Classification without identification in visual search

Joan Brand

* Institute of Experimental Psychology, Oxford University,
CLASSIFICATION WITHOUT IDENTIFICATION IN VISUAL SEARCH

JOAN BRAND†
Institute of Experimental Psychology, Oxford University

Six subjects scanned displays of random consonants for a single target which was (a) another consonant; (b) a given number; or (c) any number. A second group of six subjects took part in three comparable conditions with number displays, and letters or numbers as targets. Scanning time for a number in a letter display or a letter in a number display was more rapid than scanning for a target drawn from the same set as the background. Several unpractised subjects, and all the subjects who practised the task, were able to scan as fast through letters for "any number" as for a specific number, or conversely through digits. The finding of different scanning rates for two precisely physically specified targets, depending on which class they were drawn from, runs counter to an explanation of high-speed scanning in terms of the operation of visual feature analysers. It is suggested that familiar categorization responses may be immediate and may provide the basis for the discrimination of relevant from irrelevant items in rapid visual scanning.

Introduction

In visual search tasks such as those described by Neisser (1963, 1964), the searcher is pushed two ways: he has to scan and reject large numbers of irrelevant items as fast as possible, while attending closely enough to detect a single target when it does occur. He must strike a rather delicate balance between examining the display so carefully that he does not miss targets and "skimming" it rapidly enough so as not to waste time on non-targets. An important factor in the achievement of this balance must be the ratio of non-target to target items in the display; in the experiments cited, this has been very high indeed, usually of the order of 99:1 (Neisser, 1963). With as low a signal frequency as this the subject may well feel he has more to lose by being careful than by being fast; and phenomenal reports at least suggest that his examination of non-target items is cursory indeed. He retains no memory for them, and they even appear during the search as "just a blur" (Neisser, 1964).

A rapid perceptual scanning process which results neither in memory nor even in very detailed percepts, yet which still allows signal detection is plausibly seen as an abbreviated form of the full recognition process (Neisser and Beller, 1965) in which only low-level visual feature analysers are activated and more complex levels of analysis held back. It is the aim of this study to see whether high speeds in visual scanning might also be the outcome of a selective analysis of the display for more complex stimulus characteristics. Subjects were asked to search for targets which were defined either by an individual description, such as "K", or

† Present address: Medical Research Council, Applied Psychology Unit, Cambridge.
by a class description, such as "any letter". In scanning for an individual target, the subject knows its precise physical appearance, and is in a position to scan the display for physical features of the target by means of rapid parallel feature analysers. In scanning for a class of items, however, the physical features of the target are unknown except as a set of alternatives; and even a comprehensive listing of the set may fail to identify the class. The classes of "letters" and "digits", for example, are made up of much the same simple component features; it wants at least an analysis of combinations of features to distinguish a letter from a digit. Thus when a visual search target is defined only as "any instance" of a particular class of items, such as letters, and not given any physical description, the searcher cannot rely solely on simple feature analysis but needs to interrogate the display at more complex levels. The question this experiment was designed to answer is whether with this type of target definition he can still attain high scanning speeds such as those reported by Neisser.

The class descriptions "a letter" and "a digit" were chosen partly because of the common physical features which make up the two sets; and partly because there is already some evidence that the letter/digit disjunction is an effective basis for the classification of information. Alternate presentation of letters and digits results in poorer short-term retention than grouped presentation (Broadbent and Gregory, 1964; Sanders and Schroots, 1968), and similarly reduces the span of apprehension (Warrington, Kinsbourne and James, 1966). Brown (1960) found that instructions to report only the letters (or digits) from a tachistoscopically presented mixed set were effective if given before but not after presentation; and Posner (1970) found no adverse effect of acoustic similarity on the time subjects took to discriminate letters from digits, in contrast to its effect on the discrimination of one letter from another. These results suggest the possibility that, with these familiar classes, individual items of input may be assigned to a class without detailed processing of the individuals themselves; and that the classifying response may provide the basis of perceptual selection. Such a possibility does not fit easily into models of cognitive operation which assume an hierarchical or sequential order of increasingly sophisticated analysis of inputs, working from the perception of crude physical characteristics to that of semantic content. Two notable examples would be Neisser and Beller's (1965) explanation of their visual search findings, and Treisman's selective attention model (Treisman, 1964; Treisman and Geffen, 1967); within the terms of these models class-membership is clearly a semantic feature.

Method

Twelve subjects took part in the experiment; they were nine men and three women aged between 16 and about 26 years old, members either of Oxford University or of the Oxford College of Further Education. They were paid for their services at the rate of £0.25/hr.

Each stimulus display was a bank of letters or digits arranged in matrix form with 6 characters in every row and 35 in every column. Vowels were excluded from the alphabet of letters and 0 and I from that of digits. A PDP8 computer was used to generate and print random sequences of characters in matrix form, and to insert one of a set of possible target characters into one of five predetermined target rows, namely, row 7, 12, 19, 25 or 32. Targets were distributed randomly over the first five spaces in a row. The matrices were
displayed to the subject one at a time at the back of a simple viewing box, 56 cm distant from
his eyes as he sat with his head against an eyepiece at the front of the box. What he then
saw was a long vertical white field measuring $7 \times 17$ cm, with an area $2 \times 15$ cm in the
centre occupied by the matrix, and with its top marked by a fluorescent light as a fixation
point for the start of the search. At the beginning of each trial the box was dark save for
the fixation point, and the subject initiated the trial himself when ready by pressing a key
on his left. This illuminated the display by switching on a single 40-W light bulb, out of the
subject's field of view at the front of the box, and simultaneously triggered a centisecond
timer.

The instructions were to search the display by means of a single steady downward scan,
without pausing and without retracing; on finding the target, the subject pressed his right
hand key to stop the timer and switch off the light, and then reported the symbol which
immediately followed the target. The experimenter checked this for accuracy, recorded
the time taken, inserted a new display and told the subject the new target for the next trial.
The check on false positives which is provided by the report of the symbol following the
target was not used by Neisser in his experiments; a pilot study was run with just two
subjects to determine whether this requirement would alter searching strategy in any way,
and it was clear that it affected neither slope nor intercept of the search time function. If
the subject completed the scan without having found the target he was then allowed un-
limited time to locate it; this time was not recorded, nor was the trial repeated, but scored
as an omission error.

The experiment had a mixed design; the subjects were randomly assigned to one of two
groups of six subjects each. One group always searched through matrices of letters and the
other through matrices of digits. Within each group, every subject took part in three
conditions, defined by the different types of target:

1. A physically specified target of the same class as the background (for example,
   the target "K" in a background of letters).

2. A physically specified target of the opposite class to the background ("7" in a
   background of letters).

3. An unspecified target of the opposite class to the background ("any number" in a
   background of letters).

The three conditions were given in three blocks, in a different order to each subject in the
group, according to the Latin square principle. Within each block, the target appeared in
each of the five target rows on five occasions, making 25 trials per block and 75 per session.
Subjects were tested individually and were given 20 practice trials at the beginning of each
session, 5 for the specific target of the opposite class, 5 for the unspecified target of the
opposite class, and 10 for the same-class target.

Results

Search rate

Search rates were calculated as the ratio of the time taken to find the target and
its distance from the starting point of the search. Slope values representing
search rates in each condition are given in Table I. The linearity of the observed
slopes testified to the efficacy of the instructions to search downwards in a single
steady scan; none of the subjects claimed to be able to discriminate target rows
from non-target rows, but several said that they learned that the target never
appeared in the first five or six rows; thus for many of the graphs the intercept
may be artificially small.

The data were also subjected to two fourway analyses of variance. The first
of these compared the two conditions in which the target was physically specified.
A target drawn from a different class from the background items was more quickly detected than one drawn from the same class \((F_1, 10 = 27.9, P < 0.001)\), and there was a significant interaction between class of target and vertical position of target \((F_4, 40 = 9.53, P < 0.001)\), indicating a slower rate of search for a same-class target. The main effect of the two groups of subjects was not significant, nor were any of its interactions.

The second analysis of variance compared the two types of target, physically specified and unspecified, which differed in class from the background. The significant main effect of specification of target \((F_1, 10 = 12.7, P < 0.01)\), and its significant interaction with position of target \((F_4, 40 = 9.21, P < 0.001)\) indicated a slower rate of search for an unspecified different-class target than for one which was known prior to the search. Again, neither the main effect of groups nor any of its interactions were significant.

Inspection of the graphs for individual subjects, however, suggested that the second analysis of variance had obscured some interesting differences between individuals. Figure 1 shows the results of three individuals who showed three

![Graph showing search rates for different targets in letter contexts. Targets: ■—■, letter; △—△, known digit; O—O, unknown digit.](image)
distinct patterns of responding and who seem to represent the whole of the variation over the 12 subjects. For ease of inspection, the three subjects are all from the group who searched through letters; six subjects, three from each group, showed the pattern represented by KM, five (two from the letter group and three from the digit group) that by JH, leaving MW, from the letter group, as an isolated case. MW and JH showed two completely different patterns; for MW the physical specification of the target was the crucial factor; if it was physically specified he scanned for it rather fast and regardless of whether it belonged to the same set as the background or a different one, whereas if he knew only that it was "any number" he was much slower. JH, by contrast, scanned a display of letters faster for a number target than for a letter target, although both were physically specified, and was no faster if he knew exactly which number it was than if he did not. KM seems to fall between these two extremes; unlike MW, he scanned faster for a specific target if it was of the opposite class to the background, but was also helped by specification of the different-class target. In partial confirmation of these results, two separate analyses of variance were carried out on the subgroup of five subjects who showed the pattern represented here by the results of JH. These five subjects searched faster for a different-class specific target than for a same-class one ($F_{1, 4} = 25.9, P < 0.01$), but there was no significant difference arising from physical specification of the different-class target ($F_{1, 4} = 0.77, N.S.$).

**Errors**

The mean percentage omission error rate over all conditions was 6.5%. Table I shows that most of these were made in the conditions where the target was of the same class as the background, and that more were made by subjects searching through digits than by those searching through letters. Analysis of variance showed that this last difference approached but did not reach significance ($F_{1, 10} = 4.67, P < 0.10$), though there was a significant interaction between groups and target conditions ($F_{2, 20} = 7.86, P < 0.01$). There was a significant difference between target conditions ($F_{2, 20} = 21.97, P < 0.001$), and a Newman-Keuls test

<table>
<thead>
<tr>
<th>Target conditions</th>
<th>Same class</th>
<th>Different class</th>
<th>Error total</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Specific</td>
<td>Unspecified</td>
<td></td>
</tr>
<tr>
<td><strong>Letter Group</strong></td>
<td>0.53</td>
<td>0.39</td>
<td>0.50</td>
</tr>
<tr>
<td></td>
<td>(10)</td>
<td>(1)</td>
<td>(7)</td>
</tr>
<tr>
<td><strong>Digit Group</strong></td>
<td>0.49</td>
<td>0.30</td>
<td>0.45</td>
</tr>
<tr>
<td></td>
<td>(27)</td>
<td>(6)</td>
<td>(5)</td>
</tr>
<tr>
<td><strong>Error total</strong></td>
<td>37</td>
<td>7</td>
<td>12</td>
</tr>
</tbody>
</table>

Figures in brackets indicate the total number of errors made by the six subjects in each condition.
showed that there were significantly more errors in the same-class target conditions than in each of the two different-class conditions, and that these two did not differ.

The effects of practice

It seemed likely that the strategy shown by MW, that of reliance on knowledge of the physical features of the target, might only become feasible on better acquaintance with the particular alphabet and the particular type-face used; Rabbitt (1967) has shown that an important component of the practice effect in visual search is the learning of the critical cues which distinguish a particular target from a particular set of background letters. Four of the 12 subjects, two from each group, took part in four further sessions, making five sessions in all. The results were contrary to this expectation; Table II shows the results from the first and the fifth sessions. MW did not even maintain his initial strategy with practice; by the fifth session his rate of scan for an unknown target of different class had improved more than that for the two physically specified targets, and no longer differed from rate of scan for a specific different-class target. Both were very slightly faster than scan for a same-class target. KM was one of the subjects in the initial experiment who seemed to combine the two strategies of using physical and class information to scan for the target. By the fifth session, far from showing more reliance on knowledge of the target’s physical appearance, he had gone the opposite way—he scanned equally fast for the two different-class targets and more slowly for the same-class target.

### Table II

Slope values for the first and fifth sessions of two subjects from each group

<table>
<thead>
<tr>
<th>Target conditions</th>
<th>Same class</th>
<th>Different class</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Specific</td>
<td>Unspecified</td>
<td></td>
</tr>
<tr>
<td>MW</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Session 1</td>
<td>0.32</td>
<td>0.33</td>
<td>0.62</td>
</tr>
<tr>
<td>Session 5</td>
<td>0.22</td>
<td>0.13</td>
<td>0.19</td>
</tr>
<tr>
<td>Digit group</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>KM</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Session 1</td>
<td>0.64</td>
<td>0.38</td>
<td>0.54</td>
</tr>
<tr>
<td>Session 5</td>
<td>0.37</td>
<td>0.13</td>
<td>0.11</td>
</tr>
<tr>
<td>CM</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Session 1</td>
<td>0.33</td>
<td>0.11</td>
<td>0.21</td>
</tr>
<tr>
<td>Session 5</td>
<td>0.19</td>
<td>0.14</td>
<td>0.11</td>
</tr>
<tr>
<td>Letter group</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>QS</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Session 1</td>
<td>0.39</td>
<td>0.26</td>
<td>0.28</td>
</tr>
<tr>
<td>Session 5</td>
<td>0.28</td>
<td>0.11</td>
<td>0.17</td>
</tr>
</tbody>
</table>
CM and QS, the two subjects from the digit group who practised the task, both searched more slowly on the first session for a same-class target and equally fast for the specific and unspecified different-class targets; and both maintained their preference over practice.

**Discussion**

The results of the experiment may be summarized as follows.

1. Eleven of the 12 subjects were able to scan faster for a specific number target on a display of letters than for a letter target, or conversely.

2. Five of the unpractised subjects could scan a display of letters for an unspecified number target as fast as for a specific number target, or conversely.

3. These differences in search rate were not due to a simple case of speed-accuracy trade-off across conditions; the conditions in which target and background items were both drawn from the same set were also those in which most errors occurred.

4. The four subjects who practised the task all tended to converge on the strategy of making most use of a distinction in class-membership between target and background, and less of the unique physical representation of the target.

The second and fourth of these results show that subjects were able to scan a visual search display as rapidly for up to 20 physically different targets as for a single one. This finding is not novel in visual search: Neisser, Novick and Lazar (1963) reported scanning for ten different targets as rapidly as for one. The subjects in the present experiment, however, achieved an equivalent rate of scan for the multiple target after little or no practice and without a dramatically large error rate; Neisser, Novick and Lazar's subjects had had fourteen sessions of practice, and made errors at rates of up to 20%. Both the high error rate and the intensive practice required were consistent with Neisser's model of visual search by feature analysis, a low-level, parallel and error-prone operation which depends on thorough learning of features. The lower error rate and little practice of the present subjects might suggest that their performance demands a different explanation; and this is necessitated by result (1), showing different rates of search for two single targets whose physical features are specified.

The multiple physically different targets of the present experiment were of rather a special kind; they were drawn exclusively from a single familiar class of items, they were described to the subject only by means of the name of the class; and the boundaries of the class coincided with the boundaries between the class of target items and the class of background items in the display. Thus a response which classified an item as "a letter" or "a digit" without previously identifying it as a particular letter or digit would serve to explain the equivalent rate of search for "any one of 20 letters" or "the letter K". Two of the subjects who practised the task claimed that they were in fact doing this; in scanning through letters, for example, they would set themselves to scan for "a digit" even if told that the target was "3".
This result is in line with Posner's (1970) reaction time data, which indicate that subjects arrive at the classifications "letter" and "digit" without first obtaining the name. It also provides a direct demonstration of the cogency of Rabbitt's (1967) analogy between search and stimulus categorization. The experimental situation differs from Rabbitt's and from Neisser's (1963) in that it makes use of a non-arbitrary, finite and highly familiar category as the basis of discrimination in search, rather than an arbitrary subset of items within a category. The effect of familiarity of the category seems to be that the subject requires little or no practice before being able to use the categorization response as the basis for an efficient and rapid search, with search rate insensitive to the number of items within a class. Indeed, the categorization of letters and digits appears to be so readily available a response that it is not superseded after practice by discrimination based on visual features, although feature learning might be expected to be one of the main effects of practice.

This account of visual search performance resembles Neisser's in that it assumes that the subject does not identify the background material. It differs from Neisser's in that, instead of proposing that identification does not occur because analysis of background items is beyond the level of visual feature analysis, it sees the limitation on the full perceptual process as due to a selective factor; out of the set of possible responses to the presented items the subject focusses on one, namely their membership of a familiar class. The claim that high-speed scanning may rely on a highly-learned classification of input items rather than identification or physical feature analysis cannot be generalized to classifications other than the one used. The digit/letter disjunction is already known to affect cognitive performance in a number of ways, and other classifications, apparently equally familiar, may be less effective. Posner's (1970) finding that acoustic confusability did not affect the discrimination of letters from digits was not replicated for the discrimination of vowels from consonants. Nevertheless, the present results open the possibility of visual scanning by complex characteristics, and emphasize the importance of well-learned classifications in the selection of information.

It should be noted that Neisser and Lazar (1964) obtained a discrepant result; in contexts of letters their subjects were consistently faster in searching for the target "3" than for "any numeral". The present experiment used more single characters than Neisser and Lazar did, and the results in fact seem to vindicate their suggestion that search for "any numeral" was based on a parallel match of all the numbers with display items, and that the speed of the parallel process was held to that of its slowest component, "3" being a distinctive character and therefore one of the faster components.

This work was supported by a Scholarship for Training in Research Methods, No. G 77/1643, from the Medical Research Council. I wish to thank Dr Anne Treisman for her helpful criticisms and discussion, and Dr J. R. Millenson and Geoff Sullivan for writing a computer program for generation of the stimulus displays.

References


Received 5 November, 1970