Psychol Res (1987) 49: 221–227

Psychological Research

Spatiotopic and retinotopic components of iconic memory

Ken McRae, Brian E. Butler, and Stephen J. Popiel

Department of Psychology, Queen's University, Kingston, Ontario, K7L 3N6, Canada

Summary. Recently, there has been considerable interest in whether information in iconic memory is stored in retinotopic or spatiotopic coordinates. The present experiment examined the issue using a masking paradigm. In one set of conditions, subjects maintained fixation while a row of four letters appeared for 19 ms, centered about three degrees to the right of fixation. After a 153 ms ISI, the letters were followed by a blank field (no mask), a mask in the same position as the letters, or a mask displaced three degrees to the right of the letters. In a second set of conditions, the stimuli were the same but subjects were asked to shift fixation from the fixation point to the middle of the letter row during the interval between the letters and the mask. Subjects' eye movements were monitored in all conditions. Accuracy of report for the letters was lowered only with the mask over the letters with the no-eye-movement conditions and in both masking conditions with the eye movements. The results suggest that the icon includes two components, one that is retinotopic and one that is spatiotopic.

In 1960, Sperling used partial report to demonstrate a short-term visual information store that has since been renamed *iconic memory* (Neisser, 1967). Sperling found that subjects seemed to have more information available than they could report from a brief visual display. Subjects were shown three rows of four letters each and could report about 4.3 letters when asked to report as many as possible. However, when the display was followed by a high, medium, or low tone that served as a cue to report the upper, middle, or lower row, subjects could report the cued row with much greater accuracy, approximately 76%, than when they attempted to report the entire display. Sperling concluded that subjects had an average of 9.1 of the 12 letters available for a brief period following the offset of the stimulus. This memory appeared to consist of "a rapidly fading, visual image of the stimulus", which lasts less than 1 s. Sperling also argued that the stored information was precategorical, although other evidence suggests that this may not be the case (Mewhort and Butler, 1983; Butler 1974).

Independent of Sperling's research, Averbach and Coriell (1961) investigated visual short-term storage using a bar-probe task to study the temporal properties of the store. A 2 by 8 matrix of letters was presented and a bar, following at different time intervals, cued the one letter that was to be reported. Accuracy of report decreased as the cue was delayed, reaching asymptote at about 250 ms. They also noted a phenomenon, which they called "erasure"; during this 250 ms interval, an individual element of the matrix could be masked by a character that appeared in the same place as the original element, making the original very difficult, if not impossible, to report. Averbach and Coriell concluded that the visual system includes a storage buffer which has a very brief duration and a relatively slow read-out time. The duration is about 250 ms and new information introduced during this time has a local masking effect on what was previously stored.

While the tasks used by Sperling and by Averbach and Coriell are the principal measures of iconic memory, a third technique was introduced by Eriksen and Collins (1967) and further refined by Di Lollo (1977). Di Lollo's experiment involved the integration of two flashes, each consisting of 12 dots, which formed a 5 by 5 dot matrix with one missing element. He found that, with stimulus onset asynchronies (SOA) of up to approximately 100 ms, the missing dot could be localized very easily, presumably because the first display was held in iconic memory and then integrated with the second display.

While there is a general consensus about the existence of iconic memory, there is far less agreement about its function. It has been suggested that the icon simply serves as a continuation of the original stimulus (Haber, 1983). If this is the case, it would play a minimal role in vision under normal circumstances. Neisser (1976, p. 48) has argued that:

> The icon simply simulates, for the rest of the nervous system, the information that would be picked up if the real display were still on ... it can play little part in normal vision: by definition it does not exist while a given fixation continues, and it is destroyed by masking after every eye movement. Although the exact retinal arrangement of still unperceived forms may be briefly stored under tachistoscopic conditions, the storage is not robust enough to affect the perceptual cycle.

The icon, as described by Neisser, would require local, retinal coordinates only. However, the icon may serve to integrate across saccades, and this would necessarily involve a storage buffer that functions in the coordinates that correspond to real space. Feldman (1985), for example, has proposed that the visual system includes both a retinotopic buffer and a spatiotopic buffer.

Offprint requests to: B. E. Butler

Davidson, Fox, and Dick (1973) developed an experimental procedure to separate retinotopic and spatiotopic components of iconic storage. A horizontal row of five letters (e. g. SHVRY) was used as the principal stimulus. There were two fixation points, one located below the spot where the *H* would occur and the other located below the point of the *R*. Subjects were trained to alternate their fixation between these two points. When the speed of the fixation shifts reached 70 ms, a trial would begin. At the point in time when the subject fixated on the left fixation point, the row of letters would be presented for 10 ms. The test apparatus was designed so that when the subject's gaze reached the right fixation point, a mask (or, on some trials, a metacontrast ring) appeared over a specific letter position, for example, where the *Y* had once been. The subject's task was to report all possible letters as well as the apparent location of the mask or ring. It was then possible to determine whether the letter that was masked was the original letter in the display position occupied by the mask or the letter occupying the same retinal position as the mask. The results showed retinotopic backward masking (the *V* could not be reported), but, the mask was perceived to share its correct spatial location (at the *Y*) with the letter in the same display position. These results suggest the existence of two icons: one arranged retinotopically and one spatiotopically.

There are problems with the Davidson et al. study. Extensive training was required to reduce eye movements to within 70 ms. Hence, two of the authors served as the only two subjects in the study. Since they were familiar with the display parameters, the report of the position of the mask may have been biased by knowledge of where the mask actually occurred. Phenomenological report of the mask's position was used as the evidence for a spatiotopic component to iconic memory. Some uncertainty about these reports must have existed because the authors acknowledged that additional data were collected "... in the vain hope of obtaining clear results from each subject" (p. 113). Van der Heijden, Bridgeman, and Mewhort (1986) pointed out that no experiment has attempted to replicate the Davidson et al. results using a masking paradigm. A replication is needed, especially in view of the renewed interest in the storage characteristics of iconic memory, and the Davidson et al. study in particular.

At least two studies have attempted to examine spatiotopic coding with the icon. Jonides, Irwin, and Yantis (1982) asked subjects to locate the missing dot in a 5 by 5 dot matrix, which was displayed using two discrete flashes of 12 dots each. In the saccade condition, the first 12 dots were presented below a fixation point. An eye movement to a new fixation point ensued and all but one of the remaining 13 dots were presented. These overlapped spatially with the first portion of the array, but not retinally, because of the intervening eye movement. The control condition involved no eye movements. The disparate retinal locations of the flashes were retained by displaying the two portions of the array at different spatial locations on the screen. Responses were significantly greater than chance only in the saccade condition, making spatial superimposition the apparent important factor for visual integration.

There was a confounding factor in the Jonides et al. study. These researchers used a P-4 phosphor display device in which the dots decayed to 1% of their original brightness in less than 1 ms. Nevertheless, this 1% will remain for periods extending over a full second. All attempts to replicate the effect using equipment that does not suffer from this stimulus persistence problem have failed (Bridgeman and Mayer, 1983; Jonides, Irwin, and Yantis, 1983). It seems quite likely that the residual brightness created a situation where spatiotopic integration, because of temporal overlapping of the stimuli, was easy for the subjects. Finally, no condition that could have produced retinotopic integration was included.

In the same year, Breitmeyer, Kropfl, and Julesz (1982) conducted a very similar study. Subjects were asked to report whether or not a dot was missing from a 4 by 4 array. On a random half of the trials one element was missing. The arrays were presented in two flashes, either seven or eight dots per flash. In one condition, the first half was shown parafoveally, and the second half, after a saccade, foveally. In this case, the stimuli were presented in the same spatial location but in different retinal locations, and this was the only condition that yielded results significantly better than chance. When subjects were to integrate the arrays retinotopically, that is, seeing two flashes to the fovea at different locations on the display device, they were unable to do so. Again, there were some problems with the methods. Breitmeyer et al. also used a P-4 phosphor display device, with no control for persistence, and did not monitor the subjects' actual eye movements. Perfect control of eye movements would be critical for their procedure. If the fixations were not precise relative to the stimuli, or were not performed in the correct time interval, subjects would have no chance to integrate the images retinotopically.

The ideal paradigm for investigating the issue should be equally sensitive to the possibilities of a retinotopic or a spatiotopic icon. Davidson et al. used a procedure that was adequate but involved a complex training procedure to ensure proper eye movements. Hence, they were able to collect data from only two subjects, both familiar with the task, and the only data they reported on spatiotopic integration was phenomenological reports dealing with the apparent location of the mask. The dot studies, cited above, share a different problem. Di Lollo used a procedure similar to that of Jonides, Irwin, and Yantis and Breitmeyer, Kropfl, and Julesz, and found that if the SOA between the two arrays was of a longer duration than 160 ms, locating the missing dot was impossible. Both Jonides et al. and Breitmeyer et al. used delays greater than 200 ms, which suggests that the integration obtained in these studies was more likely due to residual phosphor persistence with the display device than to iconic memory.

In the present experiment, the subjects fixated on a point while four consonants were presented to the right of fixation for 19 ms. In the no eye movement segment of the experiment, 153 ms after the letters had disappeared, a mask in the same location as the letters, or a metacontrast control (a mask to the right of where the letters had been) was presented. In a third condition, no mask was involved. In the eye movement segment, after the letters had disappeared, subjects switched fixation to the point in the middle of the letter row. Following a 153 ms delay, a mask appeared in either the same spatial location as the consonants, or in the same retinal location (i. e., in a position to the right of the new fixation point). In a third condition, trials did not contain a mask at all. In total then, there were six types of trials. Each subject was asked to report the letters in a left-to-right order. In the no eye movement condition, the mask matched the actual and the retinal position of the letters and should have caused a decrement in performance as compared to the no-mask condition and the metacontrast control with the mask to the right. The question of interest is, when an eye movement is carried out during the interval between the letters and mask, will there be spatiotopic masking, retinotopic masking, or both?

Method

Subjects. Twelve undergraduate students at Queen's University served as subjects in this study. Two subjects were unable to complete the task and were replaced. All had normal uncorrected vision.

Apparatus. The stimulus in each trial consisted of a horiontal row of four uppercase consonants with no repeated items. They were displayed on a Tektronix point-plot display monitor (model 604), equipped with a P-4 phosphor, which was controlled by a Digital PDP-11/23 minicomputer. The monitor was located in a partially darkened room adjacent to the room that housed the computer. Each letter was defined in a 5 by 7 matrix and subtended a visual angle of 15 min by 21 min. The four letters subtended about 1° horizontally. The basic display algorithm is described by Mewhort (1978). It was altered by the experimenter in order to allow for collection of eye movement latency data. A Biometrics SGH/V-2 eye movement monitor was used to track eye movements and was connected to a Schmitt trigger on the computer clockboard.

A small pilot study was conducted in order to find a luminance level at which 1% of the original brightness (the P-4 persistence mentioned earlier) would have no effect on performance. Three judges viewed the letter display through a Kodak N. D. 2.00 gelatin filter that allows only 1% of the brightness to pass. When the luminance was set at the level used in the main experiment, subjects were unable to identify any items in the display. Throughout the experiment, the luminance of the letter stimuli was maintained at two log units above threshold in order to control for residual phosphor persistance, as noted by Bridgeman and Mayer (1983).

Procedure. Subjects were seated in a slightly darkened room and fitted into the eye movement monitor apparatus directly in front of the display screen. The subject's head was strapped into a harness in order to ensure that head movements did not occur. There were six blocks of trials (eye movement vs. no eye movement combined with mask type). Half of the subjects participated in the eye movement blocks first while the other half began with the no eye movement blocks. The trial blocks were arranged so that each block was carried out in each ordinal position twice. Subjects received 30 practice trials before participating in the eye movement or no eye movement segments of the experiment.

Typical eye movement trials proceeded as follows. The subjects focused on the fixation point. Following the depression of the start button by the subject, the four consonants appeared in the right portion of their visual field. Immediately following the disappearance of the letters, the subject executed a saccade to the point in the center of where the letters had appeared. This saccade was approximately 3° in arc. The letters were followed by no mask or, after an interval of 153 ms, by either the spatiotopic mask or the retinotopic mask. The mask consisted of six number signs (#) and extended one character beyond each end of the row of letters. The mask remained on the screen for approximately 500 ms. Figure 1 shows a diagram of each of the three types of trials. The subject was asked to report the letters that were perceived in a left-to-right order. If a subject could not report a specific letter, he or she was asked to indicate the missing position so that position data was correctly collected for the reported items. Subjects were encouraged to report four letters on every trial. The no eye movement trials proceeded in a similar fashion but the subject's fixation remained on the initial fixation point throughout the trial. Fifteen trials with proper eye movement (or no eye movement) times were required for each trial block. Subjects continued until this number was obtained. In the case of the no eye movement trials, a time that indicated a complete lack of movement until after the mask offset was necessary for a trial to be considered valid. With eye movement trials, an eye movement time that indicated movement during the 153 ms interval between the offset of the letters and the onset of the mask was required for a valid trial. Subjects were given a 15 min break between the eye movement and no eye movement sessions.

Results

Having subjects keep their eyes steady on the fixation point for the duration of a no eye movement trial was not difficult. Subjects could carry out this part of the task with

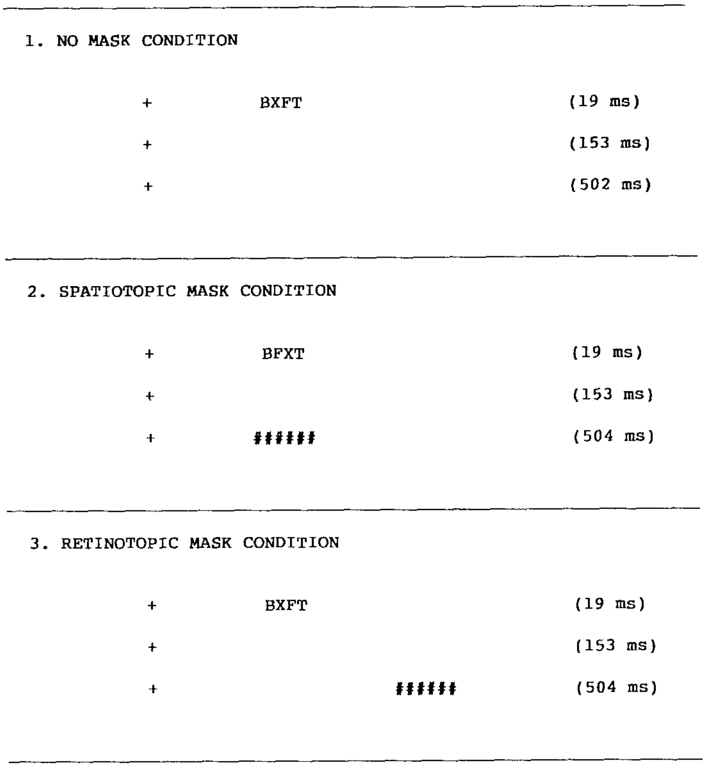

Fig. 1. An example of each of the three trials types. The labels given here correspond to the eye movement trials. In the case of the no eye movement trials, 2. would be labelled "Spatiotopic/Retinotopic Mask" and 3. would be labelled "Metacontrast control". The character "#" was used as the mask character

almost perfect accuracy. However, an eye movement that must be conducted within 153 ms was not quite so simple. On the average, approximately 55 trials were needed to obtain the required 15 trials for each condition. Furthermore, two subjects, whose data were not included, were unable to complete the task and were replaced.

The collected data were analyzed using three different measures: total number of letters correctly reported, number of letters reported in their correct position, and Kendall's tau, the correlation between stimulus order and response order. The last score relates specifically to order of report and will be discussed in detail below. The first two measures were analyzed using a two-factor analysis of variance to assess the effects of position in display for each of the six experimental conditions. The six experimental conditions could not be considered as two separate factors, that is the presence or absence of eye movements and mask type, because the resulting design would not be completely orthogonal. That is, in the no eye movement segment, there were essentially two no mask conditions and one mask condition while, in the eye movement segment, the reverse was true. The six presentation conditions, therefore, were treated as one factor in the analysis of variance and specific conditions were compared using orthogonal contrasts.

Total letters reported. The mean number of correct responses per subject for each condition in each target position can be found in Table 1. One should note that this report measure is not sensitive to order of report.

The results in Table 1 show that accuracy of report was substantially affected by position in display, $F\,(3,33) = 35.25$, $P < 0.0001$. The first letter was reported most often, the fourth slightly less often, and finally, the second and third letters were not reported as frequently. There was a significant quadratic trend in the data across the four letter positions, $F(1,11) = 100.98$, $P < 0.0001$, as well as a linear component to the trend, $F(1,11) = 12.80$, $P < 0.005$.

More importantly, the results in Table 1 show that accuracy of report varied across presentation conditions, $F\,(5,55) = 3.34$, $P < 0.02$.. A significant overall masking effect was found when the two no mask conditions plus the metacontrast control were compared to the three mask conditions, $F(1,11) = 17.43$, $P < 0.002$; in the former, the mean probability of a correct response was 0.48 but this dropped to 0.42 in the latter conditions. In the no eye movement trials, the mean probability of a correct response was 0.49 for the no mask condition and 0.46 for the metacontrast control, which did not differ significantly, $F = 1.96$, $P = 0.19$, but dropped to 0.40 for the mask condition, which differed significantly from the other two conditions, $F\,(1,11) = 26.50$, $P < 0.001$. In the eye movement trials, the results approached significance when the two masking conditions were compared to the no mask condition, $F(1,11) = 4.71$, $P = 0.051$. The mean probability of a correct response was 0.42 for the two mask conditions and 0.48 for the no mask condition. The number of correct responses was smaller in the spatiotopic mask condition than in the no mask control, a result which approached significance with an F-test, $F(1,11) = 4.53$, $P = 0.054$, and reached significance with a one-tailed t-test, $t(11) = 2.01$, $P < 0.05$. The decrement in the number of correct responses in the retinotopic masking condition, compared to the no mask control, approached but did not reach significance, $F\,(1,11) = 2.64$, $P = 0.13$, and $t(11) = 1.62$, $P < 0.1$. Finally, it should be noted that the number of correct responses given in the no eye movement control conditions were very similar to the data collected for the eye movement no mask condition, $F(1,11) < 1.0$.

Letters reported in correct position. The data were also analyzed in terms of the number of letters reported in their correct positions. These data, averaged across subjects, are shown in Table 2. Since masking affects both letter identification and order of report (Mewhort, Marchetti, Gurnsey, and Campbell 1984), this measure should be somewhat more sensitive to masking effects than the previous score for letters reported. It should be noted that all subjects reported four letters on every trial even though the instructions allowed omissions.

As with the total measure, position in display substantially affected report, $F(3,33) = 69.17$, $P = 0.0001$. The first letter was reported in its correct position most frequently, followed by the fourth letter, while the second and third letters were not correctly reported as often. The position in display effect was dominated by a quadratic trend, $F(1,11) = 327.36$, $P < 0.0001$, but was tilted to the right as evidenced by a linear trend, $F(1,11) = 34.96$; $P < 0.001$.

Figure 2 shows the data averaged across display position. As can be seen in the figure, mean accuracy of reporting letters in their correct positions differed across the six conditions, $F(5,55) = 3.22$, $P < 0.02$. An overall masking

Table 1. Mean number of letters reported

	Position in display				
	1	2	3	4	(mean)
No eye movements					
No mask	9.83	5.25	5.67	8.58	(7.33)
SP/RET mask	8.50	5.33	4.33	6.08	(6.06)
META control	9.75	5.00	5.50	7.33	(6.90)
Eye movements					
No mask	9.50	6.33	5.33	7.58	(7.19)
SPAT mask	9.42	4.83	4.83	6.92	(6.50)
RET mask	8.58	5.17	4.33	7.00	(6.27)

Note: The maximum possible correct is 15

Table 2. Mean number of letters reported in their correct position

	Position in display				
	1	2	3	4	(mean)
No eye movements					
No mask	8.17	2.67	1.17	4.67	(4.17)
SP/RET mask	6.67	2.33	1.50	2.58	(3.27)
META control	8.42	2.83	1.75	3.67	(4.17)
Eye movements					
No mask	8.33	2.67	1.67	4.08	(4.19)
SPAT mask	7.00	2.08	1.17	3.00	(3.31)
RET mask	7.00	2.25	1.42	3.42	(3.52)

Note: The maximum possible correct is 15

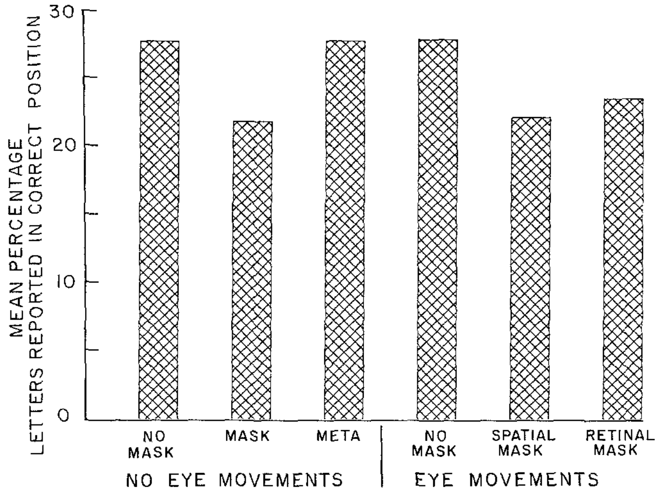

Fig. 2. Mean percentage of letters reported in their correct position for each of the six presentation conditions.

effect was again found, $F(1,11) = 25.24$, $P < 0.0007$. In the no eye movement trials, performance was equivalent in the no mask and metacontrast control conditions, $F(1,11) < 1.0$, but dropped substantially in the mask condition, $F(1,11) = 15.21$, $P < 0.003$. More importantly, in the eye movement trials, subjects reported more letters correctly in the no mask condition than in the two masking conditions, $F(1,11) = 8.60$, $P < 0.02$, and the two masking conditions did not differ significantly, $F(1,11) < 1.0$. A significantly smaller number of letters was reported in the correct position in the spatiotopic masking condition than in the no mask control, $F(1,11) = 6.01$, $P < 0.03$ and $t(11) = 2.45$, $P < 0.05$. Furthermore, when the retinotopic mask followed the letters, the decrement in report accuracy, compared to the no mask control, approached significance with an F-test, $F(1,11) = 3.57$, $P = 0.08$, and reached significance with a one-tailed t-test, $T(11) = 1.89$, $P < 0.05$. Again, there was no difference between the metacontrast control and no mask conditions with eye movements and the no mask condition without eye movements, $F(2,11) < 1.0$.

Order of Report. Kendall's tau was used to assess order of report by correlating the order of the responses with the actual order of the letters in the display. Butler and Currie (1986) have shown that order of report is a major factor with errors in tachistoscopic recognition and the differences between the scores in Table 1, and the scores in Table 2 show that order errors were common in the present task. The value of tau ranges from $+1$ (order of report matching the order in the display) to -1 (order of report opposite to order in display). It was calculated by taking all pairs of items reported and assigning a value of $+1$ to each correctly ordered pair and a value of -1 to each inverted pair of items. If less than two items were correctly reported in a trial, that trial would not enter into the tau score. The tau scores, averaged over subjects, are shown for each experimental condition in Table 3. A one-factor analysis of variance was carried out on this data.

Table 3. Mean tau scores for each condition

Experimental condition					
No EM			EM		
No mask	SP/RET mask	meta control	No mask	SPAT mask	RET mask
0.50	0.34	0.43	0.51	0.42	0.38

Overall, the presence of a mask caused some confusion with order of report. The mean tau scores for the no eye movement mask condition and the two eye movement mask conditions were significantly lower than the mean scores in the other conditions, $F(1,11) = 10.000$, $P < 0.01$. In the no eye movement segment of the experiment, the mean tau score was lower for the mask condition than for the no mask and metacontrast control conditions, $F(1,11) = 8.20$, $P < 0.02$. There was no significant difference between the no mask and metacontrast control conditions, $F(1,11) < 1.0$. In the eye movement portion of the experiment, the tau scores for the two masking conditions were lower than the tau score for the no mask condition but the difference did not quite reach significance, $F(1,11) = 4.09$, $P < 0.07$. The analysis showed no significant difference between the spatiotopic mask condition and the retinotopic mask condition, $F(1,11) < 1.0$. The difference between the no mask condition with eye movements and the retinotopic mask condition approached significance with an F-test, $F(1,11) = 3.37$, $P < 0.07$, and reached significance with a one-tailed t-test, $t(11) = 1.83$, $P < 0.05$. The difference between the no mask condition and the spatiotopic mask condition, however, did not reach significance, $F(1,11) = 3.05$, $P = 0.11$, $t(11) = 1.75$, $P < 0.1$

Discussion

Mewhort et al. (1984, p. 288) have pointed out that "masking provides prima facie evidence for iconic memory" and this is the strategy that we have used to determine whether information in the icon is coded retinotopically or spatiotopically. The present results demonstrate masking effects in three different conditions, one of which involves no eye movements. Masking in the no eye movement condition is hardly surprising; it has been found using a paradigm such as this in many different studies. Because the eyes remain stationary in this condition, masking is occurring in a situation in which the retinotopic and spatiotopic coordinate systems are completely confounded. In the other two masking conditions, however, an eye movement intervening between the stimulus and the mask effectively separates the two coordinate systems.

The present results show masking effects when the mask overlaps both spatiotopically and retinotopically with the stimulus. The statistical support for retinotopic masking is not as strong as the evidence for spatiotopic masking but the data are probably as good as any that can be obtained, given the delay of mask necessary to allow an eye movement. The two forms of masking seem to indicate the existence of two data representations, one arranged spatiotopically and another arranged retinotopically. The data gathered by Davidson et al (1973) also seemed to point to the existence of two icons. Just as in their experiment, masking in retinotopic coordinates was found while no subject reported any apparent movement of a mask.

Spatiotopic masking effects were not found by Davidson et al., possibly because of the time parameters. They used a 70 ms ISI, while the present study used an ISI of more than twice that, 153 ms. The two studies combined may indicate a transition from retinotopic to spatiotopic coding over time.

The above interpretation is based on the assumption that masking occurs only when the mask overlaps the position of the letters, either spatially or retinally. If the masking stimulus produces a decrement in performance regardless of where it occurs in the visual field, that is, if the decrement is simply produced by the onset of a second stimulus, then masking cannot be used to determine whether the information in the icon is represented in spatial or retinal coordinates. To ensure that the position of the mask was important, we incorporated the metacontrast control in the no eye movement trials and the results, especially those in Table 2, confirm that the position of the mask is crucial. In this condition, however, the mask was always peripheral to the letters. One reviewer (A. H. C. van der Heijden), has suggested, quite astutely, that metacontrast might still produce a decrement in performance if the mask was foveal and the letters were peripheral. These conditions would resemble those which we interpreted as yielding spatiotopic masking in the eye movement condition.

To examine the effects of foveal metacontrast, we ran a short experiment with 12 subjects using conditions similar to the no eye movement conditions in the present study. Subjects saw four letters, followed after 153 ms by either no mask, a mask at fixation, or a mask over the letter positions. Accuracy of reporting the letters was lower in the third condition, compared to the first two, $F(1,11) = 7.05$, $P < 0.02$, but there was no difference between the no mask and the foveal metacontrast mask, $F(1,11) < 1.0$. We are reasonably confident, therefore, that the decrement in performance caused by the mask occurred only when the mask was perceptually aligned with a representation of the target letters.

The claim that we have obtained both spatiotopic and retinotopic masking is also based on the assumption that subjects shifted fixation as instructed in the eye movement conditions. Given the nature of our equipment, we were able to record when an eye movement to the right occurred but we could not record the exact point of fixation. Under these circumstances, it is possible that subjects made some fixation errors by overshooting or undershooting their objective on some trials. The stimuli were designed to compensate for minor fixation errors because the mask extended one character beyond the letters at each end of the row. A more serious concern, however, is whether our interpretation could be negated by larger errors of fixation. Given the results, it seems unlikely that evidence for the two types of masking could be produced by either spatiotopic or retinotopic masking alone. If all masking were spatiotopic, the eye movements would have been irrelevant and we should have obtained the same results with or without eye movements; the results in Figure 2 show that this was not the case and that some aspect of masking is retinotopic. The relative weakness of our evidence for retinotopic masking, however, may be due to variability in fixation.[1]

Conversely, it seems unlikely that these results could be due to retinotopic masking alone. If the subject made an eye movement of more than 1°, the retinal trace of the letters would have moved beyond the position of the mask over the letter position and the mask would have had no effect (as our additional research showed). With an eye movement of less than 1°, some portion of the retinal trace would have aligned with the mask and this would have yielded an interaction between masking and serial position, that is the mask would have affected only the leftmost letters in the row. Given that we are certain that eye movements did occur, if all masking were retinotopic then the spatiotopic mask should have yielded a weak effect of masking and an interaction with serial position, which is not what we found. The present results are most compatible with the assumption that masking can occur in both retinotopic and spatiotopic coordinates.

The present results differ somewhat from the studies examining the effects of eye movements on the integration of dot patterns. These discrepancies may be due to the difference between visible persistence and informational persistence (Coltheart, 1980). Yeomans and Irwin (1985, p. 167) defined visible persistence as the "phenomenal visibility of a just-extinguished stimulus" while informational persistence is characterized by "knowledge about the visual properties of a just-extinguished stimulus". In the dot studies, temporal integration of visible persistence was measured. However, the length of time required to plan and execute an eye movement is long. In the dot studies mentioned above (Jonides et al., 1982; Breitmeyer et al. 1982), time parameters may have been too long for temporal integration to occur[2].

For some time, researchers have been attempting to segregate what was once collectively referred to as iconic memory into two different constructs or processes. A finding such as the present one obviously len[illegible]tself to this type of theorization. Coltheart (1980) proposed that a very important distinction is the one between visible persistence and informational persistence. With respect to the present experiment, it seems likely that the icon in retinal coordinates corresponds to a form of visible persistence that has much of its basis in neural persistence at or near the retina (Sakitt, 1976). On the other hand, a buffer providing informational persistence would involve spatial tags and, hence, essentially work in spatial coordinates. It may be responsible for spatiotopic masking in the present results. This may also correspond to the dichotomy between a feature buffer and a character buffer proposed by Mewhort, Campbell, Marchetti, and Campbell (1981). These authors have proposed a dual buffer system in which character identification is carried out at a feature buffer level, while these identified items and the spatial relations among them are stored in a character buffer. It seems reasonable that a buffer, with the essential function of holding information

---

[1] Ideally, this research should have been conducted using equipment that could record the exact point of fixation and provide input to the display routine so that the position of the retinotopic mask could be adjusted appropriately. Unfortunately, such equipment is beyond our resources and, we suspect, beyond the resources of many other experimenters

[2] Prior to the present experiment, we attempted to examine temporal integration using two groups of consonants separated by an intervening eye movement but were unable to get phenomenological integration across the time intervals needed for proper eye movements

for an identification mechanism, would have no need to reside in spatiotopic coordinates; retinal coordinates would certainly suffice. However, a buffer preserving spatial information would certainly reside in the coordinates of real space.

The visual system may include two buffers in order to form a system that can integrate images across saccades. Without a system to carry out this function, we could not possibly make sense of a world that we view via successive fixations. It is possible that the icon in retinotopic coordinates acts simply as a buffer in which the original visual information that enters the system is stored. At the time of a saccade, the spatiotopic buffer must become important. There must exist a mechanism that maps successive retinal images onto a nonvarying spatiotopic representation. It seems that the spatiotopic buffer that was revealed in this study would be fundamental to this process. Breitmeyer et al. (1982, p. 175) have stated that:

> Based on extant experimental data we tentatively identify two general types of visual persistence: one resides in activity along the afferent visual pathway and is retinotopically organized; the other resides at central levels and is spatiotopically organized. Moreover, whereas the former afferent persistence is eliminated via saccadic suppression mechanisms in order to separate successive, retinotopic frames of pattern information, the latter, central one, in contrast, is generated and enhanced via extraretinal signals accompanying saccades in order to preserve phenomenal continuity of a stable spatiotopic representation of the environment from one fixation to the next.

It is obvious that, since masking across saccades occurred in retinotopic coordinates, the "former afferent persistence" is not totally eliminated via saccadic suppression mechanisms. However, the present results certainly support the notion of a spatiotopic buffer that follows a retinotopic representation.

Acknowledgements. The present experiment was an honours thesis submitted by the first author in completion of the requirements for a B. Sc. (Cognitive Science) degree. This research was supported by grant A-9531 from the National Science and Engineering Research council of Canada to the second author. We would like to thank D. J. K Mewhort for comments on an earlier draft of this paper and for assistance with the programming of this study

References

Averbach, E., & Coriell, A. S. (1961). Short-term memory in vision. *Bell System Technical Journal, 40,* 309–328.

Breitmeyer, B. G., Kropfl, W., & Julesz, B. (1982). The existence and role of retinotopic and spaciotopic forms of visual persistence. *Acta Psychologica, 52,* 175–196.

Bridgeman, B., & Mayer, M. (1983). Failure to integrate visual information from successive fixation. *Bulletin of the Psychonomic Society 21,* 285–286.

Butler, B. E. (1974). The limits of selective attention in tachistoscopic recognition. *Candian Journal of Psychology 28,* 199–213.

Butler, B. E., & Currie, A. (1986). On the nature of perceptual limits in vision: A new look at lateral masking. *Psychological Research, 48,* 201–209.

Coltheart, M. (1980). Iconic memory and visible persistence. *Perception & Psychophysics, 27,* 183–228.

Davidson, M. L, Fox, M., & Dick, A. O. (1973). Effects of eye movements on backward masking and perceived location. *Perception & Psychophysics, 14,* 110–116.

Di Lollo, V. (1977). Temporal characteristics of iconic memory. *Nature, 267,* 241–243.

Eriksen, C. W, & Collins, J. F. (1967). Some temporal characteristics of visual pattern perception. *Journal of Experimental Psychology, 74,* 476–484.

Feldman, J. A. (1985). Four frames suffice: A provisional model of vision and space. *Behavioral and Brain Sciences, 8,* 265–289.

Haber, R. N. (1983). The impending demise of the icon: A critique of the concept of iconic storage in visual information processing. *Behavioral and Brain Sciences, 6,* 1–55.

Heijden, van der A. H. C., Bridgeman, B., & Mewhort, D. J. K. (1985). Is stimulus persistence affected by eye movements? A critique of Davidson, Fox and Dick (1973). *Psychological Research, 48,* 179–181.

Jonides, J., Irwin, D. E., & Yantis, S. (1982). Integrating visual information from successive fixations. *Science, 215,* 192–194.

Jonides, J., Irwin, D. E., & Yantis, S. (1983). Failure to integrate information from successive fixations. *Science, 222,* 188.

Long, G. M. (1980). Iconic memory: A review and critique of the study of short-term visual storage. *Psychological Bulletin, 88,* 785–820.

Mewhort, D. J. K. (1978). An n-channel tachistoscopic algorithm. *Behavior Research Methods and Instrumentation, 10,* 756–760.

Mewhort, D. J. K., & Butler, B. E. (1983). On the nature of brief visual storage: There never was an icon. *Behavioral and Brain Sciences,* 6, 31–33.

Mewhort, D. J. K., Campbell, A. J., Marchetti, F. M., & Campbell, J. I. D. (1981). Identification, localization, and "iconic memory": An evaluation of the bar probe task. *Memory & Cognition, 9,* 50–67.

Mewhort, D. J. K., Marchetti, F. M., Gurnsey, R., & Campbell, A. J. (1984). Information persistence: A dual-buffer model for initial visual processing. In H. Bouma & D. Bouhwuis (Eds.), *Attention and Performance 10* (pp. 287–298). New York: Erlbaum.

Neisser, U. (1967). *Cognitive psychology.* New York: Meredith.

Neisser, U. (1976). *Cognition and reality: Principles and implications of cognitive psychology.* San Francisco: W. H. Freeman.

Sakitt, B. (1976). Iconic memory. *Psychological Review, 83,* 257–276.

Sperling, G. (1960). The information available in brief visual presentations. *Psychological Monographs, 74,* (11, Serial No 498).

Yeomans, J. M, & Irwin, D. E. (1985). Stimulus duration and partial report performance. *Perception & Psychophysics, 37,* 163–169.

Received September 26, 1986/February 5, 1987