same-Different task in which the subject was timed and her eye fixations recorded while she decided whether two figures depicted the same object or two objects that were mirror images of each other. The stimuli were three drawings shown on the left-hand side in Figure 1 plus their mirror images, for a total of six basic figures. In the Same trials, the left-hand figure could be rotated clockwise 180° or less in the picture plane to bring it into complete congruence with the right-hand figure. The amount of rotation necessary to bring the two figures into congruence varied from 0° to 180° in steps of 20°, for a total of 10 possible angular disparities. To construct a Different pair, the right-hand figure of a Same pair was replaced by its mirror image isomer. The mirror image figure was constructed by reflecting the original figure through a plane in three-dimensional space (see Metzler & Shepard, 1974). There was a Same and a Different pair for each of the 6 basic figures at each of the 10 angular disparities, for a total of 120 pairs of stimulus figures.

The two figures were displayed side by side, with the left-hand figure randomly assigned to one of three orientations. The center-to-center distance between the figures was 15.5 cm, and each figure was between 10 and 10.5 cm wide. The stimulus pairs were displayed on a standard video monitor at a viewing distance that was set for each subject at between 53 and 68 cm so so as to keep the horizontal and vertical excursion of the eye spot constant. Thus each figure subtended about 10° of visual angle and the center-to-center angle was about 15°, depending on the viewing distance.

The eye spot was electronically superimposed on a picture of the stimulus display, and the composite was recorded on videotape for later scoring. Subjects initiated a trial by fixating a point in the middle of the left-hand side of the screen and pushing a "ready" button. The eye spot was calibrated with respect to this fixation point before each trial. After calibration, the fixation point disappeared and half a second later, the stimulus appeared. The subject responded Same or Different by pressing one of two microswitches with the index and third finger of her dominant hand. The stimulus presentation and timing of the response were controlled by a DDP-116 computer. Head movements were minimized by using a bite bar and head rest. The 120 stimuli were presented in a random order, and distributed over two testing sessions, separated by at least one day. The subjects received 60 practice trials before the experiment began. The three paid subjects were right-handed females of college age, with 20-20 corrected vision. Five other subjects were eliminated because they made more than 15% errors during the 60 practice trials.

The locus of the eye spot, relative to the ten cubes that made up each figure, was scored on each frame of the videotape, namely once every 16.7 msec. When the eye spot was located on the same cubes in a sequence of successive video frames for at least 100 msec, the frames were aggregated into a single fixation. The occasional corrective eye movement that occurs after a saccade (cf. Bartz, 1967) resulted in some very short fixations. These were aggregated into the main fixation if the fixations were on adjacent cubes of the figure.

Results. The mean response latencies for correct Same trials, shown in Figure 2, increased monotonically with increasing angular disparity. All three subjects showed a linear increase between 0° and 100°, but the curves were positively accelerated beyond 100°. The mean latencies from 0° to 100° disparity have a pattern similar to that obtained by Metzler and Shepard (1974).

Insert Figure 2 about here

Eye fixation results. One striking feature of the eye fixation behavior was that subjects systematically looked back and forth between the left and right figure.[2] For example, at 0° disparity, subjects initially fixated the left figure, then looked over at the right hand figure, then looked back at the left, and frequently looked back at the right hand figure for a second time, for a total of three switches between the two figures. The mean number of such switches between figures increased with angular disparity, as shown in Figure 3.

Insert Figure 3 about here

The next step of the analysis was designed to determine exactly what subjects were looking at, and how the pattern of their fixations might reveal the microstructure of the underlying cognitive operations. To classify the locus of the eye fixation, we divided each figure into three main segments: the arm whose third face of the end cube was visible (open), the arm whose third face of the end cube was not visible (closed), and a central angle. For example, in Figure 1a, the upper arm will be called the open arm, while the lower arm will be called the closed, and the four central cubes constitute the central angle. The locus of the eye spot was scored according to the locus of its centroid with respect to the three segments.

The simplest way to describe our scoring procedure is to apply it to a

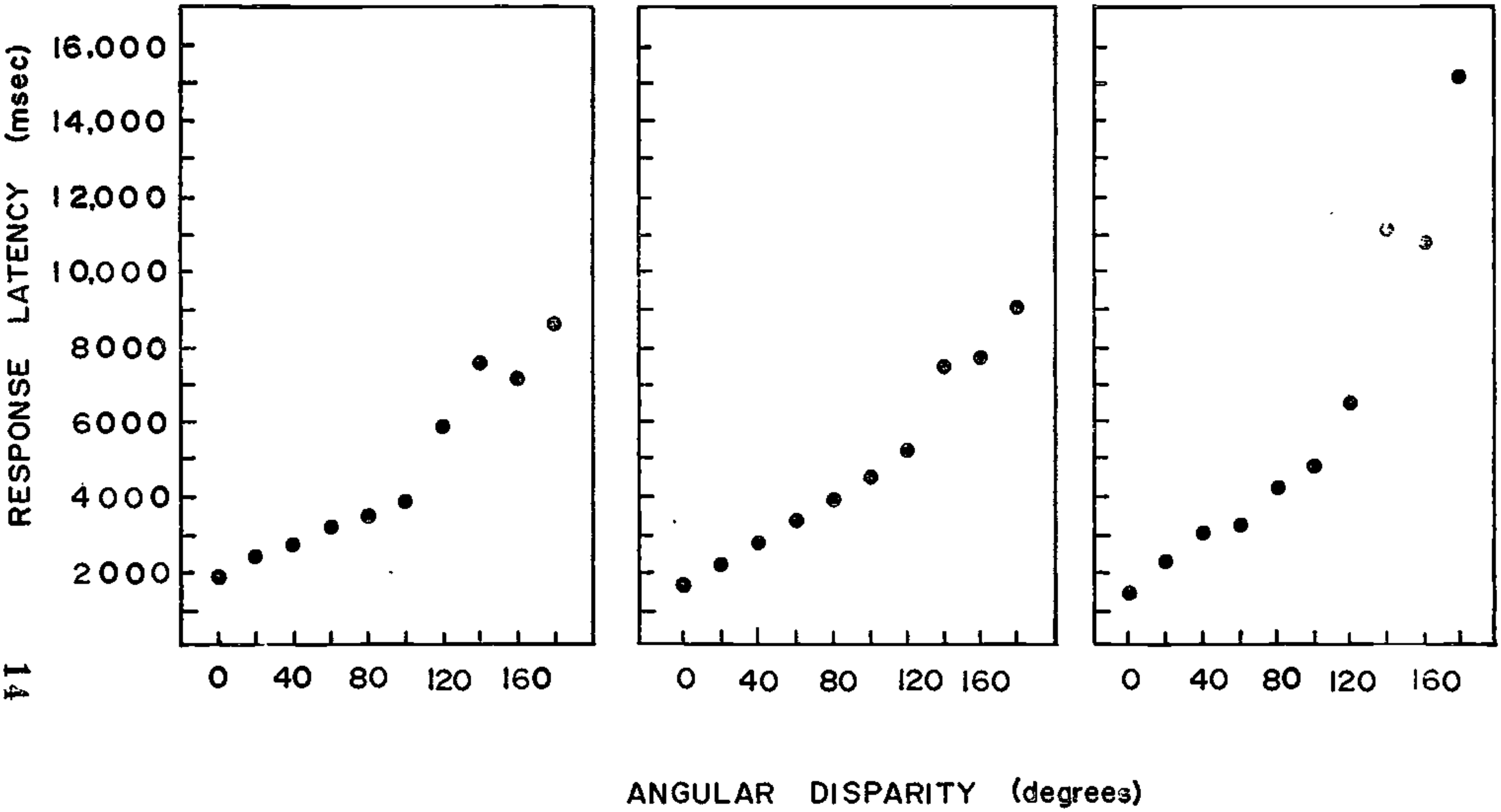

Figure 2. Mean response latency for Same trials as a function of angular disparity for the 3 subjects.

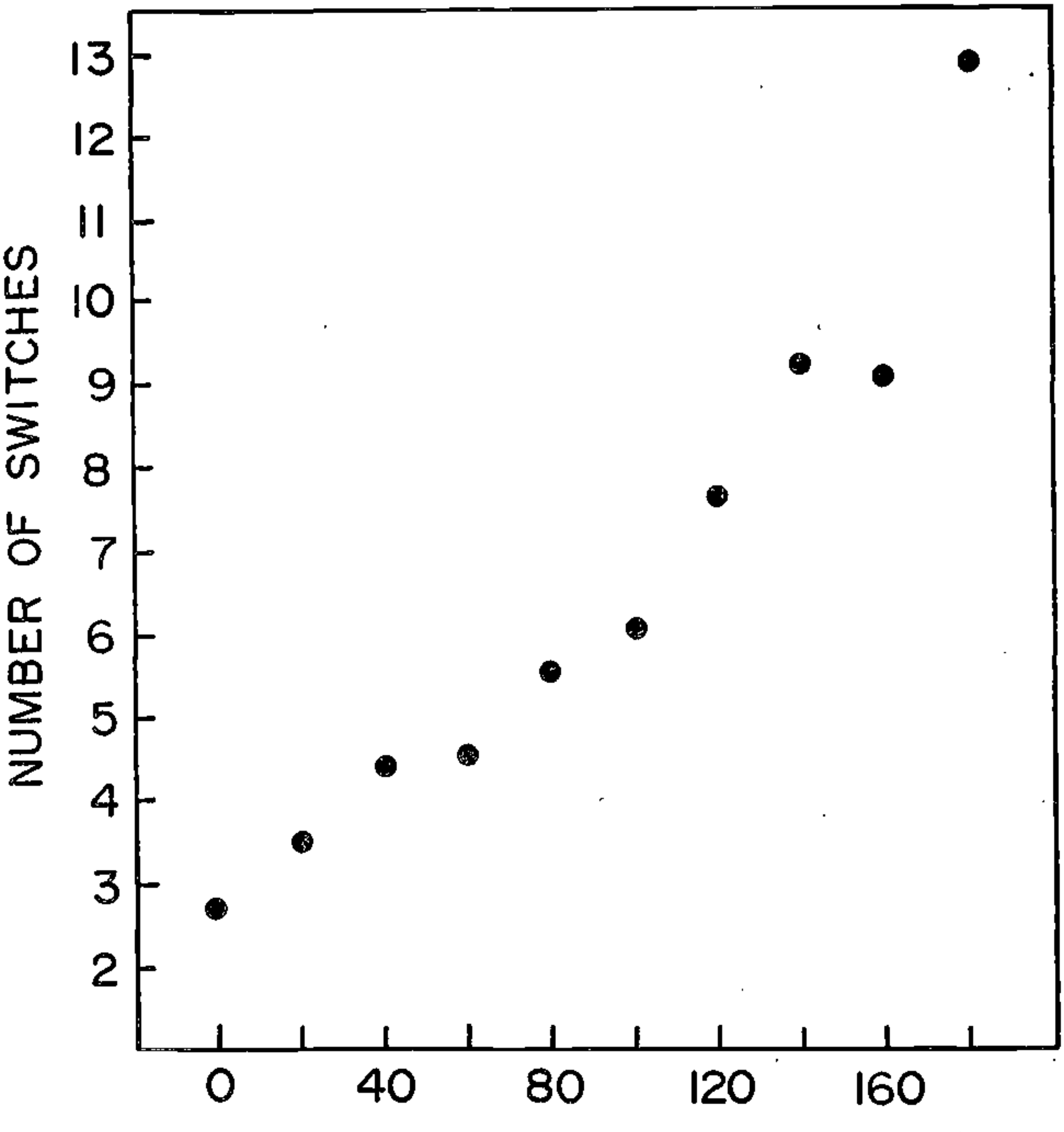

Figure 3. Mean number of switches for Same trials as a function of angular disparity.

few representative scan paths. Figure 4 shows a scan path for a Same trial with 0° disparity. After the initial fixation on the center of the left figure, the subject fixated corresponding closed arms at the upper part of each figure. Then the open arms at the bottom of each figure were fixated.

Insert Figure 4 about here

To make our analysis of the scan paths precise, we constructed rules for classifying instances of search, transformation and comparison, and confirmation. The most prominent property of the scan paths was that the subject would repeatedly look back and forth between corresponding segments of the two figures without any intervening fixations on other segments. We identified the repeated fixation of corresponding segments with the transformation and comparison process. When the same pair of segments was involved in two transformation episodes separated by other fixations, their durations were combined. The transformation and comparison process is evident in fixations 5 to 8 of the scan path shown in Figure 5, where the figures have an 80° disparity. In fixations 5 to 8, the subject looked back and forth between the closed arms of the two figures, for a total of 1185 msec.

Insert Figure 5 about here

We identified the search process with the initial portion of the scan path that preceded the first instance of transformation. Applying these rules to the scan path in Figure 5, fixations 1 to 4 would be attributed to search, for a total of 818 msec. In Figure 4, where the angular disparity is much smaller, the duration of the search process (351 msec, fixation 1) is much shorter.

We identified the third process, confirmation, as a short sequence of fixations between corresponding parts of the two figures other than the trans-

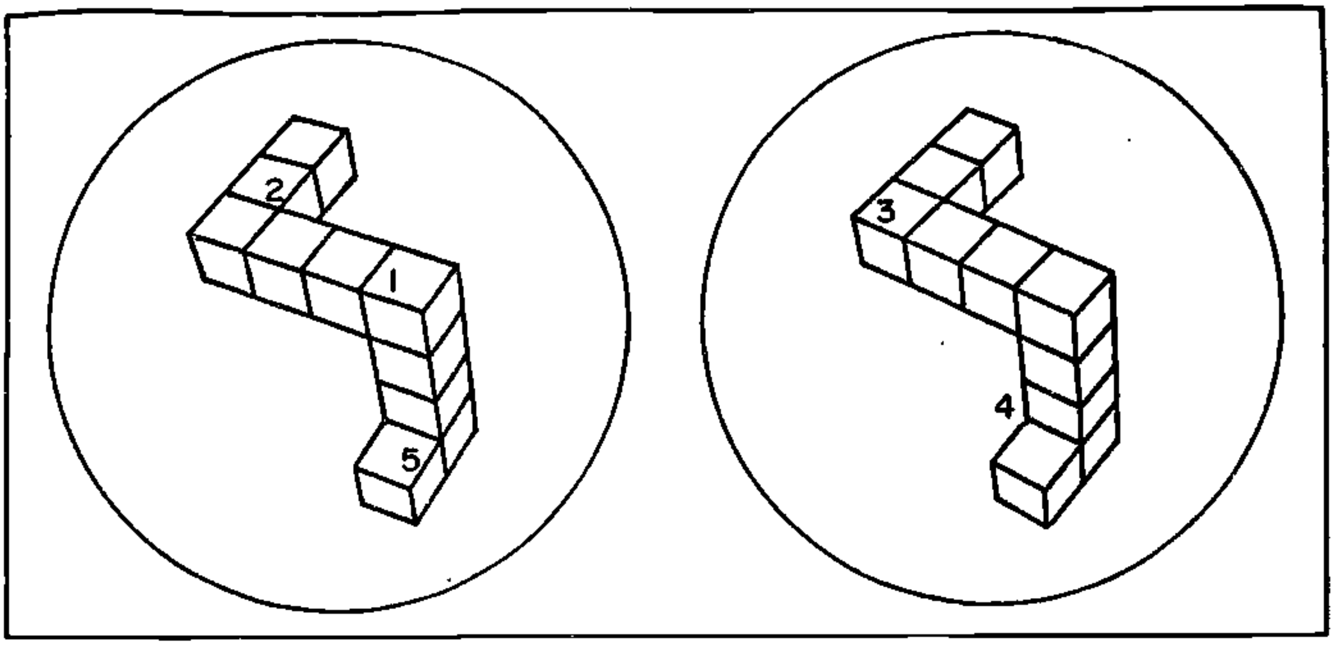

Figure 4. The figure indicates the sequence of fixations on a correct Same trial in which the disparity was 0°. The subject's total response latency was 1296 msec, of which 11% had no visible eye spot. The locus and duration of the fixations are as follows:

Fixation	Figure	Location	Duration
1.	Left	Center	351 msec
2.		Closed arm	150 msec
3.	Right	Closed arm	200 msec
4.		Open arm	200 msec
5.	Left	Open arm	250 msec

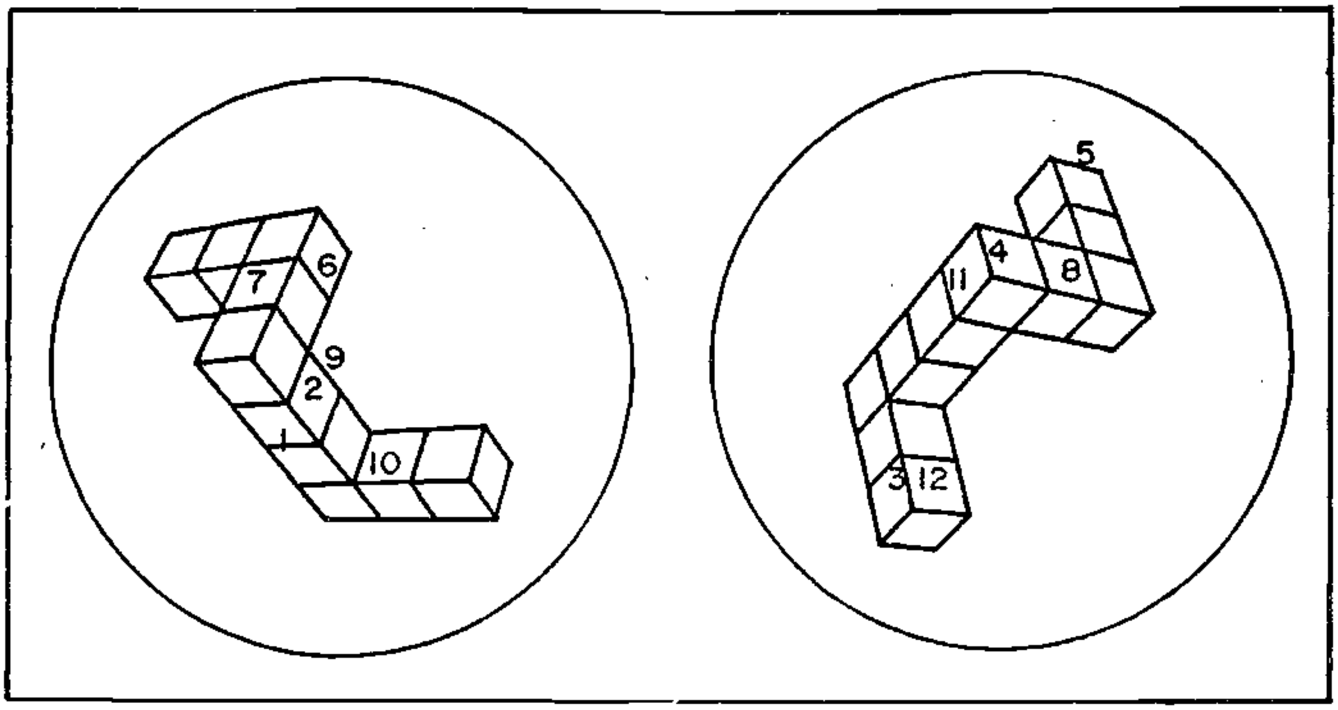

Figure 5. The figure indicates the sequence of fixations on a correct Same trial in which the disparity was 80°. The subject's total response latency was 3574 msec, of which 9% had no visible eye spot. The locus and duration of the fixations are as follows:

Fixation	Figure	Location	Duration
1.	Left	Center	200 msec
2.		Center	301 msec
3.	Right	Open arm	167 msec
4.		Center	150 msec
5.		Closed arm	167 msec
6.	Left	Closed arm	200 msec
7.		Closed arm	317 msec
8.	Right	Closed arm	501 msec
9.	Left	Center	250 msec
10.		Open arm	200 msec
11.	Right	Center	484 msec
12.		Open arm	317 msec

formed segments. Confirmation could appear as a scan from the central angle to an arm on one figure, then a similar scan on the other figure. Figure 5 shows an example of the confirmation process where the fixations proceed from the central angle to the open arm on the left figure (fixations 9 and 10) and then a similar scan is executed on the right (fixations 11 and 12). In the scan path in Figure 4, the last two fixations on the open arms (fixations 4 and 5) also exemplify confirmation.

To see how well the model fits the eye fixation data, we scored the scan paths of 100 of the 171 correct Same trials. Seventy-one trials were not scored because of apparatus failure, in which the optical system failed to capture an eye spot that was visible at least 85% of the time. The mean response latencies from the 100 trial sample are very similar to the data for all 171 correct Same trials, as shown in Figure 6, so the sample appears to be representative.

Insert Figure 6 about here

The analysis of the scan paths allows us to examine how the total processing time is distributed across search, transformation and comparison, and confirmation stages as a function of angular disparity. As Figure 7a shows, the time spent initially searching the figures increased with angular disparity, from about 300 msec at 0° to about 1600 msec at 180°. The bulk of the processing time was spent in transformation and comparison, as shown in Figure 7b. The duration of this stage increased markedly with increasing angular disparity, from about 500 msec at 0° to 3800 msec at 180°. The average time spent in the third stage, confirmation, increased from 450 msec at 0° to 2300 msec at 180°, as shown in Figure 7c. Thus, for a typical trial, say at 80° disparity, 21% of the time was consumed by initial search, 39% by transformation and comparison, and 26% by confirmation. The remaining 14%

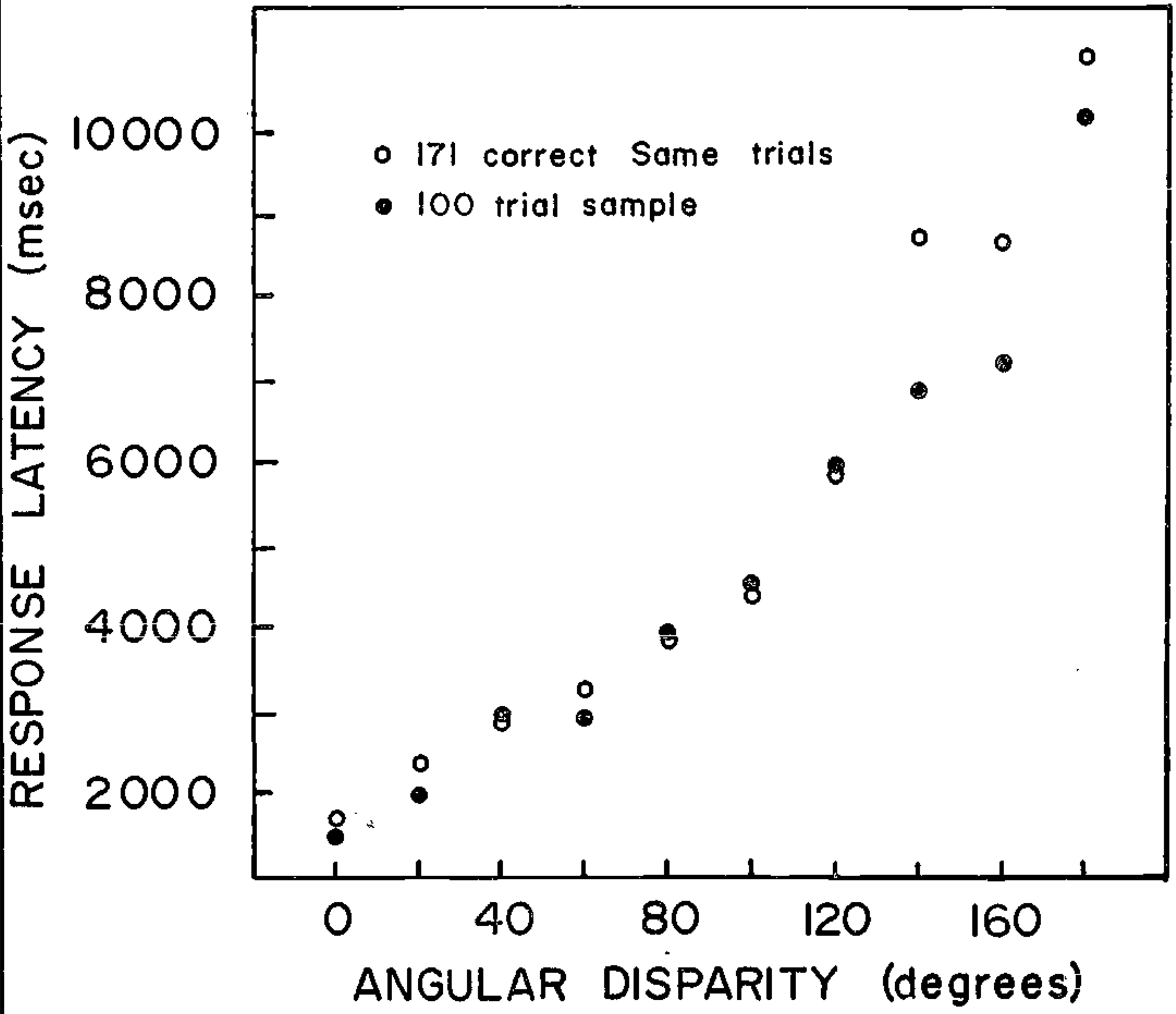

Figure 6. Mean response latency as a function of angular disparity for all correct Same trials and for 100 correct Same trials in which eye fixations were scored.

was distributed between saccades (about 10%) and remaining "other" fixations (about 4%) that did not fit any of the three categories.

Insert Figure 7 about here

Even though the duration of transformation did not increase linearly, the total response time did so at least between 0° and 100°. It is easy to see in Figure 7 (panels a through e) how the durations of several monotone functions cumulate to produce a linear increase in total latencies between 0° and 100°, shown in Figure 6. Some of the non-linearity beyond 100° is unaccounted for, even by the eye fixation analysis, and has been classified as "other" eye fixation behavior. This miscellaneous behavior may result from the subject losing track of her place in the process, or attempting to start over. Occurrences of this behavior inflated the total response times, particularly at 180°.

We mentioned earlier that a striking feature of the scan paths was that the subject repeatedly looked back and forth between the two figures, and the number of such switches between the figures increased with angular disparity. We measured how these switches were distributed across the three stages of processing, as shown in Table 1. The number of switches associated with the search stage remained quite low (usually one or less) at all disparities. Most of the switches occurred during the transform and compare process, during which the number of switches increased monotonically with angular disparity. The switching data from this stage will play a key role in the development of our model. Finally, the switches during confirmation increased with angular disparity, but not as much as for transformation. The classification procedure also categorizes the switches that occur during the transition from one stage to another, and the number of such switches remains fairly constant across angular disparities.

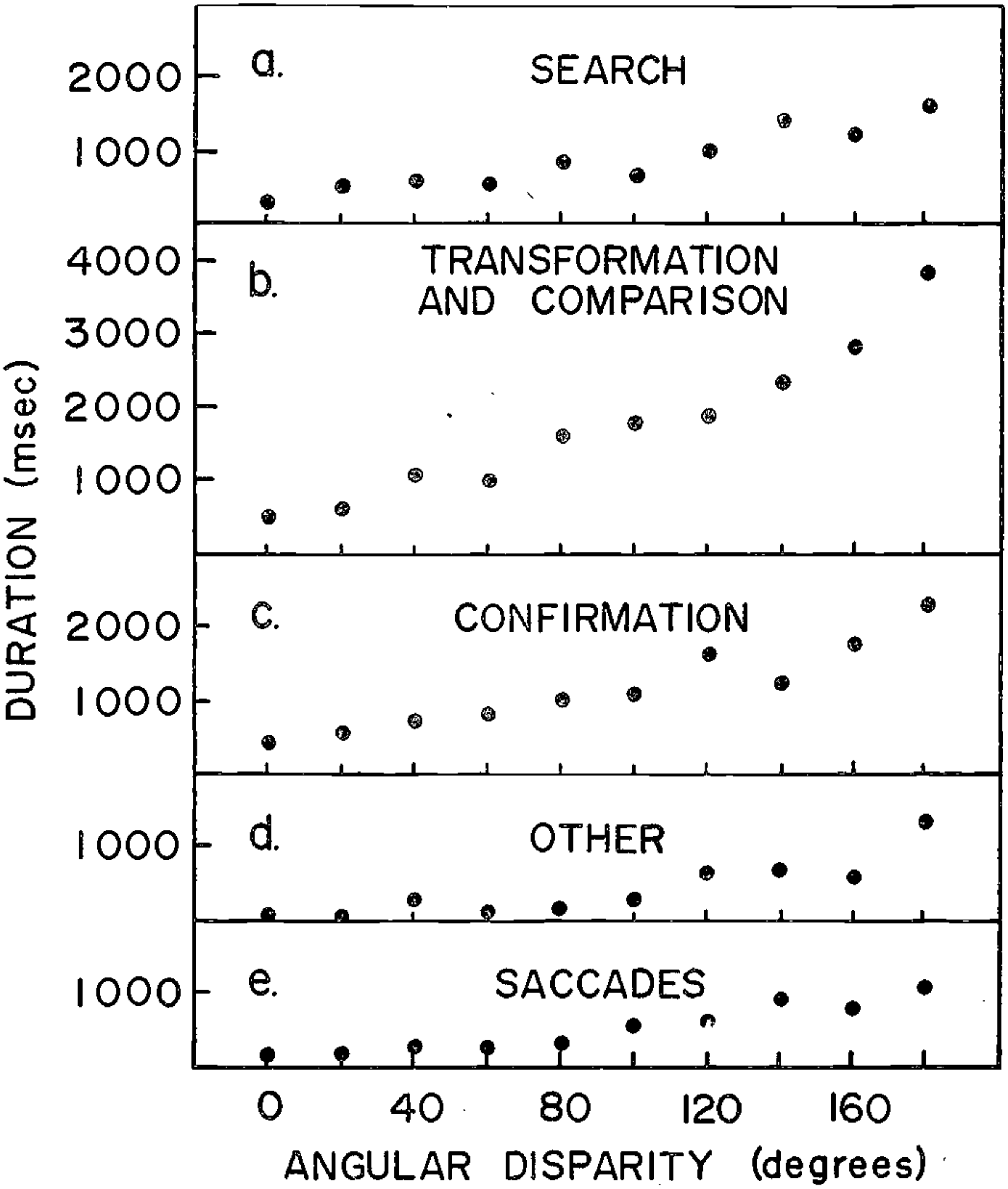

Figure 7. Mean duration of various processing stages in Same trials as a function of angular disparity.

Insert Table 1 about here

As might be expected, the average number of fixations increased with angular disparity, from 6 fixations at 0°, to 31 at 180°. Also, the average duration of a fixation increased from 200 msec at 0° to over 320 msec at 180°.

Incorrect Same trials. Error trials have often been ignored by chronometric models of cognitive processes because it is difficult to attribute errors to a particular stage of processing (one exception is the work on the speed-accuracy trade-off, cf., Wickelgren, 1974). An incorrect response in the rotation task could result from an error during any one of the stages of searching, transforming and comparing, confirming, or in executing the final motor response. An example of a transformation error would be to rotate a segment about the wrong axis and incorrectly conclude that two Same figures represent different objects. The total response latency alone provides insufficient information to localize the error on a particular trial to a particular stage. However, the eye fixations do provide clues about the reasons for some of the errors. There was a total of nine errors on the Same trials, all on angular disparities greater than 120°. On five of the nine trials, the subject attempted to transform non-corresponding segments. That is, the initial search process selected two segments that were in fact not corresponding. The subsequent transformation and confirmation stages failed to detect this error. The scan path in Figure 8 demonstrates this type of error in which the subject erroneously selected the open arm on the left and closed arm on the right as corresponding, then looked back and forth between them in fixations 2 to 8 and 10 to 12, and attempted confirmation in fixations 13 to 21.

Insert Figure 8 about here

In the remaining four error trials, subjects did successfully complete

Table 1

Distribution of switches in 100 trial sample

	Mean Number of Switches during:					
Angular Disparity	Initial Search	Transformation and Comparison	Confirmation	Transition Between Stages	Switches Not Accounted For	Total
0°	0.0	1.0	0.9	0.6	0.2	2.7
20°	0.4	1.1	0.9	0.8	0.1	3.3
40°	0.2	1.7	1.2	0.9	0.2	4.2
60°	0.3	1.7	1.5	0.7	0.4	4.6
80°	1.0	3.0	1.5	0.4	0.3	6.2
100°	0.5	2.7	1.5	0.8	0.5	6.2
120°	1.1	2.9	2.1	0.6	1.2	7.9
140°	2.2	3.6	1.8	0.7	0.7	8.9
160°	1.3	4.0	2.2	0.8	0.9	9.2
180°	1.6	5.7	2.3	0.8	1.8	12.2

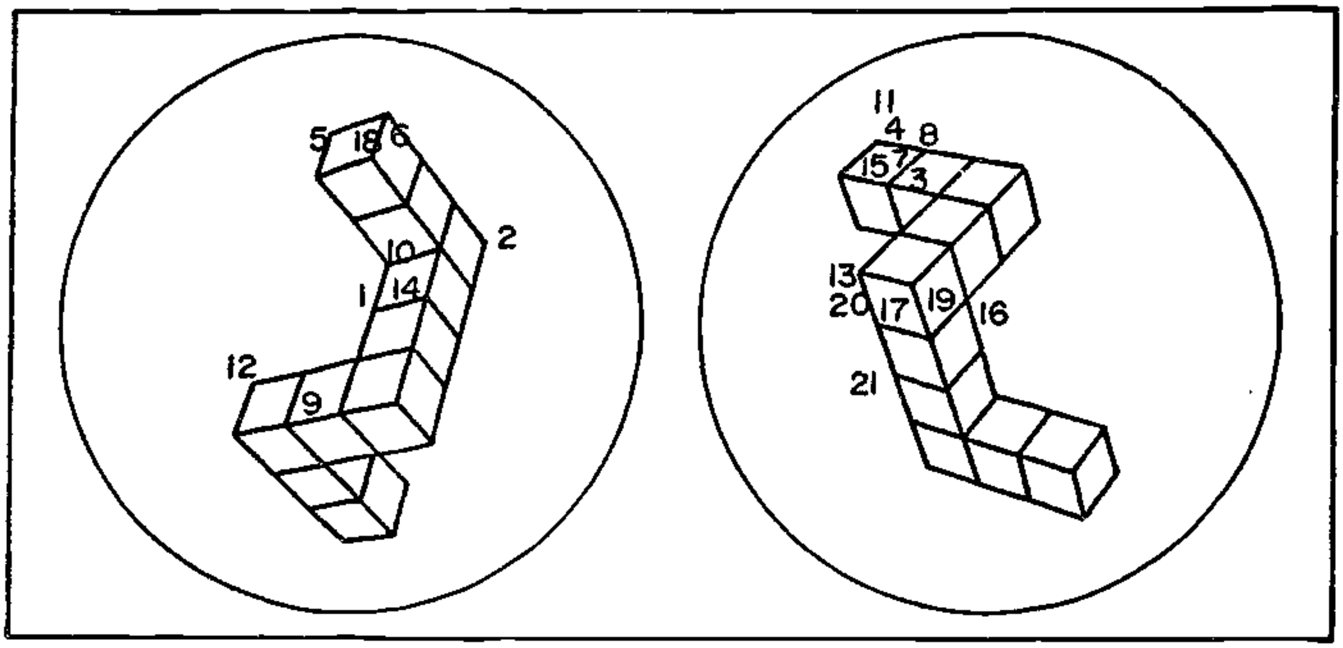

Figure 8. The figure indicates the sequence of fixations on an incorrect Same trial in which the disparity was 140°. The subject's total response latency was 8567 msec, of which 13% had no visible eye spot. The locus and duration of the fixations are as follows:

Fixation	Figure	Location	Duration
1.	Left	Center	334 msec
2.		Open arm	134 msec
3.	Right	Closed arm	200 msec
4.		Closed arm	200 msec
5.	Left	Open arm	468 msec
6.		Open arm	317 msec
7.	Right	Closed arm	200 msec
8.		Closed arm	334 msec
9.	Left	Center	334 msec
10.		Open arm	117 msec
11.	Right	Closed arm	401 msec
12.	Left	Closed arm	150 msec
13.	Right	Center	150 msec
14.	Left	Center	418 msec
15.	Right	Closed arm	251 msec
16.		Center	200 msec
17.		Center	568 msec
18.	Left	Open arm	768 msec
19.	Right	Center	450 msec
20.		Center	902 msec
21.		Center	534 msec

the initial search process and subsequent fixations alternated between corresponding segments of the two figures. This suggests that the source of the error must have occurred in some subsequent stage such as the transformation, confirmation, or response execution.

Different trials. The response latencies for Different trials were long (an average of 4 sec longer than Same trials) and variable. The angular disparity between two figures is not really well defined for a Different trial, since the two figures cannot be physically rotated into congruence. The total response latencies alone give no indication of how processing time was distributed across the three stages. However, the pattern of eye fixations allow us to follow the sequence of processing stages and to determine which stages consume the extra 4 sec of processing.

The scan paths indicate that the initial search process in Different trials starts out similarly to Same trials. However, in a Different trial, the segments selected by the search stage cannot be in complete correspondence. For example, in the Different pair shown in Figure 1c, the short arm in the left figure is closed while the short arm in the right figure is open. So subjects must select a pair on the basis of length or openness. In all seven Different trials involving stimulus pair 1c that we analyzed, the initial selection was based on the feature of length. In Different trials involving the objects depicted in Figures 1a and 1b, the two open arms have the same length, but differ in the way they are joined to the center. In two of the three analyzed trials involving these objects, the transformation was between open arms. In the third case, it was between arms that were similarly joined to the center.

The confirmation process is extremely important in the Different trials, since it leads to the discovery that the inter-segment relations are not the same in the two figures and hence that the figures are different. In fact, one of the most prominent features of the Different scan paths is the large amount of confirmation behavior that they contain. In the ten analyzed Different trials, the confirmation process consumed an average of 4195 msec, or 49% of the total duration.

The prolongation of the confirmation process is not the only reason for the very long response latencies for Different trials. On some trials, after going through a complete search-transform-and-unsuccessfully-confirm sequence, subjects make a second attempt at searching, transforming and confirming a different pair of segments. Occasionally, a lengthy search stage involved an examination of all the possible ways of pairing the segments, and that kind of search led directly to a response of Different, without any transformation. Thus, the durations of all three stages increased during Different trials, but the duration of confirmation increased the most.

One scan path that exemplifies the processing on Different trials is shown in Figure 9. Fixations 1 to 4 reflect the initial search for corresponding segments, consuming 1436 msec. Then, there is a transformation and comparison of the short arms of each figure in fixations 5, 6, and 7, consuming 919 msec. Fixations 8 to 15 reflect the confirmation process, consuming 3175 msec. We presume that it is during confirmation that the subject determined that the relation between the arm and central angles was different in the two figures. In this trial, the bulk of the processing time was consumed by the confirmation stage.

Insert Figure 9 about here

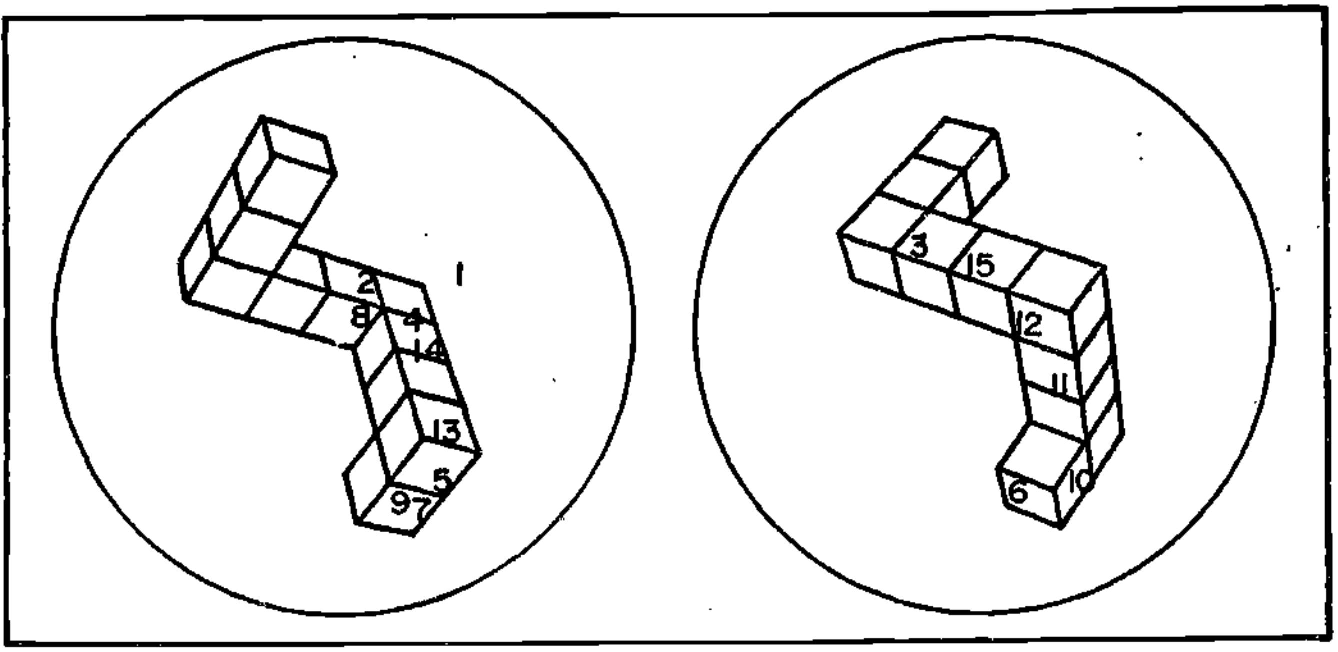

Figure 9. The figure indicates the sequence of fixations on a correct Different trial. The subject's total response latency was 5868 msec of which 6% had no visible eye spot. The locus and duration of the fixations are as follows:

Fixation	Figure	Location	Duration
1.	Left	Center	234 msec
2.		Center	518 msec
3.	Right	Closed arm	367 msec
4.	Left	Center	317 msec
5.		Closed arm	184 msec
6.	Right	Open arm	434 msec
7.	Left	Closed arm	301 msec
8.		Center	251 msec
9.		Closed arm	217 msec
10.	Right	Open arm	635 msec
11.		Center	518 msec
12.		Center	234 msec
13.	Left	Closed arm	585 msec
14.		Center	167 msec
15.	Right	Center	568 msec

The Processing Model of the Rotation Task

The internal representation. We propose that the processor operates on one segment of the figure at a time, and that the representation of the segment is schematic. The representation must include information about the segment's absolute orientation in space, as well as some defining feature such as its length or whether it is a closed or open arm. This information can be efficiently represented as the vector formed by the major axis of the segment. Moreover, if the vector has its initial point at the origin of the reference frame, then the segment can be represented by the spherical coordinates of the end point of the vector. For example, an open arm might be represented (OPEN ($\underline{r}$, θ, ϕ)) where $\underline{r}$ is the length of the segment, and θ and ϕ define the orientation of the segment.

The initial search process. The scan paths indicate that the search for corresponding segments uses a simple heuristic. Once a segment of one figure has been identified, then the search for the corresponding segment starts in the corresponding location of the other field. For example, if the long arm is in the upper-right hand corner of the left field, then the search for the corresponding segment begins in the upper-right hand of the right field. If there is no segment in the upper right, then the segment nearest the upper right is examined. The duration of this search process increases with angular disparity for two reasons. First, with increasing disparity, the corresponding segments are in successively more dissimilar locations. At 0° disparity, corresponding segments have identical locations in their respective fields. However, as the disparity increases from 0°, absolute location is a successively poorer cue for finding corresponding segments, and the heuristic must be supplemented by an active search. The second reason for the increase is that at larger disparities, the probability of selecting and attempting to transform non-corresponding segments increases and this incorrect transformation is counted as part of initial search. Figure 8 shows an example of the search process selecting non-corresponding segments that are both at the top of their respective fields. In this trial, the incorrect search led to an error. On other trials, the incorrect selection of a pair of segments was detected after some transformation had been attempted. Thus, the eye fixations allow us to trace the initial search for corresponding features and to determine the reason for the increase in the duration of the search process with angular disparity.

The transformation and comparison process. The eye fixation data also allow us to formulate a precise model of the transformation process. We propose that rotations are executed and monitored in discrete steps of approximately 50°. The estimate of the 50° step size is based on the result that there is one additional switch during the transformation stage for each additional increment of 50° in angular disparity, as shown in Figure 10.[3] A transformation may consist of applying a rotation rule that alters the representation of the orientation of a segment by 50°. For example, an open arm represented as (OPEN ($\underline{r}$, θ, ϕ)) might be transformed into (OPEN ($\underline{r}$, $\theta + 50°$, ϕ)). It is assumed that the representations of the two segments are rotated towards each other by applying the 50° rotation rules first to one segment and then to the other, until they are within 25° of each other. This form of representation and transformation does not impose any great computational burden, in contrast to a truly analogue, holistic representation of the entire figure rotated by a parallel computation of the position of all its points.

Insert Figure 10 about here

This model of the transformation stage is most easily explained by working through an example, say when two Same figures have an angular disparity

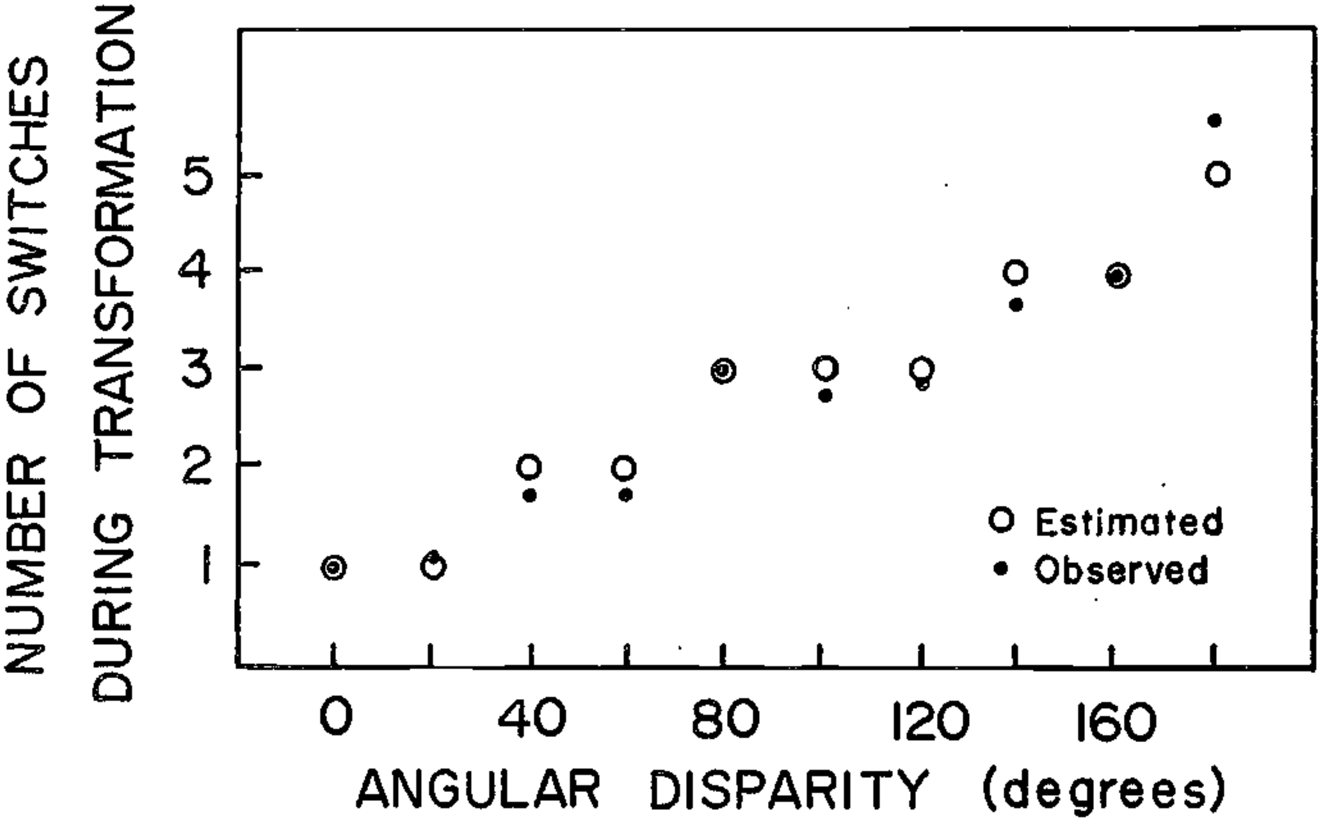

Figure 10. Mean number of observed and estimated switches during the transformation and comparison stage in Same trials as a function of angular disparity.

of 80°. Suppose that the subject has encoded a particular segment of the left-hand figure. The first switch occurs when she fixates and encodes the corresponding segment of the right figure. Then, the orientations of the two segments are compared. The orientations differ by more than 25°; therefore, she rotates her representation of the right-hand figure by 50° counterclockwise. After this transformation, she retrieves the representation of the segment on the left. In doing this retrieval, she switches her fixation to that figure. After it is retrieved, she compares the two orientations. They would still be more than 25° apart. Therefore, she transforms the representation of the orientation of the left-hand figure by 50° clockwise. Then she retrieves the representation of the segment on the right in order to compare the two again. In doing this retrieval, she switches fixation over to the right figure. After it is retrieved, she compares the two orientations. At last, after three switches and two applications of the rotation rule, the two segments are represented at fairly similar orientations (within 25° of each other). The subject would then continue on to the confirmation process.

This model can be easily summarized. After the subject has initially encoded one segment on a figure, she proceeds through the following steps:

1. Retrieve the representation of the corresponding segment on the other figure (and switch fixation to other figure).
2. Compare the two orientations. Are they less then 25° apart?
   a. No. Transform the currently fixated segment by 50° in the direction of the other figure. Go back to Step I.
   b. Yes. Go on to the confirmation stage.

The model assumes a very close relationship between eye fixations and mental operations during the transformation process. The rotation rule is always applied to the arm that is being fixated. Applying a rotation rule to the representation of one arm causes the representation of the other arm to be pushed down in the short-term memory stack. When the representation is being retrieved to the top of the stack, the other arm is fixated anew (Step 1). According to this model, the number of switches during transformation should increase monotonically with the angular disparity, but the increase should be in the form of a particular step function. There should be one switch between 0° and 25°, two switches between 25° and 75°, three switches between 75° and 125°, and so on. Figure 10 shows that this estimated number of switches corresponds very closely to the observed number of switches. The increase in switches is similar to the pattern obtained for the duration of the transformation process, shown in Figure 7b. The data in Figure 7b suggest that each step of the transform and compare process takes about 800 msec. Thus, the model of the transformation stage gives a good account of the data.

The confirmation process. Rotating two segments into similar orientations during the transformation stage is not sufficient for deciding whether the figures are the Same or Different. Therefore, the third stage, confirmation, determines whether segments other than the transformed ones correspond to each other. The scan paths indicated two methods for confirming such correspondence. One method applies the same sequence of rotation rules used in the transformation stage to another pair of segments. If this second rotation is successful, then the two figures are the Same. This method, used on about half the trials, was identified by switches of fixation between a pair of segments other than the transformed pair. A second method encodes the relation between the central angle and an arm of each figure and determines whether that relation is the same in both figures. This method, used on the other half of the trials, was identified by a scan from the center to the arm of one figure and then a similar scan of the other figure (see Figure 5 for an example). The combination of these two methods would produce a confirmation duration that increases with angular disparity, but the increase would be at a slower rate than for the transformation duration (shown in Figure 7). Either method of confirmation can determine the response of Same or Different.

Discussion. The eye fixation data lead to a detailed model of the processing in the Shepard-Metzler task, but there are questions about the generalizeability of the model. Without examining a broad range of experimental situations, there is no way of knowing which aspects of the model are invariants of the human processing system and which aspects are task-induced. Consider the proposed 50° rotation steps. It is possible that the 50° steps are fundamental and invariant over tasks. Alternatively, people may be able to tune the size of the rotation step to the particular grain and range of orientation differences they are faced with in an experiment. This is a clear empirical question of whether the rotation operation adapts itself to the task environment. Similarly, one can consider whether the representation of the figures is the same in all rotation tasks. The representations proposed in the current model are highly schematic, but they do contain sufficient information to perform the task. The representations might be more complex in tasks that demand that more information be encoded from the figures. Just as eye fixation analyses led to a precise model for the Shepard-Metzler task, this methodology should also distinguish the invariant from the transient processes, and so lead to a general theory of mental rotation.

Mental rotation has been described as an analogue process (Metzler & Shepard, 1974) in the sense that rotating a representation from one orientation to another requires that the representation pass through some internal states that correspond to intermediate orientations. Although Metzler and Shepard (1974) "...are willing to concede that sufficiently small differences may be handled in discrete jumps," the 50° step that we propose is rather large. It is possible that within each 50° step there are intermediate stages corresponding to intermediate orientations. But even with 50° steps, a 150° rotation involves intermediate stages corresponding to 50° and 100° rotations, and so the proposed process is in a loose sense analogue.

In summary, the scan paths enabled us to separate the processing into search, transformation, and confirmation stages and to measure the duration of each stage. Switches in fixation during the transformation stage indicated that the rotation was monitored in steps of approximately 50°. This analysis was applicable not only to the correct Same trials, but also provided evidence on error trials and Different trials. The research shows how eye fixations can reveal the sequence of mental operations during the processing of spatial information.

COMPARING SENTENCES WITH PICTURES

Under the assumption that visual gaze reflects underlying mental processes, we can study how people verify whether a sentence is true or false of an accompanying picture. In particular, we can determine how the total processing time is allocated among the various stages of verification. Reaction-time studies of sentence verification show that people make more errors and take longer to respond when verifying a negative sentence. The extra processing time for a negative lies between 300 and 1200 msec, depending on the linguistic structure of the negative sentence (Carpenter & Just, 1975). The processing stages involved in verification include reading the sentence and internally representing it, coding the picture, and comparing the representation of the picture to the information from the sentence (Carpenter

& Just, 1975; Chase & Clark, 1972; Clark & Chase, 1972; Trabasso, Rollins & Shaughnessy, 1971). The extra processing time for negative sentences could be consumed in any of these stages. One purpose of the following study was to monitor eye fixations in order to determine which stage of processing consumes the extra time due to negation.

Elsewhere, we have developed a model of how the information in a sentence might be represented and compared to information from a picture in sentence verification tasks (Carpenter & Just, 1975). The model suggests that elements in the sentence representation are compared sequentially to elements encoded from the picture. Mismatches between elements result in additional comparisons, thereby consuming additional processing time. The model postulates that because of the number and nature of the mismatches, the number of comparison operations increases linearly from the case of true affirmative sentences, to false affirmatives, to false negatives, to true negatives. In fact, the verification latencies in a number of studies have been found to increase linearly--corresponding to the increasing number of postulated comparisons. The current experiment examined which parts of the display were fixated longer during the conditions with longer response latencies. The primary goal was to determine how the location and duration of the gaze was related to the proposed stages of processing.

An important innovation in the current methodology was that the display was made contingent on the locus of the gaze. The only part of the display (either the sentence or the picture) that was visible to the subject was the part at the locus of the gaze, as depicted in the schematic diagram in Figure 11. This gaze-contingent display creates a functional "tunnel vision" in the subject by eliminating all peripheral information relevant to the true-false decision. The subject could not encode new information unless he looked at the relevant position in the display.

Insert Figure 11 about here

Method. The sentences in the experiment were either affirmative, like Is North, or negative, Isn't North, and involved one of the four directions, North, South, East, or West. The subject was told the phrase always referred to the location of a plus and to consider it to mean "The plus is North" or "The plus isn't North." The picture contained a plus at one of the four compass directions, and a star at the other three. (Any one of these characters, as well as the sentence, was displayed only when the subject directly fixated it). When an affirmative sentence was true, or a negative sentence was false, the plus was at the place specified by the directional term in the sentence. In the false affirmative and true negative cases, the plus could have been at any one of the three remaining locations. This design was adopted to discourage subjects from recoding negatives like Isn't North into corresponding affirmatives, like Is South. The analysis, however, is concerned only with the cases where the plus was located on the same axis as the directional term in the sentence. The sentence, centered on the video monitor, was 5.6 mm high and 45 mm wide (50 mm for negative sentences). The plus and stars were 5 mm by 5 mm, and they were at a distance of 75 mm from the center of the screen. Subject's eyes were 64 cm from the monitor, on average; however, the distance was adjusted for each subject to keep the excursion of the eye spot constant.

For scoring purposes, the viewing field was divided into an imaginary three-by-three grid, such that the sentence was located in the center square, while the stars and plus were in the middle top, middle bottom, middle left, or middle right squares. Any single fixation or sequence of fixations on one of these squares was scored as gaze on that location. During a trial, the digitizer determined the locus of the eye spot every 16 msec. Sixteen msec

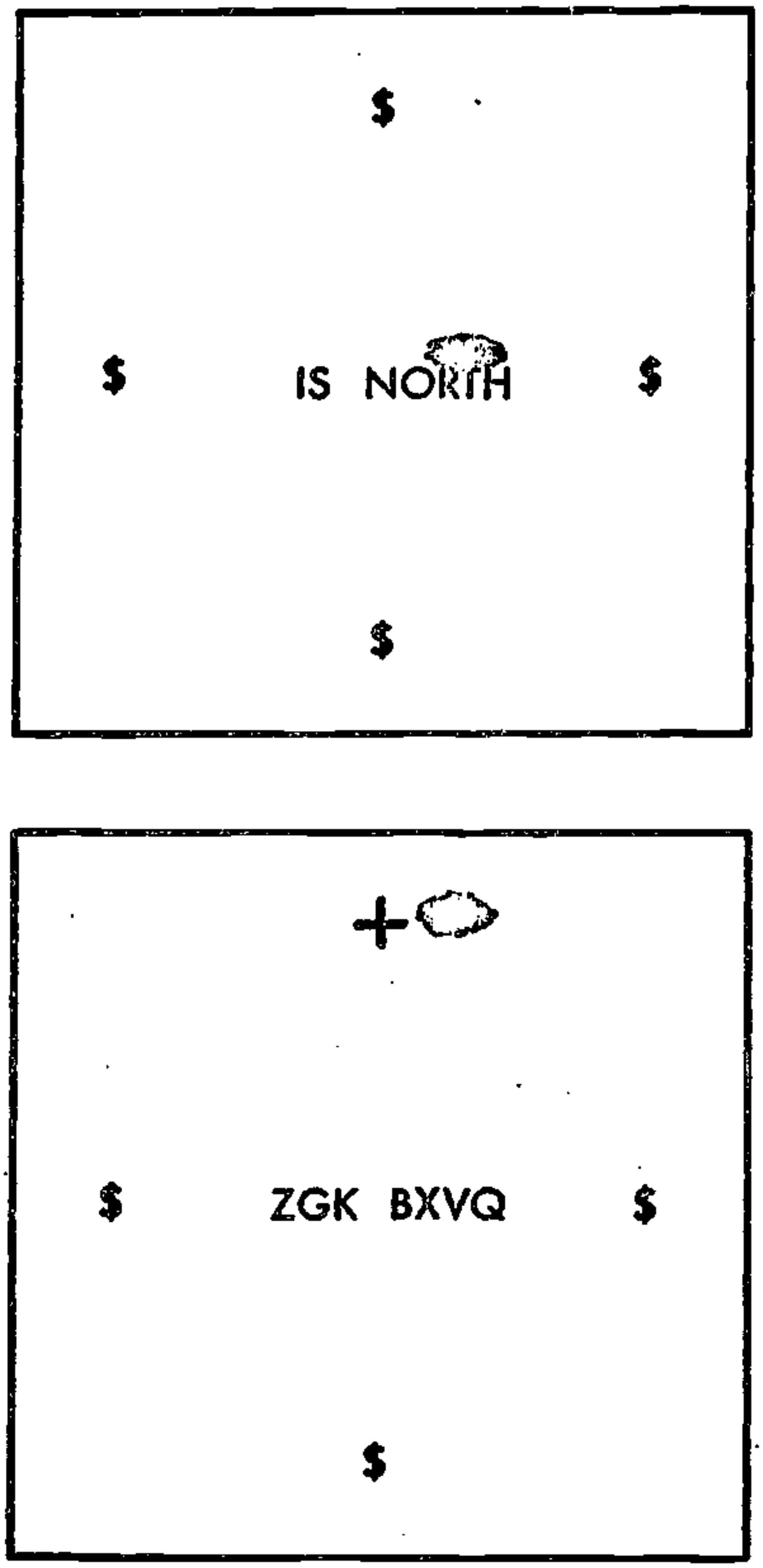

Figure 11(a) Schematic diagram of the visual display in the sentence verification task when the eye spot (denoted by black spot) is on the sentence (not to scale);
(b) visual display within the same trial when the eye spot is in the North location.

after the eye spot was first detected in a square, the stimulus material for that square appeared on the screen. As soon as the eye spot moved from that square, the stimulus was replaced by a place holder. The place holders for the sentence location were random letters like "ZGK BXVQ". The place holder for each star and plus was a dollar sign, "$". The rapidity of the replacement, within 16 msec after the initial fixation of a square, made it relatively unobtrusive. The place holders assured that the information in the periphery was only locative in nature, providing a marker of where the subject could look to get information.

Half a second after the subject fixated a target in the center of the display field and pressed a "ready" button, the sentence appeared at the central fixation place. The subject was timed from the onset of the display until his response terminated the trial. Each of the 12 subjects had 15 practice trials and two blocks of 48 test trials.

Results and Discussion. As Figure 12 shows, the total response times in the four information conditions did increase linearly from true affirmative, to false affirmative, to false negative, to true negative. In fact, a straight line accounts for 98.6% of the variance among the four means. The residual 1.4% of the variance is not significant, F (2, 33)<1. Thus, the pattern of total latencies for the current task resembles the latency pattern found in other experiments (cf., Carpenter & Just, 1975). These analyses concern only those trials in which the subject gave a correct response. The frequency of incorrect responses was very low, as indicated in Figure 12.

Insert Figure 12 about here

The important advantage of the current methodology is that the location and duration of the gaze allow us to break down the total response time into

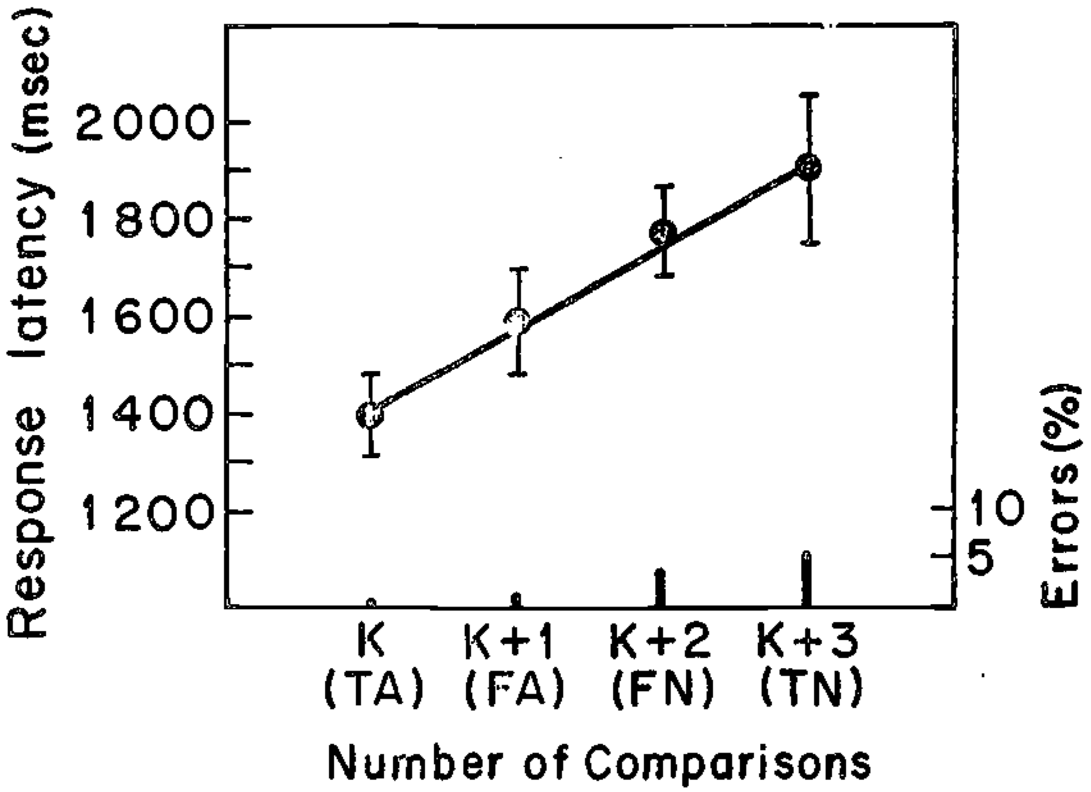

Figure 12. Mean response times for the true affirmative (TA), false affirmative (FA), false negative (FN), and true negative (TN) conditions. Response times are plotted as a function of the hypothesized number of comparison operations for each condition, where K is the number of comparisons for the true affirmative condition.

finer components. For this analysis, we divided the gazes into four categories: the initial gaze on the sentence, subsequent gazes on the sentence after having looked away, gazes on the location specified by the directional term in the sentence, and finally, gazes in any other locations. Thus, the durations of all four types of gazes add up to the total response time. The important question was whether these durations varied systematically as a function of the four information conditions.

The initial gaze on the sentence should reflect the time to read and represent the sentence. As Figure 13 shows, the duration of the initial fixation was 57 msec longer for negatives than for affirmatives, F (1, 33) = 14.93, p <.01. This result indicates that the negative sentences take about 57 msec longer to read and represent than the affirmatives. After having looked away from the sentence, subjects occasionally refixated it later in the trial. The durations of such subsequent gazes on the sentence were similar for all four information conditions, as Figure 13 shows.

Insert Figure 13 about here

The directional term in the sentence can be viewed as an instruction for where to direct the next fixation--irrespective of whether the sentence was affirmative or negative. In fact, the location specified by the directional term was the locus of the second gaze on 92% of the trials. Subjects tended to fixate this location only once during a trial. The time spent gazing at this location increased linearly with the number of hypothesized comparison operations, as Figure 13 shows. The straight line accounts for 98.1% of the variance among the four means. The residual 1.9% is not significant, F (2, 33) <1. This suggests that after reading and representing the sentence, the subject fixated the picture and completed the comparison operations while fixating the location specified by the directional term. The slope, 135 msec per operation, may be interpreted as an estimate of the

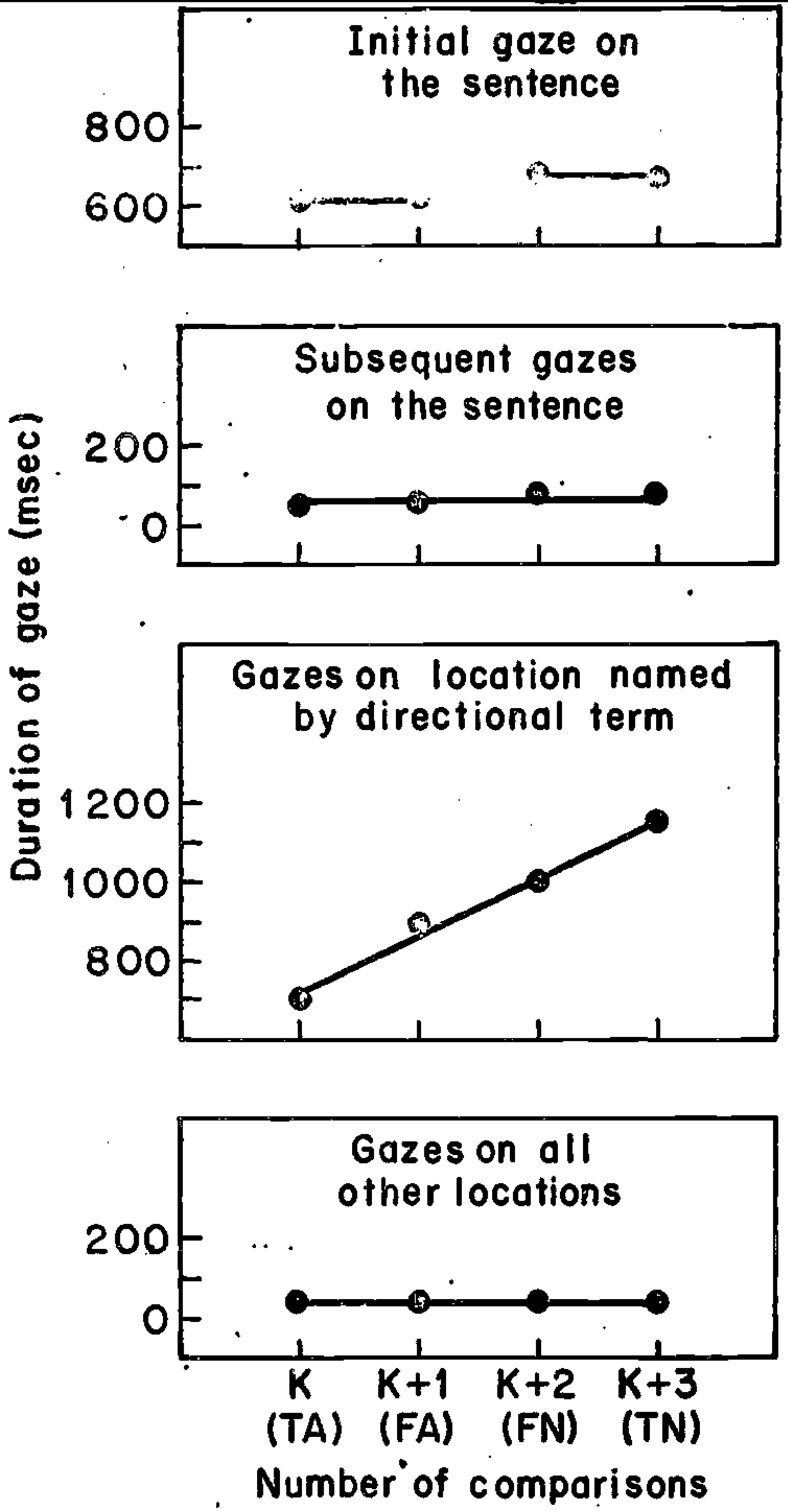

Figure 13. Average duration spent gazing on various locations of the display for the true affirmative (TA), false affirmative (FA), false negative (FN), and true negative (TN) conditions. These components add up the total response times shown in Fig. 12.

time to compare an element from the sentence representation to one from the picture representation.

Occasionally, subjects would gaze at a location other than the sentence or the location specified by the directional term. The frequency and duration of these other gazes did not vary as a function of information condition, as Figure 13 shows.

These results indicate how the total processing time in sentence verification is distributed among various stages. The duration of the initial gaze on the sentence suggests that the time needed to read and represent the sentence is 700 msec at most. This 700 msec enters primarily into the intercept of the total response time. What accounts for the difference between the response time for the fastest condition, the 1400 msec for the true affirmative, and the slowest condition, the 1900 msec for the true negative? This 500 msec is consumed by the operations that compare the sentence and picture to determine their relation. In fact, these comparison operations are reflected in the duration of the gaze on the location specified by the directional term.

This analysis can tell us why negative sentences take longer to process than affirmatives. The total response time was 346 msec longer for negatives than for affirmatives. This can be partitioned into several components. The largest component is the comparison time (reflected in the duration of gaze at the picture) which was 267 msec longer for negatives. Secondly, negative sentences took 57 msec longer to read. And thirdly, subsequent gazes on the sentence were an insignificant 20 msec longer for negative sentences. Thus the bulk of the additional processing time for negatives is consumed by the operations that compare the information from the sentence to the picture.

The results show that there is a systematic correspondence between the mental operations and eye fixations in a sentence verification task. Under well-controlled conditions, the sequence of gazes on the external display corresponds to the sequence of mental operations in the processor. Moreover, the duration of the gaze is proportional to the duration of the underlying operations.

QUANTITATIVE COMPARISON

A quantitative comparison requires an order judgment (e.g., *Which is larger?*, or *Which is brighter?*, or *Which is longer?*) of two or more objects along a common underlying dimension. The comparative judgment requires that the two objects be represented, and their representations be compared. In order to obtain more detailed evidence about the processes in this task, we devised an experiment in which subjects' eye fixations were monitored while they decided which of two groups of dots was larger. The response latencies for selecting the larger of two groups of dots strongly resemble the latencies for digit comparisons (Buckley & Gillman, 1974), so our task may produce results generalizeable to digit comparisons. Furthermore, prior data (summarized by Klahr, 1973) have shown that the time to determine how many dots there are in a group increases monotonically from about 500 msec for one dot, to 2200 msec for nine dots. These results suggest that larger groups of dots might be fixated longer if they are to be quantified. We hypothesized that the duration of fixation on each of the groups of dots might tell us how the two groups of dots were represented and processed during a quantitative comparison task.

Method. Subjects' eye fixations were monitored as they compared the sizes of two groups of dots. Each group contained from 1 to 6 dots, so there were 15 possible pairs of unequal groups. If the word *more* appeared on the left side of the display (as shown in Figure 14), subjects indicated whether the upper or the lower group contained more dots, by pressing an upper or lower response button. If the word was less, they judged which group contained fewer dots. A total of 60 stimuli were formed by orthogonally combining the two words, more and less with the 15 pairs of groups and the responses designating either the upper or the lower group. Each subject had four blocks of 60 stimuli, presented in a random order. A trial started 500 msec after the subject fixated a point at the locus of the word, and pressed a "ready" button.

Insert Figure 14 about here

The computer-generated display was presented on a video monitor at a distance of 53 to 68 cm. The word more or less, 2.8 cm wide, appeared 13 cm to the left of the dot display. The dots formed two vertical lines one above the other, separated by a vertical distance of at least 5 cm. Each group of dots was .5 to 6 cm long, depending on the number of dots in the group. For scoring purposes, the screen was divided into the four imaginary sectors indicated by the dashed lines in Figure 14. The analysis was concerned primarily with the distribution of the gaze across the four sectors.

Results. The response latencies showed that this experiment replicated the major latency results that have been previously reported for this task (Buckley & Gillman, 1974). The mean latencies ranged from 700 to 1100 msec. We will discuss the response latencies in more detail after an analysis of the eye fixation results.

First of all, we examined the duration of gaze on the smaller group. If subjects were computing the number of dots in the group, one might expect that the more dots there were in the group, the longer people would spend looking at it. As expected, the gaze duration on the smaller group in-

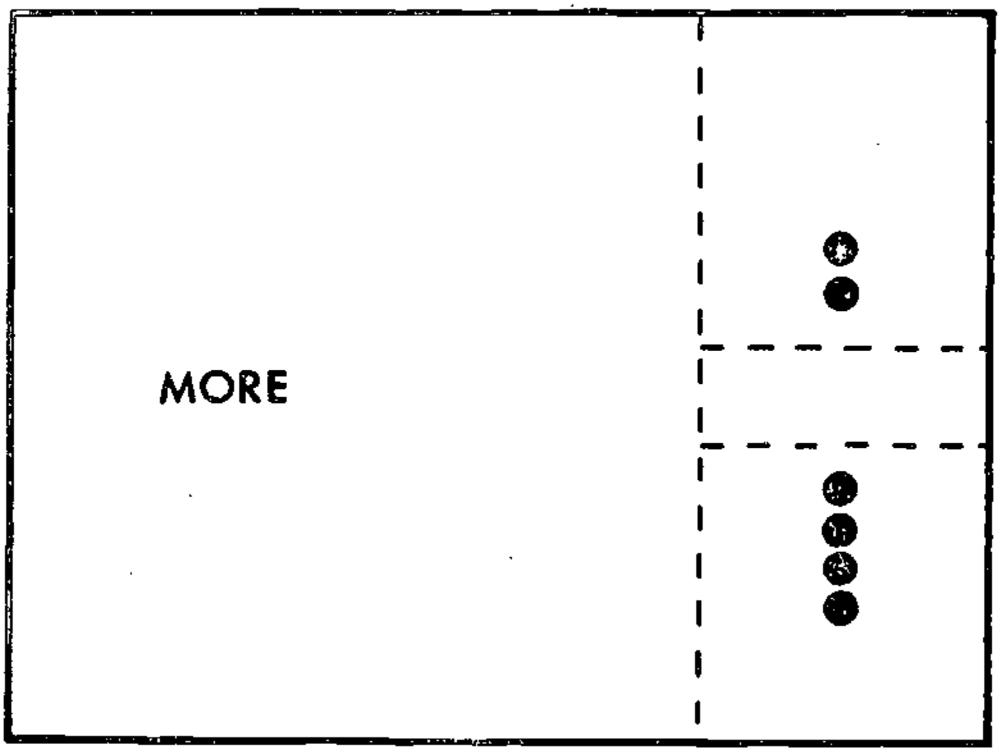

Figure 14. Schematic diagram of the visual display in the dot quantitative comparison task (not to scale). The dashed lines which indicate the boundaries between the four sectors, did not appear in the display.

creased by about 26 msec for each additional dot (see Figure 15). This is within the range of subitizing rates cited by Klahr (1973), although it is at the low end. Thus, it is plausible that the subjects compute the number of dots in the smaller group.

Insert Figure 15 about here

But what about fixation on the larger group? If subjects determine the quantity of dots in the larger group, then gaze duration on the larger group should also increase with the number of dots in that group. However, Figure 16 shows that the duration of gaze on the larger group is independent of the number of dots there. Thus, the two groups of dots are fixated differently. The size of the smaller group predicts the gaze duration on the smaller group, but the size of the larger group does not predict the gaze duration on the larger group.

Insert Figure 16 about here

The proposed model. The results are consistent with a counting model (cf. Parkman, 1971; Groen & Parkman, 1972) adapted to the dot inequality task. The process might start by counting one or two dots in each group, and checking to see if either group had been exhausted. If one group had been exhausted, it would be designated the smaller one. If neither had been exhausted, then one or two more dots might be counted in each group, and again there would be a check to see if either group had been exhausted. This process would continue until one of the groups, the smaller one, would be exhausted. If the subjects were answering the question "Which group contains more dots?", they would simply indicate the group that had not been exhausted. The number of counts or iterations in this process would be proportional to the number of dots in the smaller group. If the gaze duration is proportional to the number of increments, then it follows that duration

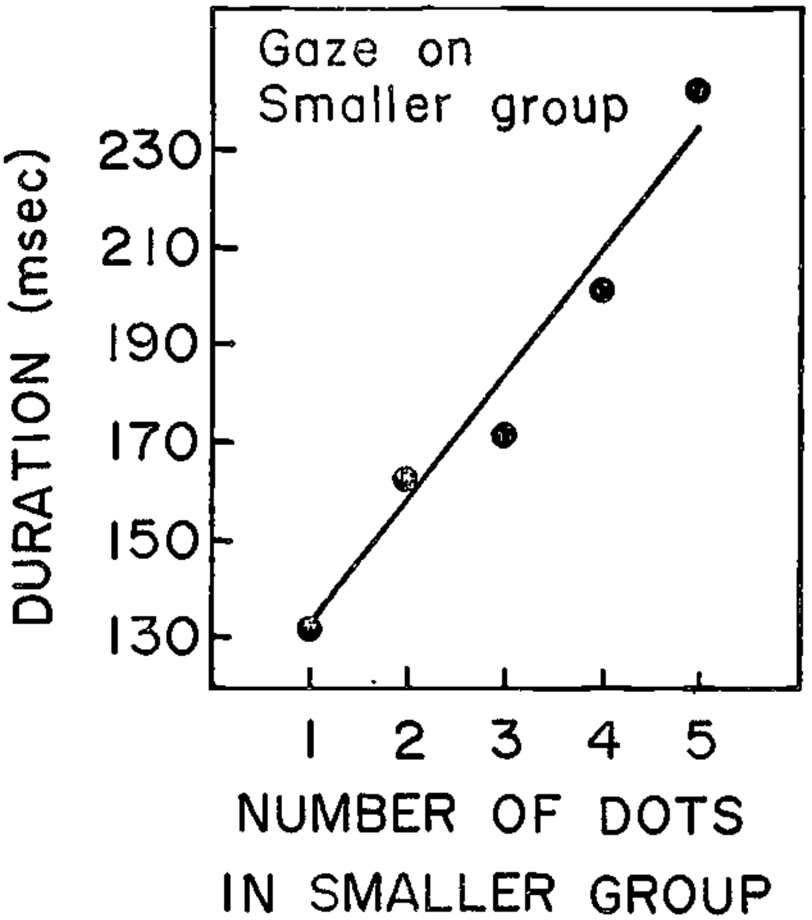

Figure 15. Mean duration of gaze on the smaller group of dots as a function of the number of dots in that group.

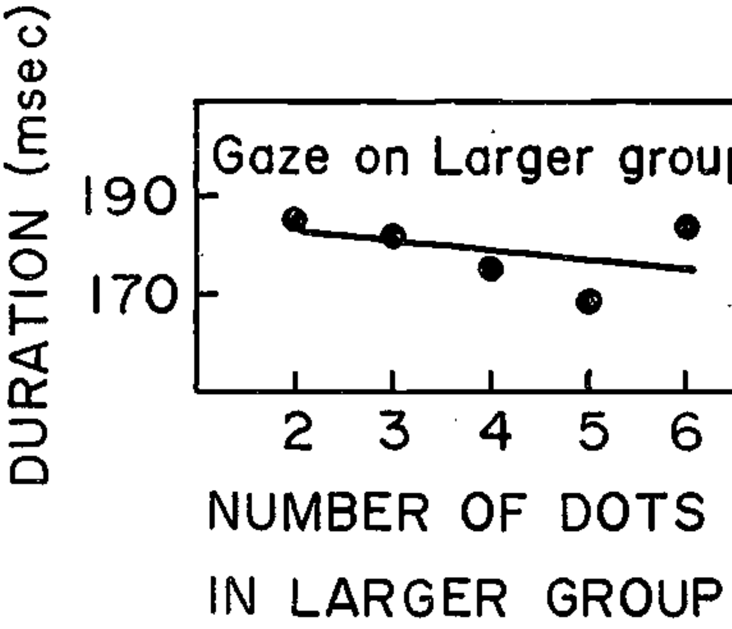

Figure 16. Mean duration of gaze on the larger group of dots as a function of the number of dots in that group.

of gaze on the smaller group should increase with the number of dots in the smaller group, as it does. (This is called the min effect, since latencies increase with the size of the smaller or minimum group.) Furthermore, the duration of gaze on the larger group should be more or less independent of the number of dots in the larger group, which it is. One further prediction of this model is that the duration of gaze on the larger group should increase with the size of the smaller group. This prediction follows from the proposal that the dots in both groups are counted only until one group (the smaller one) is exhausted. This prediction is confirmed, with gaze durations on the larger set increasing monotonically from 160 msec when the smaller group contains one dot to 296 msec when the smaller group contains 5 dots.

The proposed counting model requires supplementation to account for a persistent finding from this and previous research, namely that comparisons are faster when the absolute difference or "split" between the two groups is larger (cf., Henmon, 1906; Johnson, 1939 for the data on line length comparisons; Buckley & Gillman, 1974; Fairbank, 1969; Moyer & Landauer, 1967; Parkman, 1971; Sekuler, Rubin & Armstrong, 1971 for data on digit comparisons, and Buckley & Gillman, 1974 for data on dot comparisons). The split effect is present in both the total latencies (see Table 2) and in the gaze durations on both the smaller and larger groups of dots. We attribute the split effect to the presence of a second mechanism that can sometimes make the quantitative comparison by categorizing each of the two groups of dots as a small group or a large group. Groups of 1, 2, or 3 dots may be classified as "small", while groups of 4, 5, or 6 may be classified as "large", but the boundary may be variable across trials and subjects. If one group of dots belongs to the "small" category, and the other to the "large" category, then the one that belongs to the "large" category is larger. The category judgment mechanism may be much quicker than the counting mechanism, but it would not work when the split is small, since in those cases the two groups would tend to belong to the same category. Pairs with large splits (splits of 3, 4, or 5) could be processed with the quick category membership judgment much more often than pairs with small splits (1 or 2). The mean response latencies for any pair would be a mixture of the trials where the fast category membership judgment is used and trials where the counting mechanism is used. As the split increases, the number of fast trials contributing to the mean should increase, and so on average, the mean latencies should decrease.

Insert Table 2 about here

The two-process explanation is supported by an interaction between the min effect and the split effect observed in this experiment and others. When the split is small, the counting mechanism is more likely to be used, resulting in a strong min effect. That is, the response latencies increase with the size of the smaller group. When the split between the two groups is larger, the category judgment mechanism should be used more often, and so the min effect should decrease. The total latencies in Table 2 show this trend. When the split is small (namely, 1), then latencies increase by an average of 43 msec with each increment in the min (the smaller group). When the split is larger (2 or 3), the min effect is reduced to 16 msec. Finally, with a split of 4, there is no min effect. A similar analysis of the Buckley and Gillman (1974) dot comparison data, based on a larger range of mins and splits, further supports this conclusion.[4] For splits of 1, 2, 3, 4, 5, 6, and 7, the min effects are 66,57 49, 34, 25, 16, and 8 msec, respectively. In other words, there is a clear monotonic decrease in the min effect as the split increases. Also, there is a main effect of the split such that the latency generally decreases as the split increases. The important point, as far as the two-process explanation is concerned, is that

Table 2

Mean Latencies in msec (and % error)

for the 15 Different Pairs of Group Sizes

		Number of Dots in Smaller Group				
		1	2	3	4	5
Number of dots in Larger Group	2	817 (4.0)				
	3	757 (2.8)	844 (1.7)			
	4	718 (2.3)	744 (2.3)	873 (4.5)		
	5	730 (2.3)	729 (1.1)	770 (1.1)	916 (7.4)	
	6	716 (2.3)	731 (3.4)	750 (1.7)	800 (1.7)	997 (4.0)
	MEAN	748	762	798	858	997

when the split is larger, the category judgment mechanism may be used more often, and so the min effect decreases.

Other types of explanations account for the quantitative comparison task in terms of a quasi-logarithmic analogue representation of quantities (Buckley & Gillman, 1974; Moyer & Landauer, 1967; Shepard, Kilpatric, & Cunningham, 1975). The advantage of these alternative explanations is that they are parsimonious, and they seem readily applicable to continuous dimensions, such as sizes of animals (Moyer, 1973). However, these approaches can not easily account for the finding that the gaze duration on both groups of dots was proportional to the size of the smaller group. By contrast, a counting model is easily compatible with this aspect of the data.

The duration of gaze on the sector other than the larger and smaller group of dots did not vary from condition to condition and showed little evidence of a min effect or a split effect. Incorrect responses were rare (2.8%), and the errors were distributed as shown in Table 2. The mean processing time in this task, 793 msec, was distributed as follows. On average, 371 msec were spent gazing at the word more or less, 178 msec gazing at the larger group of dots, 165 at the smaller group, and 79 msec at the sector between the two groups. The eye fixation data also showed how this distribution of processing time was affected by the sizes of the smaller and larger groups. The results indicated that the two groups of dots are fixated in a manner consistent with an upward counting process.

OVERVIEW

The locus of the fixation. The most general assumption of the current research is that locus of the eye fixation can indicate what symbol is currently being processed. Converging lines of evidence from very diverse tasks support this general hypothesis and also allow us to refine our theoretical consideration about the relationship between eye fixations and mental processes.

In tasks where the behavioral units are fairly large and open to conscious introspection, the pattern of eye fixations correlates well with subjects' verbal reports. For example, Winikoff (1967, see also Newell & Simon, 1972) found a correlation between eye fixations and verbal report in cryptarithmetic tasks, where numbers are substituted for letters to solve a problem like DONALD + GERALD = ROBERT. The aggregation rules cumulated many fixations (as much as 5 sec of activity) into a typical processing unit. In general, Winikoff's subject tended to look at the letter whose value he was computing or trying to recall, as inferred from his concomitant verbal protocol.. Similarly, eye fixations correlate with verbal protocols when subjects are choosing among several alternatives such as cars that differ in make, year, and condition (Russo & Rosen, 1975). These studies provide evidence that the locus of the eye fixation corresponds to the information being processed in tasks where subjects can verbalize what they are processing.

Some aspects of problem solving involve operations too rapid for verbal protocols, but the eye fixations still reveal what symbols the subjects are processing. A good example are the few scan paths that have been recorded of chess masters scanning a board position for 5 sec (de Groot & Jongman, 1973; Tichomirov & Posnyanskaya, 1966). The locus of eye fixations is accounted for by assuming that the master scans between pairs of pieces that are related by attack or defense (Simon & Barenfeld, 1969). Again, these data support the hypothesis that the locus of the eye fixations reflects what is being internally processed.

Since eye fixations are sensitive to the structure of the internal representation being constructed or operated upon, they provide a valuable methodology for examining how linguistic material is interpreted. Our own research strategy in this area has been to present a linguistic stimulus, followed by a picture, and examine how the internal representation of the prior sentence alters the way the picture is scanned in a verification task. For example, we have used this methodology to examine the processing of affirmative and implicitly negative sentences (Carpenter & Just, 1972). The affirmative sentences (e.g., A small proportion of the dots are red) and the implicitly negative sentences (e.g., Few of the dots are red) have the same truth value. However, linguistic and psychological evidence suggests that the two sentences have different internal representations (Just & Carpenter, 1971). The affirmative sentence is represented as an affirmation that the small subset has some property, in this case, redness. We predicted that after reading the affirmative that refers to the small subset (e.g., A small proportion...), people should tend to fixate the small subset. By contrast, an implicit negative is represented as a negation of some property of the large subset, in this case, redness. We predicted that after reading an implicit negation about the large subset (e.g., Few of the dots are red), people would tend to fixate the large subset. As predicted, subjects looked at the location in the picture specified by the underlying representation of the sentence. The locus of the eye fixation is sensitive to the deep structure representation, even when subjects aren't consciously aware of the nature of the linguistic stimulus or of their pattern of eye fixation.

While people are listening to spoken questions or passages, they tend to fixate the pictorial referent of words that occur in the text (Cooper, 1974; Kahneman & Lass, 1971, cited by Kahneman, 1973). For example, in the Kahneman and Lass study, people were shown a schematic drawing of four objects, such as a car, person, tree and airplane and asked a question like "What makes of cars can you name?" Subjects tended to look at the schematic car while answering. More interestingly, when the picture was removed prior to the question, subjects still tended to look where the appropriate object had been located. Such fixations apparently play a place-keeping organizational role rather than an encoding role. The symbols in the short term memory may be indexed to particular spatial locations. (This formulation is reminiscent of the method of loci (cf. Bower, 1970) and spatial interference effects in retrieval (Byrne, 1974)). When the time comes to retrieve or operate on a symbol, the eye may fixate the location from which the symbol was originally encoded. It may be this mechanism that produces fixations on the referent of the symbol at the top of the stack, assuming that the referent stays in the same location.

Duration of gaze. In the tasks we investigated, the time spent gazing at a figure reflected both the time to encode that figure as well as the time to operate on the encoded symbol. Tachistoscopic recognition studies indicate that familiar figures, like alphanumeric characters or even words can be internalized within a very short exposure duration--as low as a few tens of milliseconds. Yet in these cognitive tasks, people gaze at very simple and familiar figures for much longer, often for hundreds of milliseconds. For example, in the sentence verification task, subjects looked at a star (*) or a plus (+) for 700 to 1200 msec, depending upon the relation between the sentence and the figure. Clearly, the duration of the gaze includes not only encoding time but also the time for subsequent operation on the encoded symbol.

There are a number of reasons why a subject might continue to fixate a figure after the relevant information has been encoded. If the processor is busy operating on the most recently encoded information, there is no reason for it to direct the eye to seek other information. So the eye may remain stationary simply because it is not instructed to move. An alternative view of the persistence of the gaze is that the processor might actively instruct the eye not to move during the processing of the most recently encoded information. The reason for avoiding new fixations might be that a saccade automatically initiates an encoding activity (cf. Loftus, in press) that could interrupt the ongoing processing. Perhaps the reason that people often gaze upwards or close their eyes altogether while computing the answer to a demanding question is that they are avoiding extraneous encoding operations that could interrupt processing. Thus, the persistence of the gaze could be due to the absence of an instruction to move the eye or the presence of an instruction not to move the eye. In either case, the gaze duration on a particular figure provides a measure of the time spent processing the corresponding symbol.

One of the most elegant studies of the relationship between gaze duration and mental operation examined gaze duration in a Sternberg memory-scanning task (Gould, 1973). In Gould's experiment, subjects had a memory set of 1, 2, or 3 letters, and 12 probe letters were distributed along the perimeter of an imaginary clock face that corresponded to the display. Only one of the 12 probe letters was a member of the memory set. The subject's task was to scan around the clock face (starting at 12 o'clock and proceeding clockwise) until he found the positive probe.

The duration of the time spent fixating each negative probe item increased linearly with the memory set size, at a rate of about 50 msec/item. This is compatible with the explanation that each probe item was serially compared to each of the memory set items, and as the memory set size increased, the probe had to be compared to more items in memory. The eye fixated the probe while the comparison operations occurred. The importance of this finding is that the parameter of 50 msec per item, inferred from the duration of the gaze, is very close to the time of 38 msec per item inferred from reaction time studies with only a single probe (Sternberg, 1969). The duration of fixation on the positive probe item also increased as a function of the memory set size. Of particular importance is that the rate of increase in this duration, the slope, was the same for the positive probe as for the negative probes. This result suggests that the scan through the memory set is exhaustive.

Gould's results provide an important validation for the eye fixation methodology. The eye fixation measures yield results that are completely consistent with Sternberg's careful reaction-time studies. Gould's study, as well as the current research, combines the use of additive factors techniques and eye fixation measures to estimate the time taken to execute a mental operation.

The only eye fixation research that reports a lack of correlation between fixation duration and performance concerns memory for pictures (Loftus, 1972). Loftus found that during learning, the number of fixations, not their total duration, was the best predictor of subsequent recognition memory. However, Tversky (1974) has recently found a positive correlation between the duration of individual eye fixations and later memory. Tversky suggests that the critical variable in relating fixation duration to picture memory might be the kind of features being encoded in the learning phase. The present research does not attempt to account for eye fixations in picture scanning and recognition (for relevant work on this topic, see Buswell, 1935; Mackworth & Bruner, 1970; Mackworth & Morandi, 1967; Noton & Stark, 1971; Potter & Levy, 1969). Our concern has been with ongoing computation rather than search processes in long-term memory, so the picture recognition issue would take our discussion too far afield. However, the present theoretical framework suggests that the resolution of this problem requires a model of what is encoded during the initial learning and what tests are made during the subsequent recognition phase. As yet, the data relating fixation duration to recognition memory are insufficient to construct a complete model of the mental processes in picture memorization and retrieval.

We have related mental operations to the duration of gaze, not necessarily to the duration of individual fixations. Fixations must be aggregated into a unit consistent with the underlying theory. For example, in the sentence verification task, when a person looked at the plus, whether he made one or two fixations was irrelevant to the fact that he was looking at that piece of information and no other. We measured how long he looked at the plus, not how many fixations actually went into that gaze time. The duration of gaze was systematically related to the duration of underlying processes. The duration of individual fixations may also be so related, but only under those conditions in which each fixation is a psychologically separate unit (cf. also Carpenter & Just, in press; Russo & Rosen, 1975).

Task conditions that optimize the use of eye fixations. The locus of fixation is not always synonymous with the direction of attention. Subjects can be instructed to fixate one referent while attending elsewhere. The possibility of such disassociation makes it important to specify the conditions under which eye fixations are an accurate reflection of what is being processed. One of the most important conditions is that the task require that information from the visual environment be encoded and processed. If the visual display is not relevant, there are no mapping rules between what is being fixated and what is being internally processed. A second condition is that the task goals be specified for the subject. Asking subjects simply to look at a picture or read some prose permits them to adopt their own definitions of what processing is required and this again makes it difficult to infer the relationship between eye fixations and underlying mental processes. And of course, speeded tasks discourage extraneous processing and the concomitant extraneous fixations.

Some of the rules that govern fixations are general scanning strategies, while other rules are highly specific to the processing in the task being performed. Eye fixations will reveal the mental processes in a particular task only if the task structure minimizes the use of general scanning strategies. An example of this structuring is evident in the study of how people looked at pictures after reading sentences involving affirmative quantifiers e.g., A small proportion of the dots are red, or negative quantifiers, e.g., Few of the dots are red (Carpenter & Just, 1972). The pictures always had a small subset of dots at the top and a large subset at the bottom. Thus, the subject knew to look at the top or at the bottom, depending on whether he wanted to determine the color of the small subset or the large one. This task structure eliminated the need first to search for the desired subset and then to encode its color. The relation between eye fixations and mental operations is even clearer when the role of peripheral information is controlled. The extreme case of this is the computerized "tunnel vision" in the sentence verification task, in which there is no peripheral information, so the duration of gaze at any locus cannot reflect encoding of information from another locus. These features of the task structure minimize the role of general scanning strategies and thereby make the design more sensitive to the cognitive processes of interest.

In all of these tasks, the eye scan is very much goal directed, in fact, directed by the information present "at the top of the stack." There are two possible sources of such information, namely, the task structure and information computed during the trial. Both sources influenced fixations in the sentence verification task where the instructions to fixate the sentence determined the first fixation, but the locus of the second fixation was determined by information computed during the trial. After the sentence (e.g. Plus isn't North) was fixated, the directional term in the sentence determined the locus of the next fixation, in this case, North. Since both the task structure and the ongoing processing can determine the locus of fixation, both factors must be taken into account in developing a complete processing model.

One domain of eye fixation research that has been hampered by the absence of task analyses is the area of reading. While there have been many promising empirical studies of eye fixations in reading (cf. Buswell, 1922, 1937; Hochberg, 1970; Kolers, 1970; Levin & Kaplan, 1970; Mackworth, 1974; McConkie & Rayner, 1974; Mehler, Bever, & Carey, 1967; Tinker, 1958), there is no convergence on a theory of reading. The difficulty is that there is no single "reading process", because we read differently in different situations. For example, a newspaper article is read differently from a legal contract, and the same contract is read differently depending on whether one is looking for typographical errors or buying a house. One possible solution to this difficulty is to systematically examine a number of well-defined reading tasks that are of inherent interest. For example, we are currently examining eye fixations of subjects who read and solve three term series problems (e.g., If John is leading Bill, and Tom is following Bill, then where is John?). In another study, we are examining subjects' eye fixations as they read a passage and answer questions about it in a standardized reading comprehension test. In order to develop models of reading, it will probably be necessary to study performance in a number of well-understood task environments, so as to determine the influence of the environments on the reading process.

Generalization of the models. Certain kinds of operations in the central processor appear to function similarly irrespective of the source of encoding of the operated-on symbol, be it a visual display, tactile input, semantic memory retrieval or whatever. In those cases in which the operations are invariant, conclusions gained from the eye fixation methodology may generalize to processing of symbols in non-visual domains. The invariant operations would presumably be very basic ones, such as comparing two symbols for identity, retrieving the next symbol in an ordered list, or incrementing an internal counter.

One example from recent psycholinguistic research demonstrates how sentences that refer to information from different sources (like pictures vs semantic memory) may be processed similarly. Just (1974) timed subjects while they verified quantified sentences like Some of the red figures are round with respect to a picture that included red and round figures. The overall pattern of latencies was similar to the pattern obtained when the sentences refer to concepts in semantic memory, e.g. Some men are doctors (Meyer, 1970). In fact, even though the relevant information was encoded from a picture in one case and retrieved from semantic memory in the other, both sets of data could be explained in terms of the same operations (Just, 1974). Obviously, the initial encoding stages involve different processes, but in this and certain other cases (cf. Carpenter & Just, 1975), the information seems to be manipulated similarly once it is past the encoding stages. This suggests that processing models of these subsequent stages derived from eye fixations studies may generalize to non-visual domains.

Internal rotation processes may also be somewhat independent of the visual modality. When subjects are deciding whether a visually presented, rotated "R" is normal or a mirror image (Cooper & Shepard, 1973), the response latencies resemble those for the Shepard and Metzler task in certain respects. The resemblance led Cooper and Shepard to argue that the processes in the two tasks were similar. In the Cooper and Shepard task there cannot be eye fixations switching back and forth between the two R's, since only one of them is externally present, while the other is the long term representation of a normal R. Nevertheless, it is reasonable to speculate that the sequence of internal switches of attention in the Cooper and Shepard study is related to the external sequence of fixation switches observed in our study of rotation. If this speculation is correct, the model we have proposed may have some application to rotation processes that involve mentally generated stimuli.

If processing models based on eye fixation studies are to be generalized to non-visual tasks then the factors that influence only visual encoding must be identified. For example, picture scanning processes might be affected by perceptual saliency (Williams, 1966), and there may be no parallel in semantic memory retrieval. Conversely, semantic memory retrieval may be affected by factors such as semantic distance (cf. Rips, Shoben & Smith, 1973), which has no parallel factor in picture encoding processes. If these modality-specific processes can be isolated, then eye fixations may provide a way to investigate the fundamental operations that occur in the central processor. Operations whose durations lie between 50 and 800 msec seem especially susceptible to this approach, as shown by the current work on rotation, sentence verification, and quantitative comparison. For these rapid operations, there is a very close link between the symbol that is being processed and the locus, sequence and duration of eye fixations, because of the eyes' tendency to fixate the referent of the symbol that is " at the top of the stack".

REFERENCES

Alpern, M. Movements of the eyes. In H. Davson (Ed.), The Eye. New York: Academic Press, Inc., 1962, 3, part 1.

Bartz, A. E. Fixation errors in eye movements to peripheral stimuli. Journal of Experimental Psychology, 1967, 75, 444-446.

Bower, G. H. Analysis of a mnemonic device. American Scientist, 1970, 58, 496-510.

Buckley, P. B., & Gillman, C. B. Comparisons of digits and dot patterns. Journal of Experimental Psychology, 1974, 103, 1131-1136.

Buswell, G.T. Fundamental Reading Habits, a Study of their Development. Chicago: Chicago University Press, 1922.

Buswell, G. T. How People Look at Pictures. Chicago: University of Chicago Press, 1935.

Buswell, G. T. How adults read. Supplementary Educational Monographs No. 45, 1937.

Byrne, B. Item concreteness vs. spatial organization as predictors of visual imagery. Memory & Cognition, 1974, 2, 53-59.

Carpenter, P. A., & Just, M. A. Semantic control of eye movements during picture scanning in a sentence-picture verification task. Perception and Psychophysics, 1972, 12, 61-64.

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

Carpenter, P. A., & Just, M. A. Linguistic influences on picture scanning. In R. A. Monty & J. W. Senders (Eds.), Eye Movements and Psychological Processes. Hillsdale, N. J.: Lawrence Erlbaum Associates, 1976 (in press).

Chase, W. G., & Clark, H. H. Mental operations in the comparison of sentences and pictures. In L. Gregg (Ed.), Cognition in Learning and Memory. New York: Wiley, 1972.

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

Cooper, R. M. The control of eye fixation by the meaning of spoken language. Cognitive Psychology, 1974, 6, 84-107.

Cooper, L. A., & Shepard, R. N. Chronometric studies of the rotation of mental images. In W. G. Chase (Ed.), Visual Information Processing. New York: Academic Press, 1973.

Dearborn, W. The Psychology of Reading. Columbia University Contributions to Philosophy and Psychology, XIV. New York: The Science Press, 1906.

de Groot, A., & Jongman, R. Perception and memory in chess: An experimental analysis of the master's professional eye. RITP Memorandum No. 024, University of Amsterdam, Holland, 1973.

Ditchburn, R. W. Eye movements and visual perception. Oxford, England: Clarendon, 1973.

Fairbank, B. A. Jr. Experiments on the temporal aspects of number perception. Unpublished doctoral dissertation, University of Arizona, 1969.

Gillman, C. B., & Buckley, P. B. Numeric Comparison. Paper presented at the Fourteenth Annual Psychonomic Society Meeting, St. Louis, Missouri, November 2, 1973.

Gould, J. Eye movements during visual search and memory search. Journal of Experimental Psychology, 1973, 98, 184-195.

Groen, G., & Parkman, J. A chronometric analysis of simple addition. Psychological Review, 1972, 79, 329-343.

Henmon, V. A. C. The time of perception as a measure of differences in sensation. Arch. Philos. Psychol. Sci. Method, 1906, 8, 5-75.

Hochberg, J. Components of literacy: Speculations and exploratory research. In H. Levin and J. P. Williams (Eds.), Basic Studies on Reading. New York: Basic Books, Inc., 1970.

Huey, E. The Psychology and Pedagogy of Reading. The Macmillian Co., 1908. Reprinted Cambridge, Mass.: The M.I.T. Press, 1968.

Javal, E. Essaie sur la physiologie de la lecture. Annales d'oculistique, 1878, 79, 97.

Johnson, D. M. Confidence and speed in the two-category judgment. Archives of Psych., 1939, 241, 1-52.

Just, M. A. Comprehending quantified sentences: The relation between sentence-picture and semantic memory verification. Cognitive Psychology, 1974, 6, 216-236.

Just, M. A., & Carpenter, P. A. Comprehension of negation with quantification. Journal of Verbal Learning and Verbal Behavior, 1971, 10, 244-253.

Kahneman, D. Attention and Effort. Englewood Cliffs, New Jersey: Prentice-Hall, Inc., 1973.

Kahneman, D., & Lass, N. Eye position in tasks of association and memory. Unpublished manuscript; Hebrew University, Jerusalem, 1971.

Klahr, D. Quantification processes. In W. G. Chase (Ed.), Visual Information Processing. New York: Academic Press, 1973.

Kolers, P. A. Three stages of reading. In H. Levin and J. Williams (Eds.), Basic Studies in Reading. New York: Basic Books, Inc., 1970.

Levin, H., & Kaplan, E. Grammatical structure and reading. In H. Levin and J. P. Williams (Eds.), Basic Studies on Reading. New York: Basic Books, Inc., 1970.

Loftus, G. R. Eye fixations and recognition memory for pictures. *Cognitive Psychology*, 1972, *3*, 525-551.

Loftus, G. R. A framework for a theory of picture recognition. In R. A. Monty and J. W. Senders (Eds.), *Eye Movements and Psychological Processes*. Hillsdale, N. J.: Lawrence Erlbaum Associates, 1976 (in press).

Mackworth, N. The line of sight approach to children's reading and comprehension. In S. Wanat, H. Singer, and M. Kling (Eds.), *Extracting Meaning from Written Language*. Newark, Delaware: International Reading Association, 1974.

Mackworth, N. H., & Bruner, J. S. How adults and children search and recognize pictures. *Human Development*, 1970, *13*, 149-177.

Mackworth, N. H., & Morandi, A. J. The gaze selects informative details within pictures. *Perception & Psychophysics*, 1967, *2*, 547-552.

McConkie, G. W., & Rayner, K. Identifying the span of the effective stimulus in reading. Reading and Learning Series. Research Report No. 3. Cornell University, Ithaca, New York, 1974.

Mehler, J., Bever, T. G., & Carey, P. What we look at when we read. *Perception & Psychophysics*, 1967, *2*, 213-218.

Metzler, J., & Shepard, R. Transformational studies of the internal representation of three-dimensional objects. In R. Solso (Ed.), *Theories in Cognitive Psychology: The Loyola Symposium*. Potomac, Maryland: Lawrence Erlbaum Associates, 1974.

Meyer, D. E. On the representation and retrieval of stored semantic information. *Cognitive Psychology*, 1970, *1*, 242-300.

Moyer, R. S. Comparing objects in memory: evidence suggesting an internal psychophysics. *Perception and Psychophysics*, 1973, *13*, 180-184.

Moyer, R. S., & Landauer, T. K. Time required for judgments of numerical inequality. Nature, 1967, 215, 1519-1520.

Neisser, U. Cognitive Psychology. New York: Appleton-Century-Crofts, 1967.

Newell, A. Production systems: Models of control structures. In W. G. Chase (Ed.), Visual Information Processing. New York: Academic Press, 1973.

Newell, A., & Simon, H. A. Computers in psychology. In R. D. Luce, R. R. Bush, and E. Galanter (Eds.), Handbook of Mathematical Psychology, Vol. 1. New York: Wiley, 1963.

Newell, A., & Simon, H. A. Human Problem Solving. Englewood Cliffs, N. J.: Prentice-Hall, Inc., 1972.

Noton, D., & Stark, L. Eye movements and visual perception. Scientific American, 1971, 224, 34-43.

Parkman, J. A. Temporal aspects of digit and letter inequality judgments. Journal of Experimental Psychology, 1971, 91, 191-205.

Posner, M. I. Short-term memory systems in human information processing. In A. F. Sanders (Ed.), Attention and Performance. Acta Psychologica, 1967, 27, 267-284.

Potter, M. C., & Levy, E. I. Recognition memory for a rapid sequence of pictures. Journal of Experimental Psychology, 1969, 81, 10-15.

Rips, L. J., Shoben, E. J., & Smith, E. E. Semantic distance and the verification of semantic relations. Journal of Verbal Learning and Verbal Behavior, 1973, 12, 1-20.

Russo, J., & Rosen, L. An eye fixation analysis of multialternative choice. Memory and Cognition, 1975, 3, 267-276.

Sekuler, R., Rubin, E., & Armstrong, R. Processing numerical information. Journal of Experimental Psychology, 1971, 90, 75-80.

Shepard, R. N., Kilpatric, D. W., & Cunningham, J. P. The internal representation of numbers. Cognitive Psychology, 1975, 7, 82-138.

Shepard, R., & Metzler, J. Mental rotation of three-dimensional objects. Science, 1971, 171, 701-703.

Simon, H. A., & Barenfeld, M. Information processing analysis of perceptual processes in problem solving. Psychological Review, 1969, 76, 473-483.

Sternberg, S. The discovery of processing stages: Extensions of Donder's method. In W. G. Koster (Ed.), Attention and Performance II. Acta Psychologica, 1969, 30, 276-315.

Tichomirov, G. K., & Poznyanskaya, E. D. An investigation of visual search as a means of analyzing heuristics. Voprosy Psikhologii, 1966, 12, 39-53.

Tinker, M. A. Recent studies of eye movements in reading. Psychological Bulletin, 1958, 55, 215-230.

Trabasso, T., Rollins, H., & Shaughnessy, E. Storage and verification stages in processing concepts. Cognitive Psychology, 1971, 2, 239-289.

Tversky, B. Eye fixations in prediction of recognition and recall. Memory & Cognition, 1974, 2, 275-278.

Wickelgren, W. A. Speed-accuracy tradeoff and information processing dynamics. Unpublished manuscript. University of Oregon, Eugene, Oregon, 1974.

Williams, L. G. The effect of target specification on objects fixated during visual search. Perception and Psychophysics, 1966, 1, 315-318.

Winikoff, A. Eye movements as an aid to protocol analysis of problem solving behavior. Unpublished doctoral dissertation, Carnegie-Mellon University, Pittsburgh, Pa., 1967.

Woodworth, R. S. Experimental Psychology. New York: Henry Holt, 1938.

Yarbus, A. L. Eye Movements and Vision. New York: Plenum Press, 1967.

FOOTNOTES

[1] The order of authors is arbitrary. This paper represents a collaborative effort.

We are grateful to Roger Shepard and Herbert Simon for their comments on earlier drafts of this paper. We also thank Roger Shepard for providing copies of the stimulus figures used in the rotation experiment.

The project was supported in part by Research Grant NIE-G-74-0016 from the National Institute of Education, U. S. Department of Health, Education, and Welfare, and Grant MH-07722 from the National Institute of Mental Health.

Requests for reprints should be sent to Marcel Adam Just, Psychology Department, Carnegie-Mellon University, Pittsburgh, Pennsylvania 15213.

[2] Metzler and Shepard (1974) report some preliminary observations on the eye movements of two subjects performing the mental rotation task; their subjects also looked back and forth between the two figures.

[3] The 50° steps indicated by our data are suggestively close to 45°, which has more intuitive appeal.

[4] The analysis is based on cell means estimated from a graph of the latencies for comparing random configurations of dots (Gillman & Buckley, 1973). The aggregated data appear in Buckley and Gillman (1974).