

Familiarity and pop-out in visual search

QINQIN WANG and PATRICK CAVANAGH
Harvard University, Cambridge, Massachusetts

and

MARC GREEN
Trent University, Peterborough, Ontario, Canada

In this paper, we report that when the low-level features of targets and distractors are held constant, visual search performance can be strongly influenced by familiarity. In the first condition, a **N** was the target amid **ns** as distractors, and vice versa. The response time increased steeply as a function of number of distractors (82 msec/item). When the same stimuli were rotated by 90° (the second condition), however, they became familiar patterns—**Z** and **S**—and gave rise to much shallower search functions (31 msec/item). In the third condition, when the search was for a familiar target, **N** (or **Z**), among unfamiliar distractors, **ns** (or **Ss**), the slope was about 46 msec/item. In the last condition, when the search was for an unfamiliar target, **N** (or **S**), among familiar distractors, **Ns** (or **Zs**), parallel search functions were found with a slope of about 1.5 msec/item. These results show that familiarity speeds visual search and that it does so principally when the distractors, not the targets, are familiar.

We have all learned from experience that some things are easier to find than others. For example, a person wearing a red shirt immediately stands out in a crowd of people wearing white shirts. In fact, this red shirt is detected as easily in the crowd of white shirts (distractors) as it is when presented alone. However, if you need to pick out one red shirt from several others that differ in some less noticeable way, say style, you will have to look at the shirts one by one. These differences in visual processing were documented as early as the 11th century by Ibn Al-Haytham (English translation by Sabra, 1989): "some of the particular properties of which the forms of visible objects are composed appear at the moment when sight glances at the object, while others appear only after scrutiny and contemplation." More recently, the rapid category of visual search has been referred to as parallel search or "pop-out" (Treisman, 1985; Treisman & Gormican, 1988; Treisman & Souther, 1985), because the target seems to leap from the display without any effort on the part of the observer. The slower category of search has been referred to as serial, because each item has to be processed sequentially.

These qualitatively different kinds of search are used as evidence that vision involves the interplay of two distinct systems. According to Treisman (1988), Julesz (1984), and others, the first, the preattentive system, operates in parallel over a large visual field and registers the presence of all "primitive" features. The second, the at-

tentive system, operates by serially moving a small high-resolution "spotlight" to each item in turn. On the basis of these hypotheses, Treisman (1988) used visual search as a criterion to identify primitive features (Treisman & Gormican, 1988). The distinction between serial and parallel processing was based on search rate—the increment in reaction time for each additional distractor. Low search rates indicated parallel processing and suggested that target and distractors differed by a primitive feature, such as color, length, or curvature. Note that "low" is an arbitrary term; Treisman set the low criterion at 6 msec/item or less, while other authors (e.g., Enns, 1990) might regard 10 msec/item as low. In the present experiment, the criterion for discriminating serial from parallel was preset to 8 msec/item.

Recent studies, however, have also shown parallel search for simple feature conjunctions (Nakayama & Silverman, 1986; Wolfe, Cave, & Franzel, 1989) as well as for high-level features, such as depth from shading (Ramachandran, 1988; Kleffner & Ramachandran, 1992), three-dimensional structure (Enns & Rensink, 1990), and surface representation (He & Nakayama, 1992). These results suggest that "pop-out" can also be supported by higher level processing.

In addition to the physical properties of the stimulus, nonvisual, category factors can also mediate visual search. Jonides and Gleitman (1972) asked subjects to search for an "O" among several letter distractors or among several digit distractors. Before the search task, one group of subjects were instructed that the "O" was the vowel letter O, while the other group of subjects were instructed that the "O" was the digit 0. Jonides and Gleitman reported strong category effects in that, for example, the search rate for the "O" among letter distrac-

This research was supported by NEI Grant EY09258. The authors are grateful to Reynald Comtois for his technical assistance. Correspondence concerning this paper may be addressed to P. Cavanagh, Department of Psychology, Harvard University, Cambridge, MA 02138 (e-mail:patrick@isr.harvard.edu).

SIN 2 and ZINC, both
EXPT 1 ≠ EXPT 2 → distractors familiar helps

overall:
best if
fam. distractor

tors was faster when the shape was thought to be a digit than when it was thought to be a letter. This categorical effect has also been demonstrated by Brand (1971) and Egeth, Jonides, and Wall (1972). They examined two visual-search conditions in which subjects looked for any digit or one specific digit in a background of letters. The results indicated that these two searches were accomplished at about the same rate, and that both were faster than a condition in which subjects looked for a specific letter in a background of other letters.

In the experiment reported here, we extended this evaluation of automatic processing for nonvisual factors to familiarity. Familiarity is a property that is not a characteristic of the visual pattern itself but is dependent on the subject's experience with that pattern. Our experiment examined whether familiarity could also lead to pop-out in visual search even when primitive features were held constant. Several previous studies that had addressed the influence of familiarity on visual search (Frith, 1974; Reicher, Snyder & Richards, 1976; Richards & Reicher, 1978) had manipulated familiarity by inverting letters of the alphabet. The search task was to find, for example, a single upright letter (familiar, such as **N**) in a field of inverted (unfamiliar, such as **И**) distractors. In general, familiarity speeded the visual search when it was confined to the distractors. That is, search rate was lowest when an inverted letter was the target and upright letters were the distractors. Even in this case, however, search rates remained substantial. In Frith's (1974) experiment, subjects were asked to cross out every target in a line of distractors. The number of distractors was not varied, so this procedure could not evaluate the search rate directly. In Reicher et al.'s (1976) and Richards and Reicher's (1978) experiments, a visual-search paradigm was used, but with heterogeneous distractors. Since distractor heterogeneity per se can reduce search efficiency (Gordon, 1968; Duncan & Humphreys, 1989), it is not surprising that search was serial in this task.

Schneider and Shiffrin (1977) and Shiffrin and Schneider (1977) had claimed that, with letter search, extended practice with a fixed set of letters could lead to parallel search. However, since the stimuli's low-level features were not controlled, this learning effect might have been mediated by the isolation of distinctive features in the stimuli. For example, a particular target letter may have a single feature that distinguishes it from the distractor letters, such as the letters J and H; J has curvature, while H doesn't. Once this feature is discovered, it can support preattentive search.

In our experiment, we restricted the variation of low-level features between target and distractors and varied only familiarity. Unlike most of the experiments mentioned above, the distractors were homogeneous (e.g., all Ns) rather than heterogeneous (e.g., N, M, T, etc.), and there was no extended practice.

The stimuli used as target and distractors in our experiment were mirror images of each other, for example, **N** vs. **И** and **2** vs. **5**. This strongly constrained the differences in simple features between target and distrac-

tors. In the first condition, the shapes were **И** and **И** which share all orientation, terminator, closure, and symmetry features that might mediate a low-level preattentive search. In the second condition, the same stimuli were rotated by 90° and now enjoyed an advantage of familiarity—**2** and **5**—while still lacking low-level feature differences.

In the first two conditions, both target and distractors were of the same type—either both unfamiliar or both familiar. In the last two conditions, the targets and distractors differed in familiarity. In the third condition, an **N** (or **Z**) was the target among mirror image **И**s (or **Σ**s); in the fourth condition, the roles were reversed. Note that there is also a low-level feature (the orientation of the oblique) that distinguishes target from distractors in the last two conditions.

METHOD

Subjects

Six subjects volunteered for the experiment and were paid \$8 an hour. They were 18–24 years old, and their native language was English, French, or German. All subjects had normal or corrected-to-normal vision.

Stimuli

Four different target-distractor pairs were used in the four different conditions. In Condition 1, both target and distractors were unfamiliar; in Condition 2, both target and distractors were familiar; in Condition 3, the target was familiar and the distractors were unfamiliar; in Condition 4, the target was unfamiliar and the distractors were familiar. Within each condition, the stimuli were divided into two blocks. Figure 1 shows the target and distractors for each of the resulting eight blocks. In each condition, the target and distractors appeared randomly in six possible positions on a 4.7°-diam circle, as shown in Figure 2. On each trial, the target was either present or absent, and the number of stimulus items varied from 1 to 6. Targets and distractors were 0.9° square in size (**N**, **И**, **Z**, and **Σ**), or 0.6° × 0.9° (**2** and **5**), or 0.9° × 0.6° (**И** and **И**). The computer-generated black stimuli appeared on a white background with 64.0 cd/m² mean luminance.

CONDITION		TARGET	DISTRACTORS
1	Block 1	И UNFAM	И UNFAM
	Block 2	И	И
2	Block 3	2 FAM.	5 Familiar
	Block 4	5	2
3	Block 5	N Fam	И UNFAM
	Block 6	Z	Σ
4	Block 7	И UNFAM.	N FAM.
	Block 8	Σ	Z

Figure 1. Targets and distractors for the four conditions; each condition is divided into two blocks.

STIMULI ?

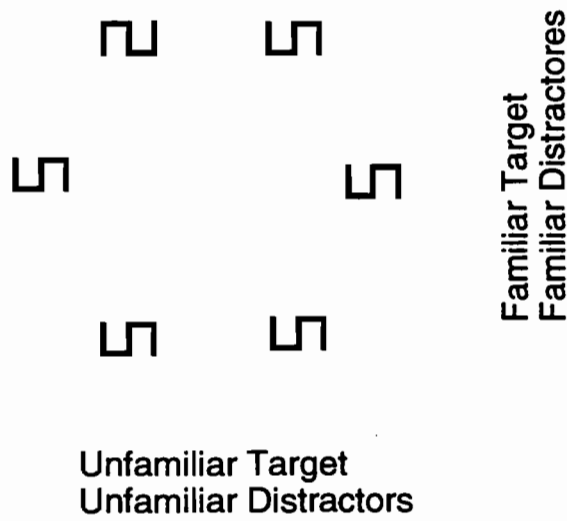


Figure 2. Example of stimuli with one target and five distractors in Condition 1. Rotating the page 90° demonstrates the stimuli in Condition 2.

Procedure

In each condition, we tested one pair of stimuli. Each of the four conditions was divided into two blocks. There were 12 cases in each block (target absent vs. target present × 6 set sizes), and each case was repeated 8 times; so there were 96 trials in each block. The order of the four conditions was randomized for each subject.

Before each block, the subject was told what the target would be and was shown examples of displays with and without the target. At the beginning of each block, a fixation point (two small black concentric circles) appeared in the center of the screen, then a sharp beep was sounded to alert the subject that the trial was about to start. The stimuli appeared 500 msec after the beep and remained visible until the subject responded. If the subject did not respond within 2,000 msec, the display would disappear and this trial was counted as an error. The intertrial interval was 1,000 msec. Through the whole process, the fixation point stayed on. No feedback was given.

The subjects were instructed to press one of two keys to indicate whether the target was present or absent “as quickly as possible while minimizing errors.” Before the experimental sessions, the subjects were given one or two sets of 24 practice trials until error rates dropped below 5%. Typically, one practice session was sufficient.

RESULTS

Error rates were low throughout, averaging from the highest value, 5% in Condition 3, to the lowest, 1% in Condition 2. Linear regressions were computed for the mean reaction times of correct responses as a function of the number of the items in the display. Slopes are reported separately for target present and target absent, and the median standard error of the mean reaction times for each panel is shown as a vertical bar ($\pm 1 SE$).

In Condition 1, when both target and distractors were unfamiliar, the search was the slowest. The average slope for the target-present responses was 82 msec/item (Figure 3).

In Condition 2, when both target and distractors were familiar, the search was much faster. The average slope for the target-present responses was 31 msec/item. These slopes were still very high, however, and were not in the range of preattentive processes (Figure 4).

In Condition 3, when target was familiar and distractors were unfamiliar, the average slope for the target-present situation was 51 msec/item. The introduction of familiarity in both Conditions 2 and 3 facilitated the search as compared with Condition 1, in which both target and distractors were unfamiliar. However, the search was still not parallel (Figure 5).

In Condition 4, in which the target was unfamiliar and the distractors were familiar, the average slope for the target-present responses dropped to 1.5 msec/item, which was not significantly different from zero. These stimuli appear to be analyzed by a parallel process. The unfa-

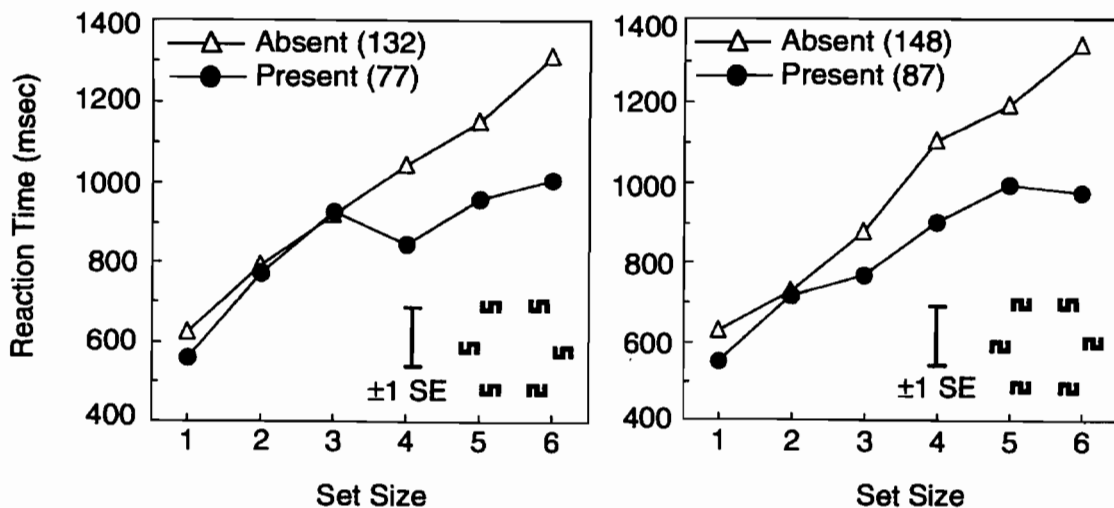


Figure 3. Reaction time as a function of set size when searching for *ru* in *uns* and *ur* in *rus*. The slopes for correct target-present and -absent regressions are shown in parentheses. Note that the vertical axis scale here is different from that used in Figures 4, 5, and 6.

Worst
by far

is better both FAM.

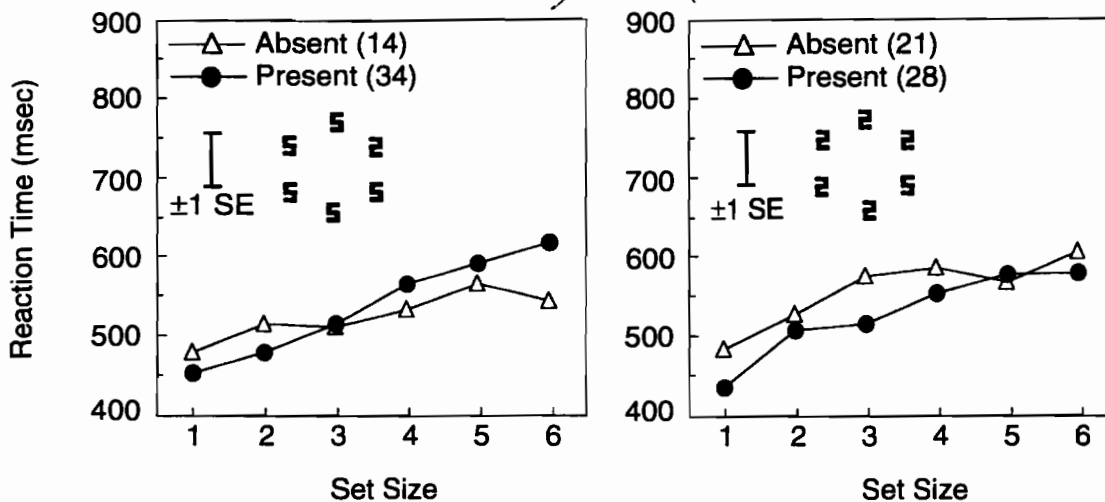


Figure 4. Reaction time as a function of set size when searching for 2 in 5s and 5 in 2s. The slopes for correct target-present and -absent regressions are shown in parentheses.

familiar target popped out from the familiar background (Figure 6).

DISCUSSION

Although familiarity is a property of memory and not of stimulus shape per se, our results suggest that familiarity can support parallel search, as long as the familiarity is confined to the background. When both target and distractor were unfamiliar (Condition 1) or both were familiar (Condition 2), or when the target was familiar but the distractors were unfamiliar (Condition 3), search was very slow. Only when the target was unfamiliar and the distractors were familiar (Condition 4) was search a parallel process. Recall that in Conditions 3 and 4 there was an alternative explanation for rapid

search, which was that the target differed from the distractors by a single feature (a left-tilted vs. right-tilted oblique) as well as by familiarity. However, since this simple feature difference was common to both conditions, it cannot have mediated the difference in search performance between them.

A number of studies have demonstrated the importance of familiarity of background items. As discussed in the introduction, Frith (1974), Reicher et al. (1976), and Richards and Reicher (1978) used different visual search paradigms. Reicher et al. (1976) and Richards and Richards (1978) used a heterogeneous distractor field composed of many different letters, whereas our distractor field was always composed of repetitions of a single pattern. Thus, although they found that a familiar background (upright letters) reduced search rates as com-

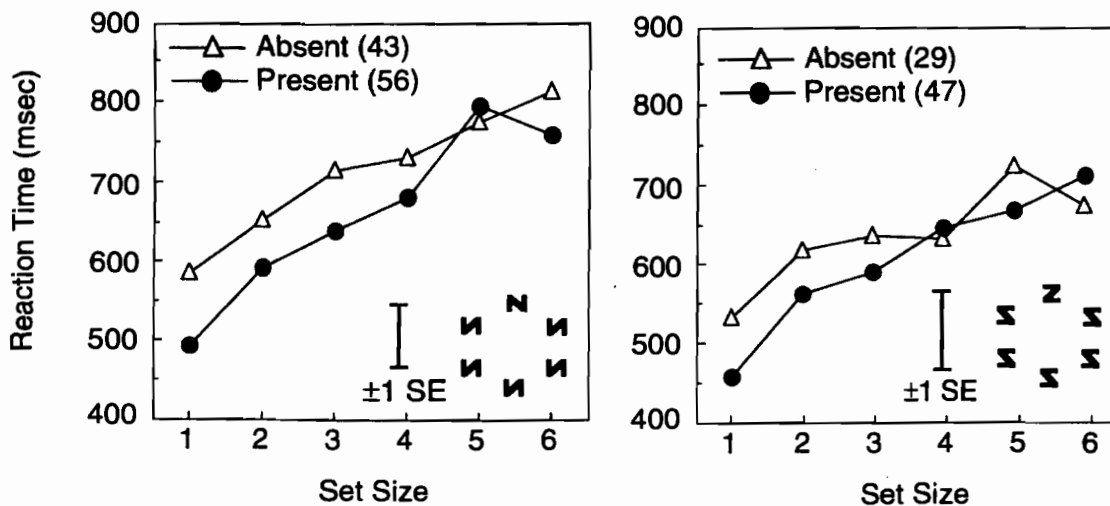


Figure 5. Reaction time as a function of set size when searching for N in Ns and Z in Zs. The slopes for correct target-present and -absent regressions are shown in parentheses.

stimulated only get for familiar distractors

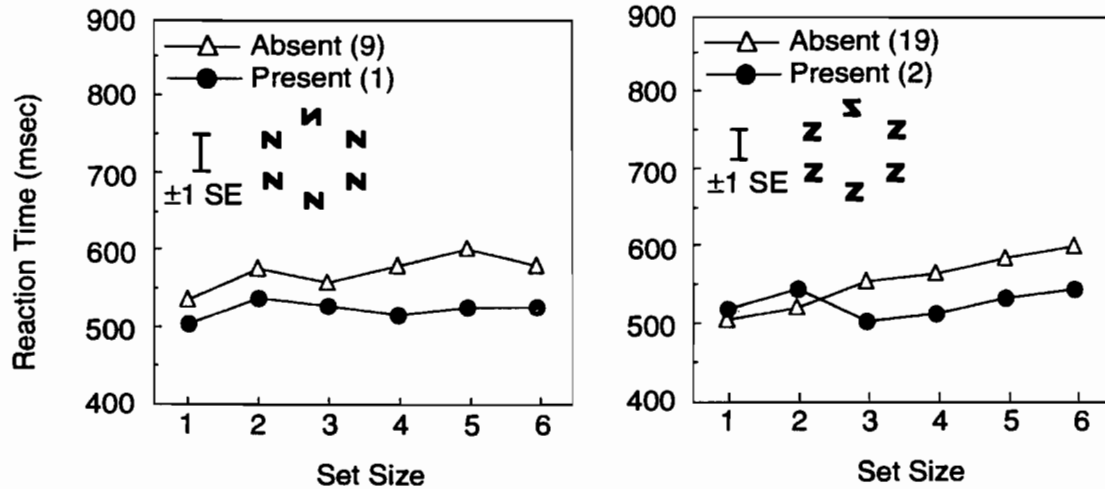


Figure 6. Reaction time as a function of set size when searching for M in Ns and Σ in Zs. The slopes for correct target-present and -absent regressions are shown in parentheses.

pared with an unfamiliar background (upside-down letters), the search remained serial (about 25 msec/item, well above our previously described criterion of 8–10 msec/item for parallel search). Frith (1974) tested stimuli identical to our N and M of Conditions 3 and 4, but she required her subjects to manually cross out the multiple targets among a number of distractors, a procedure which could not reveal whether search would be serial or parallel.

Duncan and Humphreys (1989) proposed a theory for visual search based on interitem similarity which could explain the difference between our results and those of Reicher et al. (1976) and Richards and Reicher (1978). Duncan and Humphreys (1989) claim that search efficiency increases with decreasing target–distractor similarity and with increasing distractor–distractor similarity. Our displays clearly had greater distractor–distractor similarity than did those of Reicher et al. (1976) and Richards and Reicher (1978). On the other hand, Duncan and Humphreys’s (1989) hypothesis gives a poor account of the search asymmetry we found in our experiment. Specifically, in Conditions 3 and 4, we simply exchanged the target and the distractors. The stimuli therefore had identical target–distractor similarity and distractor–distractor similarity in both conditions, yet search was much faster when the familiar patterns were the distractors. This result shows that in addition to similarity, familiarity must also mediate the efficiency of search processes.

How does familiarity affect the search process? One possibility is that familiarity facilitates the grouping of background items, which then facilitates the “pop-out” of the target. Facilitation due to the familiarity of the background was reported by Karni and Sagi (1991). In their experiments, a texture-discrimination task was used and observers “learned” the background and target patterns over extended periods of practice. During the practice, performance improved significantly. The learning was then tested with a transfer task which used, in

one comparison, either the same target pattern with a new background or the same background with a new target. The performance improvement transferred to the new task only if the background was kept the same as that used during learning. Although the familiarity of the background was the critical variable in Karni and Sagi’s (1991) experiment, it cannot explain our results. It correctly predicts faster search in Condition 4 (familiar distractors) than in Condition 3 (unfamiliar distractors), but it would also predict that search in Condition 2 (familiar target and background) should be at least as fast as that in Condition 4 (unfamiliar target, familiar distractors). This was not the case. There may be more efficient processing or “grouping” of the familiar backgrounds in Conditions 2 and 4, but this is not sufficient for parallel search when both target and background are familiar.

Clearly, a difference of familiarity between the target and the distractors was necessary for the search to be parallel if there were no other primitive feature differences. This suggests that familiarity itself might be considered a primitive feature which can be processed pre-attentively. There is a simple, almost circular, reason why this might be the case. Familiar items can be processed or identified rapidly, but unfamiliar items need additional processing resources and therefore ought to attract attention to themselves in order to begin the special analysis. This is analogous to the classical orienting response, in which an organism directs attention to a novel or unfamiliar item. In Conditions 1 and 2, since the target and distractors are equally familiar or unfamiliar to the subjects and do not have any other primitive feature differences, all items would draw attention equally. Subjects have to direct their attention to each item in turn in order to distinguish the target from the distractors. In Condition 3, the target is familiar and the distractors are unfamiliar, so the unfamiliar distractors might be processed first, before the target is found. In Condition 4, when only the target is unfamiliar and distractors

are familiar, the unfamiliar target will draw attention and be identified rapidly no matter how many distractor items are present.

If familiarity (or its inverse, novelty) did act as a primitive feature, then the asymmetrical familiarity effect in Conditions 3 and 4 would also be consistent with Treisman's (1985) explanation of asymmetrical search rates. In this view, familiar items would be coded as standards and unfamiliar (novel) items would be coded as deviations from the standards. According to Treisman, standard items elicit less activity, whereas deviations from the standards elicit more activity. Treisman and her colleagues have shown in many instances that searching for a single deviation among a field of standards, in other words, for a high-activity item among low-activity distractors (e.g., our Condition 4), is rapid. Conversely, searching for a single standard among a field of deviants, that is, for a low-activity target among high-activity distractors (e.g., our Condition 3), is much slower.

Indeed, recent physiological evidence (Miller, Li, & Desimone, 1991) supports the hypothesis of an enhanced neural response to novel visual stimuli. Desimone's group recorded single neurons in the inferotemporal cortex of two awake monkeys. They found that some of the neurons that respond selectively to particular object features, such as color or shape, gave their best response to those features when they were part of new or unexpected patterns. This enhancement for novelty declined as the stimuli became familiar to the animal.

We have found that a familiarity difference between target and background can support parallel search. This result implies that the familiarity of each item must be available very rapidly, and that the test stimuli must make contact with memory before the parallel search process terminates. Treisman's original work (Treisman, 1985; Treisman & Gormican, 1988; Treisman & Souther, 1985) suggested that parallel search might be fast because it was based on properties of the stimuli that emerged very early in visual processing, such as the features coded by early cortical receptive fields. More recent work (Enns & Rensink, 1990; He & Nakayama, 1992; Kleffner & Ramachandran, 1992) has shown that higher levels of analysis of three-dimensional surfaces are also completed in time to support rapid, parallel search. Our results indicate that memory contact can also occur at these time scales, indicating that memory access may be an early step in visual processing and not, as might have been imagined, the last step.

REFERENCES

- BRAND, J. (1971). Classification without identification in visual search. *Quarterly Journal of Experimental Psychology*, *23*, 178-186.
- DUNCAN, J., & HUMPHREYS, G. W. (1989). Visual search and stimulus similarity. *Psychological Review*, *96*, 433-458.
- EGETH, H., JONIDES, J., & WALL, S. (1972). Parallel processing of multi-element displays. *Cognitive Psychology*, *3*, 674-698.
- ENNS, J. T. (1990). Three-dimensional features that pop-out in visual search. In D. Brogan (Ed.), *Visual search: Proceedings of the First International Conference on Visual Search* (pp. 37-45). New York: Taylor & Francis.
- ENNS, J. T., & RENSINK, R. A. (1990). Sensitivity to three-dimensional orientation in visual search. *Psychological Science*, *1*, 323-326.
- FRITH, U. (1974). A curious effect with reversed letters explained by a theory of schema. *Perception & Psychophysics*, *16*, 113-116.
- GORDON, I. E. (1968). Interaction between items in visual search. *Journal of Experimental Psychology*, *76*, 348-355.
- HE, Z. J., & NAKAYAMA, K. (1992). Surfaces versus features in visual search. *Nature*, *359*, 231-233.
- JONIDES, J., & GLEITMAN, H. (1972). A conceptual category effect in visual search: O as letter or as digit. *Perception & Psychophysics*, *12*, 457-460.
- JULESZ, B. (1984). Toward an axiomatic theory of preattentive vision. In G. M. Edelman, W. E. Gall, & W. M. Cowan (Eds.), *Dynamic aspects of neocortical function* (pp. 585-711). New York: Wiley.
- KARNI, A., & SAGI, D. (1991). Where practice makes perfect in texture discrimination: Evidence for primary visual cortex plasticity. *Proceedings of the National Academy of Sciences*, *88*, 4966-4970.
- KLEFFNER, D. A., & RAMACHANDRAN, V. S. (1992). On the perception of shape from shading. *Perception & Psychophysics*, *52*, 18-36.
- MILLER, E. K., LI, L., & DESIMONE, R. (1991). A neural mechanism for working and recognition memory in inferior temporal cortex. *Science*, *254*, 1377-1379.
- NAKAYAMA, K., & SILVERMAN, G. H. (1986). Serial and parallel processing of visual feature conjunctions. *Nature*, *320*, 264-265.
- RAMACHANDRAN, V. S. (1988). Perceiving shape from shading. *Scientific American*, *259*, 76-83.
- REICHER, G. M., SNYDER, C. R. R., & RICHARDS, J. T. (1976). Familiarity of background characters in visual scanning. *Journal of Experimental Psychology: Human Perception and Performance*, *2*, 522-530.
- RICHARDS, J. T., & REICHER, G. M. (1978). The effect of background familiarity in visual search: An analysis of underlying factors. *Perception & Psychophysics*, *23*, 499-505.
- SABRA, A. I. (1989). *The optics of Ibn Al-Haytham* (English trans.). London: Warburg Institute, University of London.
- SCHNEIDER, W., & SHIFFRIN, R. M. (1977). Controlled and automatic human information processing: I. Detection, search and attention. *Psychological Review*, *84*, 1-58.
- SHIFFRIN, R. M., & SCHNEIDER, W. (1977). Controlled and automatic human information processing: II. Perceptual learning, automatic attending, and a general theory. *Psychological Review*, *84*, 127-188.
- TREISMAN, A. (1985). Preattentive processing in vision. *Computer Vision, Graphics, & Image Processing*, *31*, 156-177.
- TREISMAN, A. (1988). Features and objects: The 14th Bartlett Memorial Lecture. *Quarterly Journal of Experimental Psychology*, *40A*, 201-237.
- TREISMAN, A., & GORMICAN, S. (1988). Feature analysis in early vision: Evidence from search asymmetries. *Psychological Review*, *95*, 15-48.
- TREISMAN, A., & SOUTHER, J. (1985). Search asymmetry: A diagnostic for preattentive processing of separable features. *Journal of Experimental Psychology: General*, *114*, 285-310.
- WOLFE, J. M., CAVE, K. R., & FRANZEL, S. L. (1989). Guided search: An alternative to the feature integration model for visual search. *Journal of Experimental Psychology: Human Perception & Performance*, *15*, 419-433.