Visual search through color displays: Effects of target-background similarity and background uniformity

ERIC W. FARMER and ROBERT M. TAYLOR

Royal Air Force Institute of Aviation Medicine, Farnborough, Hants., England

The effects of target-background similarity and background uniformity were investigated in two experiments in which subjects reported the presence or absence of an achromatic target item in arrays containing irrelevant chromatic items. In Experiment 1, both "present" and "absent" decision times were an increasing function of the similarity of a predefined target item to the background, and were shortened when the interitem similarity of the background set was increased. In Experiment 2, the effects of target-background similarity on the "present" decision were replicated when subjects were not informed on each trial as to which member of the target set might appear in the array, and the "absent" decision was faster when each row of the array comprised background items of a single color than when these items were allocated randomly to array locations. The results are interpreted in terms of the role of pre-attentive grouping operations.

Several techniques have been developed to identify factors which influence the detectability of a target element in a multielement array. For example, subjects may be given the task of locating a predefined target in a large display which requires several changes of eye fixation (molar visual search), or of indicating which of a set of possible targets appears within a small array which is presented sufficiently briefly to permit only a single fixation (forced-choice detection). Despite such procedural differences, these techniques have provided converging evidence of the critical role of background, or "noise," items in target detection. The present study is concerned with two characteristics of background items, uniformity and visual similarity to the target, which may affect visual search in color displays.

Neisser's (1963) experiments showed that target-background similarity could impose serious constraints on search efficiency. An angular target letter was more easily detectable in a context of curved letters than of other angular letters, but this advantage was reversed when a curved target letter was specified. The effects of visual confusability between target and background have been confirmed in several more recent studies (e.g., Atkinson, Holmgren, & Juola, 1969; Corcoran & Jackson, 1977; Estes, 1972; Kaplan, Yonas, & Shurcliff, 1966; McIntyre, Fox, & Neale, 1970).

One line of evidence that background heterogeneity may also influence performance comes from an early

study by Eriksen (1953), who found that increasing the number of dimensions by which display elements varied produced a decrement in search efficiency. Green and Anderson (1956) reported that although search time was dependent primarily upon the total number of items in the display, it was slightly longer for multicolored than for single-color displays when subjects were not informed of the target's color. Similarly, Cahill and Carter (1976), who had subjects search for three-digit numbers in displays of 10 to 50 items coded in 1 to 10 colors, found that search time increased as a function of color code size (number of colors used in the display) even when the number of items per color category was held constant. A series of studies by Gordon and his colleagues (Gordon, 1968, 1969; Gordon, Dulewicz, & Winwood, 1971) clearly showed search time to be an increasing function of background heterogeneity. When subjects were required to cancel all relevant items (e.g., lower-case *a*s) detected in printed lists of constant size, search time/item increased with the number of *different* irrelevant items. Analogous findings have been derived from tachistoscopic detection studies. For example, Estes (1972) and McIntyre, Fox, and Neale (1970) found that increasing the redundancy of background items facilitated detection.

In the present experiments, color patches were selected as search material in preference to the more commonly used alphabetic characters in order to achieve greater control of interitem similarity. While problems are encountered in attempts to define and measure the similarity of visual forms (e.g., Attneave & Arnoult, 1956; Singer & Lappin, 1976), the similarity of colors can be specified in terms of their Euclidean distance in Munsell color space, a three-

The authors wish to thank Andy Belyavin for assistance in statistical analysis. Requests for reprints should be sent to Eric W. Farmer, General Psychology Section, RAF Institute of Aviation Medicine, Farnborough, Hants., England.

dimensional similarity space defined by a vertical axis of value (lightness), a radial axis of chroma (saturation), and a circular axis of hue. The basic postulates of this system have been supported by multidimensional scaling of subjective estimates of color similarity (e.g., Farmer, Taylor, & Belyavin, in press; Indow & Kanazawa, 1960).

The geometric structure of the Munsell system afforded an opportunity to vary background heterogeneity *independently of target-background similarity*. In contrast to most previous studies, heterogeneity could be manipulated without recourse to variation of the number of different background items.

EXPERIMENT 1

In this experiment, subjects were required to search for, and to signal the presence or absence of, an achromatic target which was equally similar to each chromatic background item in the display. The level of target-background similarity was varied by selecting targets differing in Munsell value. Further, the effects of background heterogeneity were assessed by varying the interitem similarity of the background set independently of target-background similarity.

Method

Subjects. Six male and four female volunteers, members of the staff of the RAF Institute of Aviation Medicine, each participated unpaid in two experimental sessions. All had normal color vision, as indicated by the Ishihara Tests for Color-Blindness.

Apparatus. The following apparatus were used in both experiments reported in this study: An electronic control unit directed the sequence of events within each trial. A Kodak Carousel projector with an Ealing electronic shutter was used to present the stimuli on a rear-projection screen located in a darkened room. Subjects responded "present" or "absent" by pressing one of two response keys set 80 mm apart. Depression of a key stopped an electronic digital timer which recorded reaction time (RT) to the nearest millisecond. A Hewlett-Packard HP 5055A digital recorder printed RT and response for each trial.

Stimuli. The stimuli were color chips from the Munsell Color File, Glossy Finish Collection. The target set comprised neutrals with Munsell notation N3/, N4/, N5/, N6/, and N7/.¹ Four different background sets were selected, each comprising five colors of constant value and chroma (5 and 4, respectively) but differing in hue. Two of the sets comprised hues selected from the 100-step Munsell hue circuit at 20-step intervals (low-similarity backgrounds); the remaining two comprised hues selected at 2.5-step intervals (high-similarity backgrounds). These sets are listed in Table 1.

Table 1
Low- and High-Similarity Background Sets Used in Experiment 1

Set	Within-Background Similarity	
	Low	High
1	5R, 5Y, 5G, 5B, 5P	10RP, 2.5R, 5R, 7.5R, 10R
2	5RP, 5YR, 5GY, 5BG, 5PB	10G, 2.5BG, 5BG, 7.5BG, 10BG

Note—Scores are Munsell hue notations of each item (R = red, T = yellow, G = green, B = blue, P = purple). Munsell value and chroma scores were constant at 5 and 4, respectively.

This method of selecting stimuli ensured that a target item was equally similar to each background item. Applying Godlove's (1951) color-difference formula, target-background "distances" were 4 units for target N5/, 5.7 units for target N4/ or target N6/, and 8.9 units for target N3/ or target N7/.

Six hundred 15-element arrays (5 rows by 3 columns) were constructed from these materials. Half of the arrays contained no target, and were allocated equally to each background set. The order of appearance of items within the negative (target-absent) arrays was determined by a PDP-12 computer, which randomly allocated equal numbers of each of the five background colors within a set to the 15 array locations, with the constraint that identical items could occupy no more than two consecutive array locations horizontally or vertically. Positive (target-present) arrays were generated by substituting a target item for one background item in each negative array. Within the group of 75 positive arrays constructed from each background set, each target appeared once in each of the 15 array locations and replaced each background item equally often.

In order to prepare the targets and arrays for presentation to subjects, Munsell color chips were placed at appropriate positions on a baseboard (Munsell notation N2.5/) and photographed on Ektachrome Professional Film. As a preliminary measure, examples of the material were photographed at a range of exposures, and the resulting transparencies were matched visually to the originals. This permitted the selection of an exposure that provided the optimal visual approximation to the Munsell colors. Although the need to reproduce the material photographically, for practical reasons of cost, necessarily produced some degree of color distortion, visual matching indicated that this distortion did not seriously influence the similarity relationships of the stimuli, and was unlikely to produce significant contamination of the experimental results.

When projected, a target item appeared near the top of the screen and to the left of the area occupied by the array. Each projected color chip subtended a visual angle of 2.6 deg horizontally $\times$ 3.2 deg vertically. The dimensions of an entire array were 9 deg horizontally $\times$ 19 deg vertically. Luminance ranged from 26 cd/m² for target N3/ to 394 cd/m² for target N7/.

Procedure. Subjects were informed that on each trial they would be presented with an achromatic target item followed by an array of colors in which the target was equally likely to be present or absent. It was explained that no member of the target set, except the item defined as the target for that trial, would appear in the array. Subjects were instructed to indicate the presence or absence of the target by pressing the appropriate response key. Conventional speed/accuracy instructions were given, emphasizing greatest possible speed within the limitations imposed by maintenance of a high level of accuracy. Half of the subjects responded "present" with the dominant hand and "absent" with the nondominant hand; the remaining subjects adopted the reverse response configuration.

On each trial, a 500-msec auditory warning signal was followed, after an interval of 250 msec, by presentation of the target for 2,000 msec. The array was presented after a further interval of 1,500 msec and was cancelled when the subject depressed a response key. The intertrial interval was 1,500 msec.

Each subject completed a total of 640 trials in two sessions, separated by approximately 2.5 h. There was a short rest period after the completion of each set of 40 consecutive trials. The first 40 trials, in which each experimental condition was equally represented, were for practice only and were excluded from statistical analysis. Trial presentation order was randomized anew for each subject.

Results and Discussion

The data obtained from one male subject were discarded because of his failure to conform to the experimental instructions. For the remaining subjects,

Figure 1. Mean RT in each experimental condition of Experiment 1.

mean $\ln$ -transformed² RTs were calculated for correct responses on the 30 trials (15 trials for each of two background sets) within each condition. These data, which, for the purpose of illustration (Figure 1), have been back-transformed to the original scale and then corrected for transformation bias (John & Quenouille, 1977), were analyzed by a four-way, within-subjects ANOVA (positive/negative trials, within-background similarity, target value, subjects). Since the prediction had been made that RT would be an increasing function of target-background similarity, and hence have an essentially quadratic relationship with target value, the df for all terms which incorporated target value were split between quadratic and residual components.

"Present" decisions were reliably faster than "absent" decisions [$F(1,8) = 97.34$, $p < .001$]. It was also confirmed that the quadratic trend for the main effect of target value was reliable [$F(1,8) = 37.10$, $p < .001$], and that the residual component did not approach significance ($F < 1$). Figure 1 shows that RT was influenced systematically by target-background similarity on positive trials, and that a comparable, but slightly less consistent, effect was obtained on negative trials. Analysis of the interaction between target value and positive/negative trials indicated that the quadratic term was reliable [$F(1,104) = 18.93$, $p < .001$], but that the residual term was not [$F(1,104) = 1.39$, $p > .2$]. Thus, although both positive and negative trials yielded quadratic RT functions, these functions differed in form.

High-similarity backgrounds produced considerably faster RTs than low-similarity backgrounds [$F(1,8) = 191.48$, $p < .001$]. This result obtained both on

positive trials [$F(1,104) = 43.75$, $p < .001$] and negative trials [$F(1,104) = 499.26$, $p < .001$], but was reliably greater for the latter, as confirmed by the presence of a significant interaction between background similarity and positive/negative trials [$F(1,104) = 123.70$, $p < .001$].

The final significant effect was the quadratic component of the interaction between target value and within-background similarity [$F(1,104) = 9.05$, $p < .005$], which indicated a discrepancy in the search-time functions for high- and low-similarity backgrounds. The steeper function obtained for the latter suggests that the benefits of decreasing target-background similarity were greater in the more difficult, low-similarity background conditions.

The mean error rate was 2.63%. Error rates computed for each condition were too low to be subjected to detailed statistical analysis. However, visual inspection of these data indicated that they were in reasonable agreement with RT, an exception being the tendency for accuracy on negative trials to be superior to that on positive trials.

Clearly, the major experimental hypotheses were confirmed. RT was an increasing function of target-background similarity both on positive and on negative trials. Further, background heterogeneity was a major determinant of RT, especially on negative trials.

EXPERIMENT 2

It has been shown that within-background similarity can influence RT in a search task. The next experiment is addressed to the question of whether interitem similarity represents only one instance of a more general perceptual variable of background uniformity. This issue is explored by requiring search through heterogeneous, low-similarity backgrounds whose organization is varied. Thus, background items may be allocated randomly to the display locations (unstructured condition) or the same items may be allocated such that each of the five rows within the array comprises items of a single color (structured condition). Cahill and Carter (1976) speculated that increasing background regularity in this manner might improve search performance. Moreover, it has been reported that the perceptual organization of multielement arrays can influence target detection (e.g., Banks & Prinzmetal, 1976; Prinzmetal & Banks, 1977).

A further aim of this experiment was to determine whether target-background similarity influences RT when subjects are not precued as to which member of the target set may appear within the array. Replication of the findings of Experiment 1 under these conditions would preclude the simple explanation that increasing target-background similarity merely

increases the time required to reject each irrelevant item.

Method

Subjects. Six male and four female volunteers, all members of the staff of the RAF Institute of Aviation Medicine, served as unpaid subjects. Six of these subjects had also participated in Experiment 1. All had normal color vision.

Stimuli. Only three achromatic target items (N3/, N5/, and N7/) were employed. Background items comprised Low-Similarity Set 1 of the previous experiment (see Table 1).

A total of 180 arrays were prepared. The structured and unstructured negative conditions each comprised 15 sets of three identical arrays. In the structured conditions, the arrays were generated by assigning the three examples of each color to a single row, such that the frequency of occurrence of each color in each row was balanced; in the unstructured condition, the three examples of each color were assigned randomly to array locations with the constraints that (1) each background item appear equally often in each location, and (2) items of the same color not occupy consecutive locations horizontally or vertically. Positive arrays were produced by replacing one background item in each set of three identical negative arrays by each of the three possible targets. Within the 15 sets of structured and unstructured positive arrays, each target appeared equally often in each array location and replaced each background item equally often.

The arrays were identical in all other respects to those described in Experiment 1.

Procedure. The procedure was similar to that followed in Experiment 1, with the following exceptions. A target item was *not* presented on each trial, but, instead, subjects were allowed to inspect the target set at the beginning of the session and were informed that each array was equally likely to contain one of the targets or no target. On each trial, the auditory warning signal was followed after 250 msec by presentation of an array. The intertrial interval was 2,000 msec.

Each subject completed 24 practice trials, drawn equally from each experimental condition. The 180 experimental trials were then administered in a single session, with a short rest pause on completion of each set of 60 consecutive trials.

Results and Discussion

Mean ln-transformed RTs for correct responses were analyzed by a four-way, within-subjects ANOVA (positive/negative arrays, structured/unstructured backgrounds, background replication, subjects). The term "background replication" denotes comparison of RTs within the sets of three arrays that were identical in background composition. On positive trials, these arrays differed only in the Munsell value of the target item; on negative trials, the arrays were identical.

Back-transformed RTs are shown in Figure 2. The analysis revealed that, as in Experiment 1, the positive decision was more rapid than the negative decision [$F(1,9) = 92.69$, $p < .001$]. Background replication was a reliable main effect [$F(2,18) = 16.84$, $p < .001$] but interacted with positive/negative arrays [$F(2,18) = 18.97$, $p < .001$]. The Newman-Keuls procedure indicated that there was a reliable effect of background replication on positive trials, that is, a difference in RT for arrays identical in background composition but differing in target value ($p < .01$) but, as expected,

Figure 2. Mean RT in each experimental condition of Experiment 2.

no reliable effect of background replication on negative trials. On positive trials, RT for targets N3/ and N7/ did not significantly differ, but each was significantly faster than that for target N5/ ($p < .01$). In other words, RT increased as a function of target-background similarity.

Background structure emerged as a reliable main effect [$F(1,18) = 172.02$, $p < .001$] and interacted with positive/negative trials [$F(1,9) = 54.23$, $p < .001$]. The source of this interaction was a strong advantage for structured backgrounds on negative trials [$F(1,18) = 77.85$, $p < .001$] but no reliable difference between RTs for structured and unstructured backgrounds on positive trials ($F < 1$).

The mean error rate in this experiment was 2.72%. Although errors were in general too infrequent to permit meaningful statistical analysis, the fairly large difference in error rates for structured and unstructured backgrounds containing target N5/ (5.33% and 10.00%, respectively) was examined in order to assess whether a speed-accuracy tradeoff had contributed to the failure to obtain a reliable effect of background

structure on RT in the positive conditions. However, comparison of arcsin-transformed error data failed to reveal a reliable difference [$t(9) = 1.32, p > .10$].

These results provide evidence that failure to specify which member of the target set may appear within the array does not abolish the effects of target-background similarity on RT. Moreover, they extend the generality of background uniformity effects to the organization of background items within the array, although only negative RT was reliably affected by this factor.

GENERAL DISCUSSION

The findings of these experiments provide clear evidence that increasing the visual similarity between target and background items, and decreasing the uniformity of background items, may each have a detrimental effect on RT. Moreover, they confirm that decreasing background uniformity influences RT even when target-background similarity is controlled.

Consider first the role of target-background similarity in these experiments. RT increased regularly as the "distance" in Munsell color space between target and background items decreased, even when the target did not appear within the array (Experiment 1) or the target was not uniquely specified (Experiment 2). The latter finding argues against the adequacy of the notion (e.g., Neisser, 1963) that visual confusability between target and background inhibits search performance by increasing the processing time required for the rejection of irrelevant items. Rather, it appears necessary to postulate that dissimilarity may also help to guide search towards regions of the display likely to contain the target.

A logical implication of this assumption is that the display may be subjected initially to a rapid global scan before more detailed information is extracted from suspect array locations. Recent two-stage models of search proposed by Hoffman (1978, 1979) and Snyder (1972) are consonant with this possibility. A common theme of these models is that parallel, but error-prone, visual analysis of the display occurs at an early, preattentive stage of processing, to be succeeded by a slower, serial, processing stage which is responsible for identification (cf. Kahneman, 1973; Neisser, 1967).

During preattentive processing, grouping of input items is assumed to proceed in accordance with Gestalt principles such as similarity and proximity. Recent studies by Banks and Prinzmetal (1976) and Prinzmetal and Banks (1977) have demonstrated the importance of Gestalt factors in determining the ease with which a target may be detected in the presence of noise items. As an alternative to feature-analytic approaches, in particular Estes' (1972, 1974) interactive channels

theory, these authors emphasized the critical role of preattentive grouping in mediating a variety of effects, including those of target-background similarity and background homogeneity. Central to their argument was the assumption that performance may be predicted from the degree to which preattentive grouping causes the target item to merge with or emerge from irrelevant items in the display. In support of this assumption, they showed that targets which possess a relationship of good form or of good continuity with noise elements are detected less readily than targets which lack these relationships.

Banks and Prinzmetal's "configurational" hypothesis correctly predicts that decreasing visual similarity between target and background items will improve performance even when subjects are not cued as to which target may be presented, because a dissimilar item will emerge from other items and so be detected with greater ease. But it is necessary to account for the finding that target-background similarity had demonstrable effects on *negative* RT. This is, in fact, also consistent with a two-stage conception of search. Since it is assumed that preattentive processes direct focal attention towards suspect array locations, negative RT should benefit from decreasing confusability of target and background items because of a concomitant decrease in the probability that focal attention will be directed towards irrelevant items. In terms of Hoffman's (1979) model, for example, only those irrelevant items reaching a criterion level of visual similarity to the target will be placed on a candidate list for transfer to Stage 2 (focal-attentive) processing.

In summary, it appears that at least two factors may contribute to the effects of target-background similarity. One is guidance of search towards items which are grossly dissimilar to other items in the display; the second is variation in the degree to which irrelevant items are processed.

The effects of background uniformity, the second major variable investigated in these experiments, support the notion that preattentive grouping operations help to determine the speed with which a target is detected. The reduction in RT observed in the high-similarity background conditions of Experiment 1 is predicted because similarity grouping of irrelevant items should facilitate emergence of the target. Experiment 2 showed that the relatively poor performance for low-similarity backgrounds could be improved by organizing background items into orderly regions of local homogeneity. Not surprisingly, this manipulation was rather less effective than increasing the similarity of background items, since it created distinct subgroups of irrelevant items rather than a relatively homogeneous background.

The results of both experiments indicated that background uniformity exerted a stronger influence

on negative than on positive RTs. This finding suggests that preattentive grouping of irrelevant items plays a more important role in permitting the rapid rejection of these items than in guiding search towards the item considered likely to be the target.

REFERENCES

- ATKINSON, R. C., HOLMGREN, J. E., & JUOLA, J. F. Processing time as influenced by the number of elements in a visual display. *Perception & Psychophysics*, 1969, **6**, 321-326.
- ATTNEAVE, F., & ARNOULT, M. D. The qualitative study of shape and pattern perception. *Psychological Bulletin*, 1956, **53**, 452-471.
- BANKS, W. P., & PRINZMETAL, W. Configurational effects in visual information processing. *Perception & Psychophysics*, 1976, **19**, 361-367.
- CAHILL, M. C., & CARTER, R. C. Color code size for searching displays of different density. *Human Factors*, 1976, **18**, 273-280.
- CORCORAN, D. W. J., & JACKSON, A. Basic processes and strategies in visual search. In P. M. A. Rabbitt & S. Dornic (Eds.), *Attention and performance VI*. London: Academic Press, 1977.
- ERIKSEN, C. W. Object location in a complex perceptual field. *Journal of Experimental Psychology*, 1953, **45**, 126-132.
- ESTES, W. K. Interactions of signal and background variables in visual processing. *Perception & Psychophysics*, 1972, **12**, 278-286.
- ESTES, W. K. Redundancy of noise elements and signals in visual detection of letters. *Perception & Psychophysics*, 1974, **16**, 53-60.
- FARMER, E. W., TAYLOR, R. M., & BELYAVIN, A. J. Large color differences and the geometry of Munsell color space. *Journal of the Optical Society of America*, in press.
- GODLOVE, I. H. Improved color-difference formula, with applications to the perceptibility and acceptability of fadings. *Journal of the Optical Society of America*, 1951, **41**, 760-772.
- GORDON, I. E. Interactions between items in visual search. *Journal of Experimental Psychology*, 1968, **76**, 348-355.
- GORDON, I. E. Eye movements during search through printed lists. *Perceptual & Motor Skills*, 1969, **29**, 683-686.
- GORDON, I. E., DULEWICZ, V., & WINWOOD, M. Irrelevant item variety and visual search. *Journal of Experimental Psychology*, 1971, **88**, 295-296.
- GREEN, B. F., & ANDERSON, L. K. Color coding in a visual search task. *Journal of Experimental Psychology*, 1956, **51**, 19-24.
- HOFFMAN, J. E. Search through a sequentially presented visual display. *Perception & Psychophysics*, 1978, **23**, 1-11.
- HOFFMAN, J. E. A two-stage model of visual search. *Perception & Psychophysics*, 1979, **25**, 319-327.
- INDOW, T., & KANAZAWA, K. Multidimensional mapping of colors varying in hue, chroma and value. *Journal of Experimental Psychology*, 1960, **59**, 330-336.
- JOHN, J. A., & QUENOUILLE, M. H. Experiments: Design and analysis. London: Griffin, 1977.
- KAHNEMAN, D. *Attention and effort*. Englewood Cliffs, N.J.: Prentice-Hall, 1973.
- KAPLAN, C. A., YONAS, A., & SHURCLIFF, A. Visual and acoustic confusability in a visual search task. *Perception & Psychophysics*, 1966, **1**, 172-174.
- MCINTYRE, C., FOX, R., & NEALE, J. Effects of noise similarity and redundancy on the information processed from brief visual displays. *Perception & Psychophysics*, 1970, **7**, 328-332.
- NEISSER, U. Decision-time without reaction time: Experiments in visual scanning. *American Journal of Psychology*, 1963, **76**, 376-385.
- NEISSER, U. *Cognitive psychology*. New York: Appleton-Century-Crofts, 1967.
- PRINZMETAL, W., & BANKS, W. P. Good continuation affects visual detection. *Perception & Psychophysics*, 1977, **21**, 389-395.
- SINGER, M. H., & LAPPIN, J. S. Similarity: Its definition and effect on the visual analysis of the complex displays. *Perception & Psychophysics*, 1976, **19**, 405-411.
- SYNDER, C. R. R. Selection, inspection and naming in visual search. *Journal of Experimental Psychology*, 1972, **92**, 428-431.

NOTES

1. In the Munsell system, the notation NV/ describes a neutral with value V, where value extends from 0 (absolute black) to 10 (absolute white).
2. The ln transformation was employed because means and variances were found to be positively correlated.

(Received for publication August 15, 1979;
revision accepted December 11, 1979.)