

Perceptual objects and the cost of filtering

A. TREISMAN, D. KAHNEMAN, and J. BURKELL
University of British Columbia, Vancouver, British Columbia, Canada

The latency of reading a single word is increased by 20 to 40 msec if another object is present in the display. The delay is affected by the spatial organization of the display: a colored frame causes less delay when it surrounds the word than when it is shown on the opposite side of fixation. A small gap in the frame is also more efficiently detected as a secondary task when the frame is around the word than when the two are spatially separate. The advantage of integrated over separate presentation suggests that a "filtering cost" is incurred when two distinct perceptual objects compete for attention. Attention in filtering tasks operates on perceptually distinct objects rather than on nodes in a semantic network.

The act of reading is usually assumed to be automatic, in the sense that it occurs both without voluntary control and without requiring attentional resources, if a word or letter is sufficiently clear and close to the fovea. Yet a consistent delay in naming a single letter is produced by the addition of irrelevant objects to the display, even when these objects are highly discriminable from letters (Eriksen & Hoffman, 1972; Eriksen & Schultz, 1978, p. 18). Distractors such as black disks or color patches add about 30 msec to the latency of letter naming. Eriksen and Schultz labeled this effect cognitive masking. We have found a similar delay in reading a single word when an irrelevant but highly discriminable object is added to the display. We further found that the delay in reading increases as more objects are added, although probably at a decreasing rate; it can be eliminated by precuing the location of the word on each trial, and it is reduced or eliminated, with the same displays, when the subject is asked to press a key whenever a word is shown, instead of to read it (Kahneman, Treisman, & Burkell, in press).

The fact that it disappears with precuing links the delay to attention rather than to peripheral interference. The fact that it is reduced in search or detection suggests that the delay arises not in *finding* the word but in allocating attention to it and/or in filtering out the irrelevant objects. Focused attention to the word is not required when the response is determined directly by the detection of its presence. We suggest attention must be narrowed down to the relevant stimulus, however, when the choice of a response demands further processing. We therefore interpret the observed delay of read-

ing as a filtering cost. The cost of filtering should be distinguished from interference caused by confusion between the relevant and irrelevant items and from interference that originates in response conflict: Filtering costs occur even with highly discriminable stimuli, or with irrelevant stimuli that cannot be read and are not easy to name. (Of course, both these factors could be sources of further interference if they were introduced.)

It has been suggested that selective attention to objects is more effective than attention to properties (Kahneman, 1973; Treisman, 1969), and that the allocation of attention to an object facilitates the processing of all its properties, irrelevant as well as relevant. Kahneman and Henik (1981) supported this hypothesis in several experiments. The subjects in one of their studies were shown two words on each trial, one printed in black and the other in color. They were instructed to name the color of the colored item, which appeared unpredictably on either side of fixation. A large Stroop effect was observed when the attended item was itself a color name, but there was little or no interference when the irrelevant black word was a color name. Evidently, the allocation of attention to the colored object facilitated all the responses associated to it, irrelevant as well as relevant. Similar results were reported by Francolini and Egeth (1980).

The present experiments related the notion of object-oriented attention to the phenomenon that we have called filtering cost. The subjects were to read a tachistoscopically presented word, and the latency of reading was measured. Filtering costs were induced by presenting a colored shape on the other side of fixation. The critical condition was one in which the word was framed by the colored shape instead of separated from it. We assumed that the word and the frame were likely to be integrated into a single perceptual object in such displays, an assumption anchored in the phenomenological observation that an object spatially en-

This research was supported by grants from the Natural Sciences and Engineering Research Council of Canada. Address reprint requests to Anne Treisman, Department of Psychology, University of British Columbia, #154—Main Mall, Vancouver, British Columbia V6T 1Y7, Canada.

closed in another tends to join with it in a single hierarchical or part-whole configuration. From the idea that competition for attention arises between perceptual objects rather than between the parts or properties of any one object, we derive the hypothesis that the frame will cause less interference with reading in the integrated displays than in the separate displays, in which the two objects appear on opposite sides of fixation.

Other theories might predict the opposite outcome, that the closer and harder to separate the irrelevant stimuli are, the greater the interference they should create. This would follow from accounts involving local sensory interference between relevant and irrelevant objects (Bjork & Murray, 1977; Estes, 1972), and also from accounts that emphasize the difficulty of pinpointing the relevant object within the spotlight or focus of attention (Eriksen & Hoffman, 1973).

It should be noted that the effects of integrated presentation of the relevant and irrelevant stimuli are in opposite directions in the Stroop design and in filtering costs, because the mechanism of interference is different in the two cases. Stroop interference arises because the color word evokes a conflicting response, which it does most strongly when it is part of the attended object. Filtering costs, on the other hand, are produced by disrupting the allocation of attention to the relevant object. Integrated presentation should therefore increase interference in Stroop experiments but reduce filtering costs.

EXPERIMENT 1: READING DELAY WITH SEPARATE OBJECTS

In Experiment 1, we compare the reading delay produced by a colored frame that is placed around the word or separated from it and on the opposite side of fixation. If the filtering cost is incurred when an irrelevant object competes to attract attention away from the relevant one, the separate frame should cause more delay than the integrated frame. If the delay results either from sensory interactions or from the difficulty of separating relevant from irrelevant stimulus information, the opposite prediction should hold.

Method

Stimuli. Stimuli were presented in a Cambridge two-field tachistoscope, at a background luminance of 4.0 mL. Each display contained one word typed with the Orator head of an IBM Selectric typewriter on a rectangular white adhesive label (1.76×1.04 deg). The words were four or five letters long and subtended 1.17 or 1.37 deg in length and 0.46 deg in height. One-third of the displays included only a word. The other displays also included one of the following shapes: a rectangle (3.58×2.15 deg), an elongated hexagon (3.91×2.47 deg), or an ellipse

(4.10×2.34 deg). The shapes were outlined in black and filled in pink, pale green, or yellow (using highlighter pens) averaging 3.0 mL in luminance. They were positioned so that their centers were 1.43 deg above or below fixation. The adhesive label containing the word was pasted either over the center of the shape or at the same distance on the other side of fixation. Eighteen different cards were made for each of six conditions: word alone above fixation, word alone below fixation, word in frame above or below fixation, and word above or below fixation with frame on the opposite side. Thirty-six different words were used; their mean frequency was 38 per million. Each of these 36 words appeared once in each of the three conditions. Two cards were made with each of the 18 possible combinations of frame shape, frame color, and frame location.

Procedure. The displays were shown for 200 msec and were preceded and followed by a black fixation dot. The subjects were asked to read the word as quickly and as accurately as possible. Their vocal response latencies were measured to the nearest 5 msec using a voice-operated relay to stop a sweep clock that had been triggered by the onset of the stimulus display. The subjects were shown examples of the different types of cards and were given some practice trials before the experiment began. They were then run twice through the entire deck of 108 cards. The deck was shuffled before each use. The session lasted about 45 min.

Subjects. Ten subjects were paid to take part in the experiment. Eight of the subjects were women; all were University of British Columbia students. Their ages ranged from 18 to 29 years.

Results and Discussion

The means of response latencies, individual standard deviations, and error rates are shown in Table 1.

The difference of 42 msec between reading latencies in the word-alone and word-opposite conditions [$t(9) = 8.94$, $p < .001$] replicates the effect that has been labeled cognitive masking or filtering cost. A colored object that is physically very dissimilar to the word and spatially separated from it causes a substantial delay in reading. The novel observation of Table 1 is that the filtering cost is significantly *reduced* by placing the irrelevant object as a frame around the word [$t(9) = 5.93$, $p < .001$, for the comparison of mean reaction times (RTs) for the word-in-frame and word-opposite conditions]. The improvement of RT in the integrated displays is not produced by a speed-accuracy tradeoff, since errors and RTs show the same pattern, although not significantly in the case of errors.

The interpretation of filtering costs is substantially constrained by the fact that these costs are reduced when the word and the potentially distract-

Table 1
Reading Latency (in Milliseconds) and Errors in Experiment 1

	WA	WF	WOF
Mean RT	589	610	631
SD	30	31	38
Percent Errors	3.6	3.9	6.8

Note—WA = word alone; WF = word in frame; WOF = word opposite frame.

ing frame are grouped together. In particular, the costs cannot be due to sensory interaction, which should increase with spatial proximity. And the interference does not arise in the process of separating the relevant object from the irrelevant one, since this should also be more difficult when the word is in the frame. Rather, it appears that the interference occurs when two separate objects are present—either because there is competition in initially orienting to the correct object, or because of continued competition in their processing.

The filtering cost is reduced by half in the word-in-frame condition. The residual interference is ambiguous: it could be due simply to peripheral interactions, or to a milder form of the attentional competition observed when the word and the frame are separate. The lack of any effect of the frame on accuracy argues against the hypothesis of peripheral interference.

EXPERIMENT 2: STRATEGY EFFECTS IN FILTERING COSTS?

A question can be raised about Experiment 1: Did subjects attend first to the conspicuous colored frame to check whether it contained a word and then switch attention if it did not? Such a strategy would not have been applicable to other studies of filtering costs, in which the relevant item and the distracting object never overlapped. It could, however, explain the advantage of the word-in-frame over the word-opposite-frame condition, since the strategy of focusing on the frame would always be correct in the first case, always wrong in the second. Experiment 2 eliminated this possibility by the use of separate blocks, one testing the word alone and the word in the frame, and the other testing the word alone and the word separate from the frame. The stimuli were the same as those used in Experiment 1.

Method

The subjects were run in the two conditions—word in frame with word alone and word separate with word alone—in counter-balanced ABBA order, half starting with one condition and half with the other. They were given two blocks of 72 trials in each condition and were told the nature of the cards before beginning each condition. The first few trials were treated as practice. Ten paid subjects were run in this experiment, three men and seven women, all students at the University of British Columbia.

Results and Discussion

The results are shown in Table 2.

Separate analyses indicated no consistent effect of the order in which the subjects encountered the two conditions, and the results were pooled. The main question in this experiment was whether the large delay of reading in the word-opposite condition would be reduced when the subjects knew that the word could not appear in the frame. This did not happen.

Table 2
Reading Latency (in Milliseconds) and Errors in Experiment 2

	Condition 1		Condition 2	
	WA	WF	WA	WOF
Mean RT	563	594	565	615
SD	41	40	47	41
Percent Errors	1.8	2.1	0.7	5.0

Note—WA = word alone; WF = word in frame; WOF = word opposite frame.

The filtering cost of 52 msec was in fact slightly larger than in Experiment 1, and the difference between the integrated and separate conditions was replicated [21 msec; $t(9) = 2.29$, $p < .05$]. There is no indication that strategies contributed either to the filtering cost or to the mitigating effect of placing the word inside the frame.

EXPERIMENT 3: DUAL TASKS AND OBJECT INTEGRALITY

One possible reason for the filtering costs induced by the separate frame is that the frame attracted processing resources away from the word. A conventional analysis of attentional competition suggests that any manipulation that improves the allocation of attention to the word should reduce the attention available for the processing of the frame. Since the presentation of the word inside the frame reduced filtering cost, this hypothesis suggests that the frame might be perceived less well in the integrated than in the separate displays.

Our analysis, on the other hand, implies the opposite prediction: once attention is allocated to an object, it facilitates processing of all its parts or properties. Objects compete for control of attention, but different parts or properties of the attended object do not compete in the same manner. Consequently, both the frame and the word may benefit from focused attention if they are seen as belonging together. Experiment 3 was designed to test these alternative hypotheses. The subjects were given two tasks on each trial: (1) to read the word as quickly as possible and (2) to locate the position of a gap in the outline of the frame.

Method

Stimuli. The displays were similar to those in Experiments 1 and 2, and the words and luminances were the same. However, the frame was always rectangular, and one of its edges always passed through the fixation point. The remainder of the frame was above or below fixation, and the word was either in the center of the frame or at an equal distance from fixation but on the opposite side. The frame was 3.78 deg wide and 2.47 deg tall. On each trial on which the frame was presented, a small gap of 0.39 deg was present in the black outline of the frame, 0.86 deg to the left or to the right of fixation on the edge that was centered on the fixation point. The two possible positions of the gap (left or right of

center) were identical relative to fixation, whether the rest of the frame was above or below fixation and whether or not it contained the word. The same three colors, pink, green, and yellow, were used to fill the frame as before. Each display was preceded and followed by a card containing a strip of black and white random checkerboard 4.0 deg wide and 0.96 deg tall, centered at fixation, to mask the gap without impairing the normal persistence of the word.

Procedure. The subjects were asked to read the word as quickly as possible; this was the primary task. They were also asked to report the position of the gap in the frame (left or right of center). The exposure duration was continuously varied for each subject so as to maintain accuracy on the gap discrimination task at about 75%. Because the word was not pattern masked, its visibility was not critically affected by these variations of exposure time. The median durations of exposure for individual subjects ranged from 105 to 200 msec, with a mean of 158 msec.

The subjects were given a few practice trials on reading the word alone, then on reading the word with the frame present, and finally with the dual task of reading the word and locating the gap. They were then run on two blocks of 108 trials, 36 in each of three conditions (word alone, word in frame, and word opposite to frame). The conditions were randomly mixed within each block.

Subjects. The 12 paid subjects were students at the University of British Columbia, 9 women and 3 men, aged 18 to 29 years.

Results and Discussion

The mean reading times, error rates in reading, and error rates in gap location are shown in Table 3.

Although the subjects had been instructed to treat reading as the primary task, they clearly did not assign absolute priority to that task. Reading times were generally slower than in Experiments 1 and 2, and the difference between the word-alone and word-opposite conditions was 108 msec, compared with 42 and 52 msec in the two previous studies.

Of main interest in the present experiment was the comparison of reading and of gap detection when the word was in the frame and when it was separate from the frame. The performance of both tasks was worse when word and frame were separate. The difference of 17 msec in reading delay was significant [$t(11) = 2.86, p < .01$], and so was the difference in error rates in locating the gap [$t(11) = 4.40, p < .01$]. The advantage of the integrated condition in reading latency replicates the previous results. The advantage of the same condition in the gap-location task shows that dividing attention between parts or properties of the same object is easier than dividing attention between two separate objects. There is no evidence

for a tradeoff in processing resources such that perception of the frame benefits when it delays the reading of the word.

Hoffman and Nelson (1981) and Hoffman, Nelson, and Houck (in press) have recently reported a related result, showing that when attention is directed to a particular spatial location it facilitates the processing of any stimuli presented in or near that location. A probe stimulus (whether a brief offset of one of four lights or a U shape in one of four different orientations) was better detected or discriminated when it was near a target in a concurrent search task than when it was located farther from the target, next to a distractor. The notion of attention as a spotlight that selects an area of space has been discussed by several authors (Eriksen & Hoffman, 1973; Posner, 1980; Shulman, Remington, & McLean, 1979). In our experiment, however, the spatial distance of the gap from the target word was constant; differences in discriminability were induced by perceptual grouping or segregation of frame and word. The results suggest that the objects of attention are defined by perceptual grouping, not just by spatial distance. The gradient of interference is defined not in retinal distance but in "belongingness." Banks and Prinzmetal (1976) discovered a related effect of perceptual grouping in a search task. Their experiments showed that search was hampered by grouping the target with the distractors and facilitated by separating them. This is the opposite of the effects that we find in the filtering task. To reconcile these apparently conflicting results, we suggest that spatial segregation may help when the problem is to detect a target in a display containing distractors that are otherwise confusable. It can also help, as mentioned earlier, when the relevant and irrelevant objects induce competing responses, as in Stroop experiments (Kahneman & Henik, 1981). However, in the experiments described here, the frame was neither confusable with the word nor likely to evoke a conflicting response. The possible advantages for separation were therefore absent.

We account for the disadvantage of separation as follows: When the task requires subjects to process the relevant object in order to determine a choice response, attention must be directed to the object and a filtering cost is incurred. In this case, segregation of relevant from irrelevant objects can be damaging because it sets up a conflict for attention between potential objects. Any object on which attention could, but should not, be focused must be excluded at a cost, in time or in accuracy. Segregation is also detrimental when attention must be divided between two objects, as in Experiment 3. Information about the different properties of one object can be registered with less difficulty than the same information originating from two different objects. Attention seems to be limited in the number of separate objects that it

Table 3
Reading Latency (in Milliseconds) and Errors in Reading and in Gap Localization in Experiment 3

	WA	WOF	WF
Mean Reading Time	623	731	714
SD	66	112	108
Percent Errors, Reading	0	2.5	2.1
Percent Errors, Gap Localization		27.3	16.0

Note—WA = word alone; WOF = word opposite frame; WF = word in frame.

can encompass, but not (or much less so) in the number of properties that can be concurrently registered when they belong to a single object (Lappin, 1967; Treisman, 1969).

GENERAL DISCUSSION

Each of the three experiments showed a decrement in performance when two objects were present rather than one, whether both were relevant or only one was relevant. In each experiment, the decrement was smaller when the two objects could be perceptually grouped into one complex entity. Certainly, none of the results fits well with an account in terms of peripheral interference, since spatial proximity reduces, rather than increases, the interference.

The results are also difficult to reconcile with the treatment of attention in many current theories. The most generally accepted view of perception equates it with the temporary activation of some set of internal nodes representing the properties and the identity of the objects currently in view (Johnston & Dark, 1982; LaBerge, 1975; Shiffrin, 1976; Shiffrin & Schneider, 1977). These nodes form part of an interconnected network that also represents learned associations and relations between familiar objects. Nodes can be activated by excitation from other nodes as well as by sensory input. We have called this the "display board" view of the mind (Kahneman & Treisman, 1983). It is well suited to explaining the effects of expectancy and set; selected nodes can be primed or preactivated and thus require fewer sensory data before reaching their threshold. However, display board models seem less effective at representing the operation of selective attention in filtering tasks (Broadbent, 1958).

We note two observations that are hard to accommodate within the display-board model: (1) the fact that efficient selection in standard filtering tasks (such as selective reading or dichotic listening) can be controlled by one property of the attended input (e.g., its color, location, or ear of arrival) and the response by another (e.g., the identity of the words); (2) the present finding that filtering costs are reduced and concurrent processing facilitated by perceptual grouping or integration of separate objects. Thus, it is easy to shadow the word "house" when it is said to the right ear and to ignore it when it is said to the left, and it is easier to read the word "cow" when it is surrounded by a yellow frame than when a yellow frame appears on the other side of fixation. Both findings illustrate the importance of the spatial conjunctions of properties at a particular moment of time—the current episode. The objects of selection in these tasks could not economically be represented by conjunction nodes in semantic memory. No one could plausibly require that a complete mental vocabulary be replicated for every possible spatial

source or speaker's voice or that a node exist specifically for the word "cow" in a rectangular yellow frame. Attention in filtering tasks cannot work only by selectively priming nodes.

We suggest that an account of filtering must invoke a separate perceptual representation that retains the integrity of spatially distinct objects, a temporary episodic structure that is distinct from the permanent structures of semantic memory that mediate perceptual identification (Kahneman & Henik, 1981; Treisman & Schmidt, 1982). We use the metaphor of an "object file" for the temporary representation in which the information that pertains to a particular object accumulates and is updated when the object changes. The identity of the object and of its properties may be retrieved by comparing the contents of the object file with nodes in perceptual and semantic memory. However, perception does not depend on a match to a known object's being found; we do not remain "blind" to totally unfamiliar sights, even if we are unable to make sense of them.

A selective mechanism that permits relevant object files to be favored in processing is added, in this account, to the priming mechanism that provides the sole explanation of attention in display-board models. Attention is directed, not to a node in memory, but to a particular object or event. The selected object file is located by the critical property that defines the object as relevant. Information accruing in the selected object file is privileged in subsequent processing. Selection is necessary because constraints on parallel processing of different object files are substantially more severe than constraints on the processing of different aspects of the same objects. Filtering costs arise when two objects compete for attention and when further processing of the relevant object is necessary beyond the detection of its defining property. These assumptions provide an account both of the cause of filtering costs and of the advantages of integrated displays in the present experiment.

REFERENCES

- BANKS, W. P., & PRINZMETAL, W. Configurational effects in visual information processing. *Perception & Psychophysics*, 1976, 19, 361-367.
- BJORK, E. L., & MURRAY, J. T. On the nature of input channels in visual processing. *Psychological Review*, 1977, 84, 472-484.
- BROADBENT, D. E. *Perception and communication*. London: Pergamon Press, 1958.
- ERIKSEN, C. W., & HOFFMAN, J. E. Temporal and spatial characteristics of selective encoding from visual displays. *Perception & Psychophysics*, 1972, 12, 201-204.
- ERIKSEN, C. W., & HOFFMAN, J. E. The extent of processing of noise elements during selective encoding from visual displays. *Perception & Psychophysics*, 1973, 14, 155-160.
- ERIKSEN, C. W., & SCHULTZ, D. W. Temporal factors in visual information processing: A tutorial review. In J. Requin (Ed.), *Attention and performance VII*. Hillsdale, N.J.: Erlbaum, 1978.

- ESTES, W. K. Interaction of signal and background variables in visual processing. *Perception & Psychophysics*, 1972, **12**, 278-286.
- FRANCOLINI, C. M., & EGETH, H. A. On the nonautomaticity of "automatic" activation: Evidence of selective seeing. *Perception & Psychophysics*, 1980, **27**, 331-342.
- HOFFMAN, J. E., & NELSON, B. Spatial selectivity in visual search. *Perception & Psychophysics*, 1981, **30**, 283-290.
- HOFFMAN, J. E., NELSON, B., & HOUCK, M. The role of attentional resources in automatic detection. *Cognitive Psychology*, in press.
- JOHNSTON, W. A., & DARK, V. J. In defense of intraperceptual theories of attention. *Journal of Experimental Psychology: Human Perception and Performance*, 1982, **8**, 407-421.
- KAHNEMAN, D. *Attention and effort*. Englewood Cliffs, N.J.: Prentice-Hall, 1973.
- KAHNEMAN, D., & HENIK, A. Perceptual organization and attention. In M. Kubovy & J. R. Pomerantz (Eds.), *Perceptual organization*. Hillsdale, N.J.: Erlbaum, 1981.
- KAHNEMAN, D., & TREISMAN, A. M. Changing views of attention and automaticity. In R. Parasuraman, R. Davies, & J. Beatty (Eds.), *Varieties of attention*. New York: Academic Press, 1983.
- KAHNEMAN, D., TREISMAN, A. M., & BURKELL, J. The cost of visual filtering. *Journal of Experimental Psychology: Human Perception and Performance*, in press.
- LABERGE, D. Acquisition of automatic processing in perceptual and associative learning. In P. M. A. Rabbitt & S. Dornic (Eds.), *Attention and performance V*. New York: Academic Press, 1975.
- LAPPIN, J. S. Attention in the identification of stimuli in complex visual displays. *Journal of Experimental Psychology*, 1967, **75**, 321-328.
- POSNER, M. I. The orienting of attention. *Quarterly Journal of Experimental Psychology*, 1980, **32**, 3-25.
- SHIFFRIN, R. M. Capacity limitations in information processing, attention, and memory. In W. K. Estes (Ed.), *Handbook of learning and cognitive processing* (Vol. 4). Hillsdale, N.J.: Erlbaum, 1976.
- SHIFFRIN, R. M., & SCHNEIDER, W. Controlled and automatic human information processing: II. Perceptual learning, automatic attending and a general theory. *Psychological Review*, 1977, **84**, 127-190.
- SHULMAN, G. L., REMINGTON, R. W., & McLEAN, J. P. Moving attention through visual space. *Journal of Experimental Psychology*, 1979, **5**, 522-526.
- TREISMAN, A. M. Strategies and models of selective attention. *Psychological Review*, 1969, **76**, 282-299.
- TREISMAN, A. M., & SCHMIDT, H. Illusory conjunctions in the perception of objects. *Cognitive Psychology*, 1982, **14**, 107-141.

(Manuscript received September 29, 1982;
revision accepted for publication February 9, 1983.)