

# The role of eye-fixation research in cognitive psychology

MARCEL ADAM JUST and PATRICIA A. CARPENTER

*Psychology Department, Carnegie-Mellon University, Pittsburgh, Pennsylvania 15213*

Eye-fixation research can be used to address issues of theoretical interest in several areas of cognitive psychology, such as linguistic processing, spatial processing, and reading. The eye-fixation behavior can lead to processing models, assuming that the eye fixates the referent of the symbol being operated on.

In this paper, we will attempt to classify the kinds of questions a cognitive psychologist can hope to answer with the help of an on-line eye-tracking system. Some of the unique capabilities of the eye fixation laboratory at Carnegie-Mellon University will be described. Where possible, research that uses the technology of eye-fixation research to answer questions about cognitive processes will be cited.

Eye-tracking systems, in one form or another, have been available for nearly 100 years (Javal, 1878). While eye-tracking systems have become much more sophisticated, it appears that the research questions posed and answered have not kept pace. In the study of eye fixations in reading, for example, Buswell's work (1922, 1937) and Tinker's summary (1958) are still appropriate references and are not out of date. Unfortunately, little of the richness of cognitive psychology has seeped into the eye-fixation research over the past 15 years. If we were to search the psychological literature of the past 5 years for publications on memory scanning, we might find a few hundred papers, but if we were to search for eye-fixation studies of cognition, we would find a relatively small number of papers. (If we were to search for the intersection of these two classes, namely eye-fixation studies of memory scanning, we would find one study [Gould, 1973]).

Current cognitive psychology views the mind as an information-processing system within which most 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). We propose that the rapid mental operations of the central processor can be revealed by an analysis of the eye fixations

during a task involving visual input. To support this thesis, we will discuss such tasks as the comparison of rotated figures (Shepard & Metzler, 1971) and sentence verification (Carpenter & Just, 1975). On the practical side, these tasks generally take less than 5 to 10 sec, so equipment calibration and alignment problems are minimized. On the theoretical side, they seem to be composed of elementary mental operations that take between 200 and 500 msec to execute. Thus, the durations of the individual mental operations are highly compatible with the speed with which a person can execute a separate fixation. Such characteristics make these and similar tasks fruitful areas for exploring cognitive processes with eye-fixation measures.

The critical assumption that underlies our analysis is that under certain circumstances, the eye fixates the referent of the symbol currently being processed. That is, the fixated object may reflect what is "at the top of the stack" in active memory. 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 gaze on each referent may be proportional to how long the corresponding symbol is operated on. The obvious advantage of monitoring eye fixations is that the behavior within any one particular trial can potentially be decomposed into a sequence of information-processing 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 fixations can be sampled at high densities per unit time. A high sampling rate (say once every 16 msec) is necessary and appropriate when studying mental operations whose durations range from tens to hundreds of milliseconds. If one were to study the behavior of glaciers, it might be sufficient to take a photograph once every few months, but to study the blossoming of a flower, one might want hourly photographs. Similarly, to study the rapid mental operations of the

This work was supported in part by Research Grant MH-07722 from the National Institute of Mental Health. We thank David Klahr for his comments on an earlier draft of the paper. Requests for reprints should be sent to Marcel Adam Just, Psychology Department, Carnegie-Mellon University, Pittsburgh, Pennsylvania 15213.

central processor, it is desirable to monitor its behavior many times per trial, so as to separate the behavior into stages. The trace of the stages may specify their respective durations and the sequence in which they occur.

### INSTRUMENTATION

In our laboratory, eye tracking is done with a corneal-reflection eye camera. The eye spot is electronically amplified and superimposed on the signal from the display, yielding a video-taped record of the position of the eye spot relative to the display. Meanwhile, the coordinates of the eye spot are computed every 16 msec, and these data are stored on the disk on a Honeywell DDP 116. At the end of the testing session, the data are shipped via a high-speed line to the University IBM 360 computer for analysis. The subject's display monitor takes its signal either from the computer or from a closed-circuit television camera.

The program always knows where the subject is looking, and it also controls the subject's display. Therefore we can make the display contingent on the locus of the subject's gaze. So far this feature has been used in two ways. First, the program will not start a trial unless the subject is fixating the pretrial fixation point. This ensures that the subject is cooperating by fixating the appropriate place, and it also ensures that his eye spot is in calibration before each trial. The other main use of the gaze-contingent display is to eliminate peripheral information. We create "computerized tunnel vision" by presenting only that part of the display upon which the subject is fixating. In this situation, we can be sure that if a subject is looking at a particular location, he is not encoding information at other locations, simply because we display no information at other locations. When a subject moves his eyes from Point A to Point B, the information at A disappears and the information at B appears within 34 msec of the movement. This rapid replacement, closely linked to the saccade, makes the change in the display relatively unobtrusive. We are currently developing a gaze-contingent display that need not be computer generated, but can take a video signal from the closed-circuit television camera. This will allow us to have gaze-contingent displays from any stimulus that can be transmitted through a television camera.

The low cost of collecting and analyzing eye-fixation data (relative to human-scored data) makes it possible to explore many new areas of human information processing.

### RESEARCH AREAS UTILIZING EYE-FIXATION MEASURES

#### Processing Linguistic Information

The major development in the psychology of

language over the past 15 years has been the idea that linguistic discourse is internally represented as an abstract symbolic structure, i.e., we do not internalize words or sentences, but units of meaning. The new questions in psycholinguistics concern the derivation of the representation of a sentence (or any linguistic structure), the format of the representation, and the operations brought to bear on this representation. That is, how do we extract information from a sentence, and what can we do with the information after we have extracted it?

Our strategy 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 by the eye in a verification task. The sentence provides a program for the central processor to execute, and the eye fixations provide a trace of the execution of the program. For example, we have used this methodology to examine the processing of affirmative and implicitly negative sentences (Carpenter & Just, 1972). The affirmative sentences either referred to a large subset of dots (e.g., Many of the dots are black.) or to a small subset (e.g., A small proportion of the dots are red.). The implicitly negative sentences have a surface structure similar to affirmatives that refer to the small subset (e.g., Hardly any of the dots are red.). In fact, both the implicitly negative sentences and the affirmative sentences that refer to a small subset are true with respect to the same picture. However, linguistic and psychological evidence suggested that the implicit negatives have an internal representation quite different from the surface structure properties (Just & Carpenter, 1971). An implicit negative is represented as a denial that the large subset is red. We predicted that after reading such an implicit negation about the large subset, people would tend to fixate the large subset. By contrast, after reading the affirmative that refers to the small subset (e.g., A small proportion . . .), people should tend to look at the small subset. As predicted, subjects looked at the location specified by the underlying representation, rather than at the location specified by the simple surface structure properties. The locus of the eye fixation is sensitive to the deep structure representation, even when subjects are not consciously aware of the nature of the linguistic stimulus or of their pattern of eye fixation.

One experiment using the gaze-contingent display required that subjects determine whether a printed sentence was true or false with regard to an accompanying picture, and we monitored fixations on the sentence and on the picture (Just & Carpenter, 1976). Either the sentence or the picture, depending on which was fixated, appeared on the display. With this design, we determined that a negative sentence (e.g., The plus is not North.) takes 57 msec longer to read than an affirmative sentence (e.g., The plus is North.). The total sentence-verification time was

about 350 msec longer for negative than for affirmative sentences. But subjects fixated negative sentences for only 57 msec longer than affirmative sentences (in the gaze-contingent display, only the sentence was presented during this gaze). The remaining 293 msec due to negation was consumed by operations that compare the sentence representation with the picture representation. Had it not been for the gaze-contingent display, we would not have been able to separate the sentence reading time from subsequent sentence-picture comparison operations.

### Eye Fixations in Reading

Eye-fixation research on reading has been hampered by the absence of adequate task analyses. 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 way one reads a legal contract is dependent on whether one is looking for typographical errors or checking the detailed stipulations of the contract. Cognitive research has repeatedly shown that adequate models of psychological processes require adequate models of the task structure. In the case of reading, the task structure includes the linguistic properties of the material being read, as well as task characteristics such as speed, thoroughness, or reading goals. Current research in linguistics and psycholinguistics offers powerful tools that can be brought to bear in such analyses. Thus, reading research must begin to specify both the nature of the material being read and the conditions under which it is being read.

We are examining eye fixations of subjects who read and solve three-term series problems, for example, "If John is leading Bill, and Tom is following Bill, then where is John?" In this kind of problem, the first premise describes the relative positions of two objects along some dimension; the second premise states the position of a third object relative to one of the two mentioned in the first premise. The subject's main processing requirement is to integrate the information from the two premises. The linguistic form in which the relations are expressed in the two premises is what determines the difficulty of this integrative process. We will be examining eye fixations on the two premises, with special interest in fixations on the object mentioned in both premises, that is, the linking object. There may be regressive fixations to the first premise on reading the second, or fixations on the repeated object in the

second premise might be especially long. The point is that we know exactly what the structure of this kind of problem is, so we should be able to make sense of the eye fixations during reading. If we are able to interpret the sequence and duration of eye fixations sensibly, then we should be on the road to constructing a model of how people read and solve three-term series problems. In order to develop models of reading in more complex situations, 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.

### Problem Solving

The behavioral units in problem solving are fairly large and often open to conscious introspection. In these cases, the pattern of eye fixations generally 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 cryptarithmic 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 in decision making, when subjects are choosing among several alternatives such as cars that differ in make, age, and condition (Russo & Rosen, 1975). These studies provide evidence that the locus of the eye fixation corresponds to the symbol 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 eye fixations still reveal which symbols the subjects are processing. For example, scan paths 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).

Before eye-fixation research can make substantial contributions to the study of human problem solving, the durations of the processes of interest (generally several seconds long) must be brought into correspondence with the durations of fixations (generally a fraction of a second). This can be done in one of two ways. Either the behavior can be segmented into meaningful units whose durations are of the same order of magnitude as a few fixations (Tichomirov & Posnyanskaya, 1966), or the fixations can be aggregated into higher order units (Winikoff,

1967). When either of these research strategies is used, the internal processes in problem solving are amenable to an eye-fixation analysis.

### Processing Spatial Information

There may be some mental operations that are specialized for processing spatial information, and these operations should be especially susceptible to an eye-fixation analysis. For example, operations having to do with the size, position, or orientation of an object seem inherently spatial. One of these operations, on orientation, has received considerable attention as a result of the elegant experiments and theories of mental rotation developed by Roger Shepard, Lynn Cooper, and Jackie Metzler. Their research has shown that the time taken to determine whether two objects are the same is a monotonically increasing function of the angular disparity between their orientations. The increase in response latency with angular disparity has been attributed to the time it takes to mentally rotate the two figures into congruence (Cooper & Shepard, 1973; Shepard & Metzler, 1971).

Using the Shepard and Metzler task, Just and Carpenter (1976) were successful in decomposing the psychological processes into three successive stages. In the first stage, subjects searched for segments of the two figures that looked like they might correspond to each other. They looked for segments of the same length or segments that pointed in the same direction, and they rotated the corresponding segments into alignment. During the rotation stage, the subjects switched fixation back and forth between the two corresponding segments, with one additional switch (and 800 msec) for each additional 50 deg difference in angular disparity. We interpret this result to mean that the subjects were rotating in steps of 50 deg and monitoring their rotation with a comparison after every 50-deg rotation. In the third stage, subjects determined whether the rotation operation that had brought the two corresponding segments into congruence would also bring the rest of the figure into congruence. This study showed that the operations on the spatially distributed information are revealed by the eyes' scan path. This preliminary effort invites investigations of other spatial operations.

### Processing Real-World Scenes

Although processing real-world scenes seems, at first glance, a natural domain for eye-fixation research, it presents some formidable obstacles. One major difficulty is that there are no adequate scene analyses to provide a "grammar" of pictures. One possible research approach is initially to study the scanning of highly simplified scenes whose "grammatical" structure may be transparent. An alternative is to search for human psychological processes that may resemble computer processes in machine recognition of scenes (cf. Guzman, 1968). In

either case, it will be necessary to work within a well-structured task environment, so that the subject and the experimenter know what they are looking for in the picture and why. Such research might determine the specific information that people extract from a picture, how they operate on this information, how they store it, and how they later recognize the picture. As yet, the data relating fixation duration to recognition memory (cf. Buswell, 1935; Loftus, 1972; Mackworth & Bruner, 1970; Mackworth & Morandi, 1967; Noton & Stark, 1971; Potter & Levy, 1969; Tversky, 1974) are insufficient to construct a complete model of the mental processes in picture encoding, memorizing, and retrieval.

### REFERENCES

- 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.
- CARPENTER, P. A., & JUST, M. A. Semantic control of eye movements during picture scanning in a sentence-picture verification task. *Perception & 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.
- 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.
- DE GROOT, A., & JONGMAN, R. *Perception and memory in chess: An experimental analysis of the master's professional eye*. RITP Memorandum No. 24, University of Amsterdam, Holland, 1973.
- GOULD, J. Eye movements during visual search and memory search. *Journal of Experimental Psychology*, 1973, 98, 184-195.
- GUZMAN, A. *Computer recognition of three-dimensional objects in a visual scene*. Unpublished doctoral dissertation, M.I.T., 1968.
- HOCHBERG, J. Components of literacy: Speculations and exploratory research. In H. Levin & J. P. Williams (Eds.), *Basic studies on reading*. New York: Basic Books, Inc., 1970.
- JAVAL, E. Essai sur la physiologie de la lecture. *Annales d'oculistique*, 1878, 79, 97.
- JUST, M. A., & CARPENTER, P. A. Comprehension of negation with quantification. *Journal of Verbal Learning and Verbal Behavior*, 1971, 10, 244-253.
- JUST, M. A., & CARPENTER, P. A. Eye fixations and cognitive processes. *Cognitive Psychology*, 1976, in press.
- KOLERS, P. A. Three stages of reading. In H. Levin & J. Williams (Eds.), *Basic studies on reading*. New York: Basic Books, 1970.
- LEVIN, H., & KAPLAN, E. Grammatical structure and reading. In H. Levin & J. P. Williams (Eds.), *Basic studies on reading*. New York: Basic Books, 1970.
- LOFTUS, G. R. Eye fixations and recognition memory for pictures. *Cognitive Psychology*, 1972, 3, 525-551.
- MACKWORTH, N. The line of sight approach to children's reading and comprehension. In S. Wanat, H. Singer, & 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.
- 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, & 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, 1972.
- NOTON, D., & STARK, L. Eye movements and visual perception. *Scientific American*, 1971, **224**, 34-43.
- 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.
- RUSO, J., & ROSEN, L. An eye fixation analysis of multi alternative choice. *Memory & Cognition*, 1975, **3**, 267-276.
- SHEPARD, R., & METZLER, J. Mental rotation of three dimensional objects. *Science*, 1971, **171**, 701-703.
- SIMON, H. A., & BARENFIELD, M. Information processing analysis of perceptual processes in problem solving. *Psychological Review*, 1969, **76**, 473-483.
- 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.
- TVERSKY, B. Eye fixations in prediction of recognition and recall. *Memory & Cognition*, 1974, **2**, 275-278.
- WINIKOFF, A. *Eye movements as an aid to protocol analysis of problem solving behavior*. Unpublished doctoral dissertation, Carnegie-Mellon University, Pittsburgh, Pa., 1967.