

## Dynamics of Automatic and Controlled Visual Attention

ERICH WEICHSELGARTNER AND GEORGE SPERLING

**The time course of attention was experimentally observed using two kinds of stimuli: a cue to begin attending or to shift attention, and a stimulus to be attended. Precise measurements of the time course of attention show that it consists of two partially concurrent processes: a fast, effortless, automatic process that records the cue and its neighboring events; and a slower, effortful, controlled process that records the stimulus to be attended and its neighboring events.**

**A**N ESSENTIAL COMPONENT OF VISUAL attention is attentional gating (1), the process whereby some incoming information is selected for further analysis or for memorization, while other information is ignored or attenuated and lost (2-8). Normally, eye movements and attentional shifts are tightly coupled. In this report we discuss the attentional processes that can occur while the eyes are stationary (9, 10).

Early in our investigations of the dynamics of attentional gating, we noticed curious bimodal distributions of attention shift times that suggested that we were observing not merely a single act of attention but two consecutive, partially overlapping acts. During the investigation of these phenomena (11), we learned how to attain separate, and almost independent, control of the time course of each attentional process. The first process is a quick, effortless, automatic process triggered by target detection; the second is a slower, effortful, controlled process (12) whose latency depends on practice and task difficulty.

The procedure we used to measure attention was based on the rapid serial visual presentation (RSVP) attention shift paradigm of Sperling and Reeves (13). RSVP (14, 15) produces many regularly spaced events that can be used to resolve the temporal properties of attentive processes. In the Sperling and Reeves procedure, a subject monitored a stream of alphabetic characters flashed one on top of the other to the left of visual fixation for the presence of a target letter. On detection, the subject immediately shifted attention to a numeral stream at visual fixation and attempted to remember the earliest occurring numeral from the numeral stream. Typically, subjects remembered a numeral that occurred 300 to 400

msec after the target. The onset time of a remembered numeral defined the attention-reaction time. When target difficulty or target probability was altered, the attention-reaction time varied much the same way as the motor-reaction time (13).

In this investigation we studied an additional attentional process that had completely different properties, using two procedures. In both procedures, the subject maintained visual fixation on a numeral stream throughout the trial. Procedure 1 (Fig. 1a), an adaptation of the RSVP attention shift paradigm, served as a control condition. Subjects began a trial by attending to a letter stream to the left of the numeral stream. They were instructed to shift attention from the letter stream to the numeral stream as quickly as possible on detection of the target letter C and to remember the earliest four numerals they could (simultaneous with or subsequent to target detection). In procedure 2 (Fig. 1b), the experimental condition, there was only one stream. Subjects fixated a numeral stream in which the target was embedded. Two kinds of targets were used to open the attentional gate: (i) outline square condition (OS), an outline square around one of the numerals, and (ii) high-lighted numeral condition (HN), a brightening of one of the numerals.

For each trial of both procedures, the

target was randomly embedded near the middle of a stream of 25 numerals. At the end of stimulus presentation, subjects typed the earliest occurring numeral they could remember as well as the three subsequent ones. The four recalled numerals comprised the response sequence. This recall task enabled us to probe the distribution of attention over a longer time period [and at no additional cost (13)] than would be possible with just one recalled numeral.

Over 10,000 trials were conducted with three subjects. Representative data from all three are shown (Fig. 2, a-i). Recall data were analyzed in terms of the estimated probabilities  $p(t_j, k)$  of reporting a numeral from frame  $j$  (time  $t_j$ ) in position  $k$  of the response. The target occurred at time  $t_o = 0$ . We first considered  $p(t_j)$ , the probability of reporting a stimulus item from time  $t_j$  anywhere in the response (the envelope curves of Fig. 2):

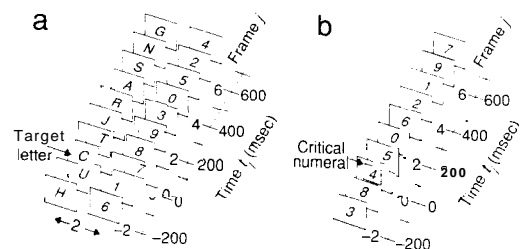
$$p(t_j) = \sum_{k=1}^4 p(t_j, k)$$

The basic finding from the control experiment is shown in Fig. 2a. The numeral simultaneous with the target and the one following it were almost never reported. Reported numerals occurred 200 to 700 msec after the target. The trace of the first reported numeral showed that 50% of the time it was the numeral occurring 300 msec after the target; the remainder of the time it was the numeral occurring 200 or 400 msec later. In all significant respects, these data replicated the data of Reeves and Sperling (1). They indicated that detecting the target C in the letter stream to the left of visual fixation and shifting attention to the numeral stream at visual fixation takes  $300 \pm 100$  msec.

Typical data from procedure 2 (target OS) are shown in Fig. 2b. The distribution of the  $p(t_j)$  clearly is bimodal. Most numerals were recalled from two separated periods in time, the first mode between 0 and 100 msec after target presentation, the second mode between 300 and 400 msec after target presentation. For all subjects, the

**Fig. 1. (a) Illustration of procedure 1 (control experiment) and (b) procedure 2 (main experiment). On the face of a cathode-ray tube, 25 frames, composed of letters and numerals in procedure 1 or numerals alone in procedure 2, appear in serial order. Presentation rate is usually 10 frames per second and sometimes 12.5 frames per second; frame duration is 18 msec. With**

**equal probability, 7 to 14 stimulus frames precede the target, and the subject's task is to attempt to report the four earliest occurring numerals simultaneous with or subsequent to the target. In order to avoid ambiguity in the analysis of the response sequence, numerals in frames -2 to +7 are all different.**



Human Information Processing Laboratory, Psychology Department, New York University, New York, NY 10003.

recall probability of the critical numeral (the numeral displayed simultaneously with the target) was well above 0.9. Indeed, all subjects reported that they could achieve high recall of the critical numeral effortlessly; it seemed to "pop out" of the background numerals, frequently enhancing the subsequent numeral as well. From this high level of response (and from more elaborate statistical tests) it was obvious that the bimodality did not result from a mixture of early-mode and late-mode trials: a single, bimodal distribution characterized all the responses.

Subjects claimed to be able to distinguish introspectively between numerals that they remembered occurring with the target and those that seemed to have occurred subsequent to the target (in fact, these sometimes occurred earlier). An extended response procedure was developed to utilize this infor-

mation. After reporting the four numerals of the response sequence, subjects further reported whether each one seemed to be perceptually attached to the target. We refer to items judged by subjects to be attached to the target as first glimpse, the others as second glimpse.

Subjects had no particular difficulty making first glimpse-second glimpse judgments. First glimpse items were described as being effortlessly acquired; recall of second glimpse items required effort. Experiments with subjects E.W. and S.W. in which these judgments were obtained with a highlighted numeral are shown in Fig. 2, c and d. When the response data were obviously bimodal (E.W.), the glimpse judgment mirrored the bimodality. Indeed, for appropriate conditions, we were able to clearly observe bimodal response data for every subject. How-

ever, even when the data were not obviously bimodal (S.W.), the glimpse judgment still revealed the same two underlying processes.

When subjects were asked to make glimpse judgments in procedure 1 (detection of a C in a neighboring stream), they reported only second glimpse. The content of the first glimpse in procedure 1 would correspond to a memory of the target C. Indeed, when the entire response distribution of procedure 1 was normalized for comparison with the second glimpse distributions of procedure 2, the three distributions were seen to be essentially equivalent (Fig. 2e), which confirmed objectively what the subjects reported introspectively.

Decreasing the intensity of the outline square illustrated independent control of the two glimpse processes. A dim but clearly visible square left the first glimpse complete-

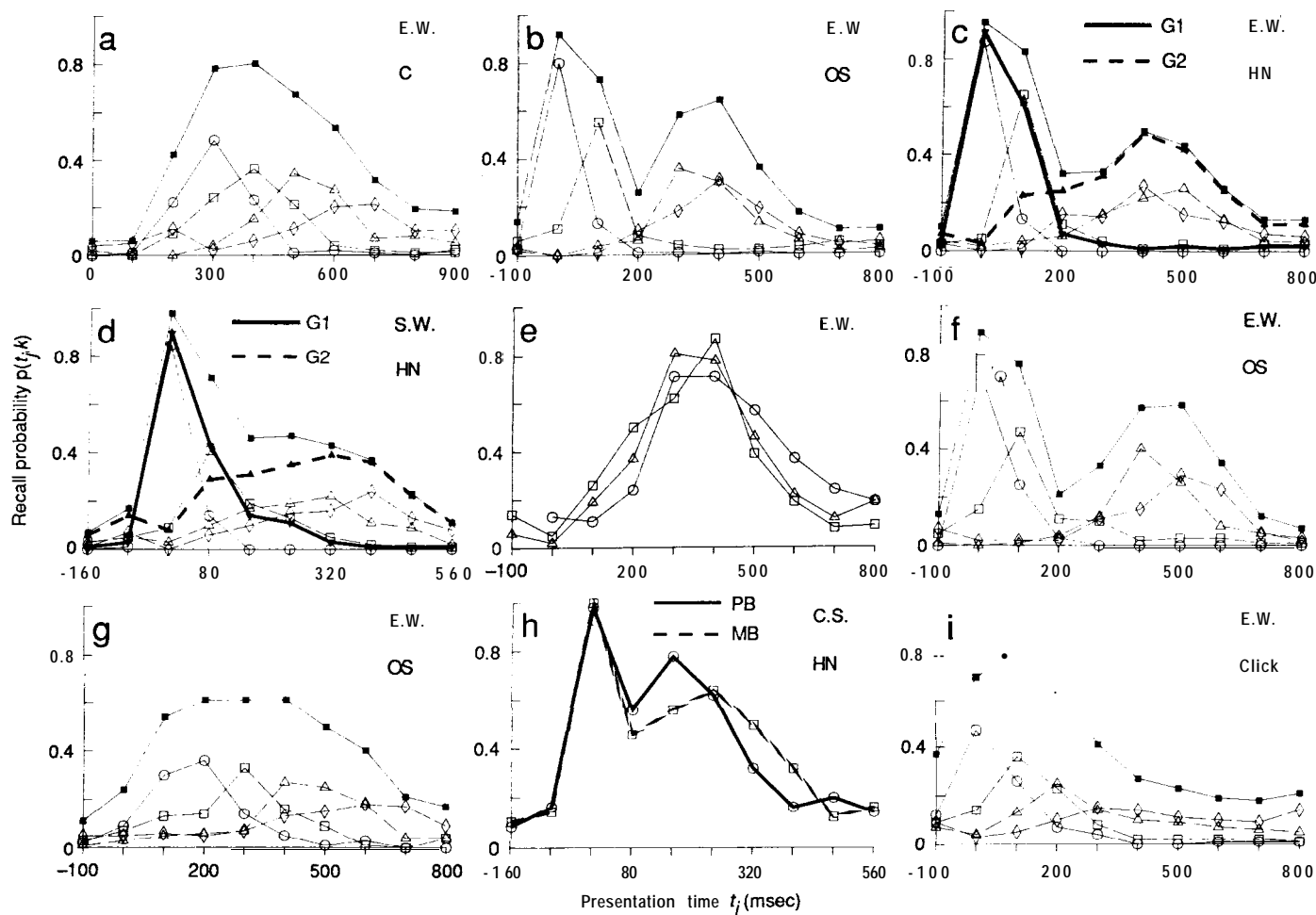


Fig. 2. Recall probability as a function of presentation time of numerals from the numeral stream; the target occurs at time 0. The three subjects were E.W., S.W., and C.S. The envelope curve  $\blacksquare$  describes the probability of recalling a numeral presented at time  $t$  anywhere in the response sequence. In panels a-d, f, g, and i, the inscribed curves give individual recall probabilities for each of the four responses [ $p(t_j, 1) \circ$ ,  $p(t_j, 2) \square$ ,  $p(t_j, 3) \triangle$ ,  $p(t_j, 4) \nabla$ ]. In panels c and d,  $\blacktriangledown$  and  $\blacktriangle$  represent glimpse 1 (G1) and glimpse 2 (G2) data, respectively. In panel e, the inscribed curves  $\circ$ ,  $\square$ , and  $\triangle$  are redefined ( $\circ$ , C;  $\square$ , HN; and  $\triangle$ , OS). Numeral presentation rates were 10 per second unless otherwise noted. To better display the data, abscissas vary between panels. (a) Target C, procedure 1. (b) Target outline square (OS), procedure 2. (c and d) Target highlighted numeral (HN), procedure 2 plus glimpse judgments. The heavy solid and dashed lines indicate items judged to have occurred in G1 and G2, respectively. The numeral presentation rate for S.W. was 12.5 per second. (e) Comparison of data from procedure 1, target C with G2 judgments from procedure 2, targets OS and HN. Curves are normalized to have the same area. (f) Target dimmed OS. Second peak is shifted +100 msec relative to bright OS. (g) Further-dimmed target OS. (h) Target HN, comparison of recall in pure blocks (PB) with mixed blocks (MB) of HN and OS. (i) Target auditory click.

ly intact but delayed the peak of the second glimpse by 100 msec (Fig. 2f). Further dimming of the outline square (to an intensity level at which the square was still always detected but motor reaction times were slowed by 65 msec relative to the full-intensity square) greatly altered first glimpse responses (Fig. 2g). Interposing a bright masking field composed of a dense dot matrix between successive numeral frames also destroyed the first glimpse but left the second glimpse relatively intact. These two manipulations dissolved the perceptual glue (16) binding the outline square to the interior numeral.

In the case of the highlighted numeral target, when we randomly varied the intensity of all the other numerals from frame to frame, and we made the highlighted numeral only somewhat more intense than the strongest of its competitors (a signal-to-noise manipulation), the first glimpse remained intact, but the second glimpse was delayed by about 100 msec. Another factor, level of practice, had strong effects on second glimpse responses but almost no effect on first glimpse responses.

Late in the stimulus sequence, before a target had occurred, the instantaneous probability that a target would occur became much greater than early in the sequence. Correspondingly, the subjects' expectancy of a target increased as the end of the stimulus sequence was approached. Like motor-reaction times, second glimpse responses sped up with this increased expectancy; first glimpse responses were unaffected. Another form of expectancy occurred in mixed-list stimulus presentations when the subject did not know which target to expect: OS or HN. Compared with pure blocks of OS or HN trials, the second glimpse mixed-list responses were later in the numeral sequence; the first glimpse responses were unaffected by this form of stimulus uncertainty (Fig. 2h).

We also used an auditory click as the stimulus for triggering an attentional gate. The distribution of numerals reported in response to the click stimulus is shown in Fig. 2i. The cross-modality data showed an appreciable proportion of recalls of the numeral that preceded the target by 100 msec. These early recalled stimuli indicated that, in the absence of competing visual stimuli, there was some memory for visual stimuli that preceded the attention trigger; they also indicated that an intermodality attention shift could be extremely quick. The click procedure demonstrated the feasibility of directly measuring the gating of visual inputs by auditory signals, a critical feature of the partial report procedure for measuring visual very short term memory (2) and of

the Reeves-Sperling theory of temporal order judgments (1).

Bimodal distributions of reports from a rapid numeral stream are explained in terms of two consecutive, partially overlapping attentive processes. A fast process (interpreted as a first glimpse by the subject) occurred automatically upon target detection and resulted in almost perfect recall of the target numeral as well as in occasional recall of a numeral immediately after the target. The automatic process was effortless, required no practice, and was virtually independent of task difficulty. For the subject to have recalled a numeral from the numeral stream in the automatic process required that his or her attention was not withdrawn to another location away from numerals (as in procedure 1) and that the target was sufficiently distinctive (for example, not too dim). With outline square targets (but not with highlighted numeral targets), the automatic process was destroyed by masking fields interposed between successive numeral frames.

An effortful, slow attentive process (interpreted as a second glimpse by the subject) mediated recall of numerals occurring more than 200 to 300 msec after the target. This was a controlled attentional process whose latency depended on factors that classically affect task difficulty: practice, expectation, stimulus probability in mixed-list designs, and target signal-to-noise ratio.

#### REFERENCES AND NOTES

1. A. Reeves and G. Sperling, *Psychol. Rev.* 93, 180 (1986).
2. G. Sperling, *Psychol. Monogr.* 74, No. 11 (whole No. 498) (1960).
3. M. Coltheart, *Percept. Psychophys.* 27, 183 (1980).
4. G. M. Long, *Psychol. Bull.* 88, 785 (1980).
5. M. L. Shaw and P. Shaw, *J. Exp. Psychol.: Human Percept. Perform.* 3, M1 (1977).
6. J. A. Swets, in *Varieties of Attention*, R. Parasuraman and D. R. Davies, Eds. (Academic Press, New York, 1984), pp. 183-242.
7. D. E. Broadbent, *Psychol. Rev.* 64, 205 (1957).
8. A. Treisman, *ibid.* 76, 282 (1969).
9. R. M. Shiffrin, in *Stevens' Handbook of Experimental Psychology*, R. C. Atkinson, R. J. Herrnstein, G. Lindaeu, R. D. Luce, Eds. (Wiley, New York, cd. 2, in press).
10. G. Sperling and B. A. Doshier, in *Handbook of Perception and Performance*, K. Boff, L. Kaufman, J. Thomas, Eds. (Wiley, New York, 1986), vol. 1, pp. 2.1-2.65.
11. E. Weichselgartner, *thesis*, New York University, (1984).
12. R. M. Shiffrin and W. Schneider, *Psychol. Rev.* 84, 127 (1977).
13. G. Sperling and A. Reeves, in *Attention and Performance*, R. Nickerson, Ed. (Erlbaum, Hillsdale, NJ, 1980), vol. 8, pp. 347-360.
14. M. C. Potter and E. I. Levy, *J. Exp. Psychol.* 81, 10 (1969).
15. G. Sperling, J. Budiansky, J. G. Spivak, M. C. Johnson *Science* 174, 307 (1971).
16. A. M. Treisman and G. Gelack, *Cog. Psychol.* 12, 97 (1980).
17. The experiments described herein were submitted by E.W. in partial fulfillment of the requirements for the Ph.D. degree from New York University. The experimental work was performed at the Human Information Processing Laboratory, New York University supported by U.S. Air Force Office of Scientific Research, Life Science Directorate, grant 85-0364.

7 May 1987; accepted 8 September 1987

Erich Weichselgartner and George Sperling.  
Dynamics of Automatic and Controlled Visual  
Attention. *Science*, 1987, 238, 778-780.