

The order of visual processing: Top-down, bottom-up, middle-out, or none of these?

J. GREANEY and A. W. MACRAE
University of Birmingham, Birmingham, England

To investigate whether global or local visual information is processed first, several researchers have used a large letter (global information) made up of smaller letters (local information) and measured speed of response and interference effects between global and local components. Kinchla and Wolfe (1979) argued that processing was neither invariably top-down (global-to-local) nor bottom-up (local-to-global). Instead, they argued that the fastest identification of letter shapes is given at an optimal visual angle of about 2° and proposed a middle-out processing model wherein elements of optimal size are processed first. In the present study, three-level figures were used in a direct test of that hypothesis, with subjects monitoring all three levels to detect target letters occurring at any level. The results are not compatible with any of the three orders of processing that have been proposed, and an interpretation in terms of the discriminability of the information available at each level is advocated. It is concluded that the two-level paradigm with letters is inadequate for investigations of the order of processing.

Three possibilities for the order of visual processing were outlined by Kinchla and Wolfe (1979). In "bottom-up" processing, local parts are registered first and are then combined to form global images. In "top-down" processing, recognition of a unitary, global whole is followed by that of progressively lower and more local levels of form. In "middle-out" processing, an intermediate level of form is recognized first, followed by higher and lower levels.

One way to investigate the order in which visual information is processed is to use a large letter (global information) made up of smaller letters (local information). Navon (1977) claimed an "invariable global precedence" based on faster response times to the global letters and asymmetrical interference between global and local components. However, subsequent studies demonstrated that local precedence can also occur. Kinchla and Wolfe (1979) found that the "local" letters of two-level letter patterns were detected faster than the global ones when the global shape exceeded about 7° in height, with the opposite order when they were smaller. Rather than any invariable "top-down" (global-to-local) or "bottom-up" (local-to-global) order, they hypothesized an optimal angle of about 2° for the fastest identification of letter shapes and proposed "middle-out" processing, in which elements of optimum size are processed first, followed by those that are larger or smaller.

In the present experiment, we used a three-level configuration in which letters were made up of smaller letters which in turn were made up of yet smaller letters. The subjects had to monitor all three levels for the occurrence of one of two target letters. The use of three-level stimuli allows the direct testing of the Kinchla and Wolfe (1979) "middle-out" hypothesis. If the intermediate size of letter subtends an angle of about 2° , response times to this level should be fastest.

METHOD

Apparatus and Stimuli

The stimuli consisted of a large letter made up of smaller letters in a 5×5 matrix, each of these in turn being made up of yet smaller letters, as illustrated in Figure 1. We follow Lamb and Robertson (1989) in using H and S as targets with A and E as distractors. Each configuration contained exactly one target letter (either H or S) and two distractors (an A and an E). All three letters in a given display were different, and the target letter could appear at any of the three levels; 12 displays were thus possible.

In two separate blocks of trials, the global letter subtended vertical angles of 8° and 12° , so that the intermediate letter subtended angles of 1.25° and 2° , respectively, and the smallest letters were 0.16° and 0.25° . Both global and local letters were 1.8 times as tall as they were wide. All stimuli were in black on white at the center of a tachistoscope card at a viewing distance of 50 cm. The visible part of the card measured approximately 100×110 mm (12° wide \times 13° high).

Subjects

Ten staff and students at the University of Birmingham, who were 22-40 years old and had normal vision, volunteered for the experiment without payment.

Procedure

The subjects sat in a dimly lit, soundproof room, looking into the tachistoscope. A switch controlled by the experimenter made an audible click to announce the beginning of each trial. The screen became illu-

J. Greaney holds a Research Studentship from the Science and Engineering Research Council. We are grateful to Westland Helicopters Ltd. for their additional support for our work. Correspondence should be addressed to A. W. MacRae, School of Psychology, University of Birmingham, Edgbaston, Birmingham, B15 2TT, England.

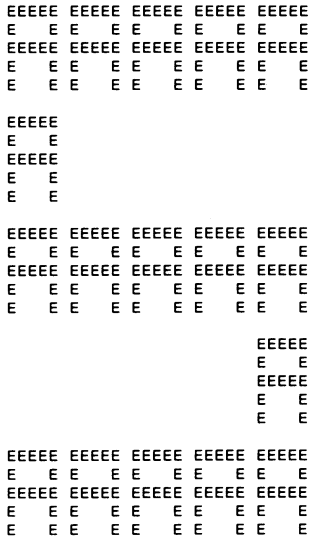


Figure 1. One of the 12 three-level patterns presented for the subject to detect the presence of either an S or an H.

minated for 1 sec, showing a fixation point. The test figure was then displayed at the same illumination; it remained visible until canceled by the subject's response. Response time (RT) was measured from the time the stimulus appeared until the subject pressed one of the buttons on the response panel to record whether an H or an S had been present in the figure. The response also caused the stimulus to disappear. The subjects were asked to respond as quickly and as accurately as possible.

On each of 2 days of testing, two blocks of trials were administered, separated by a short break. Sets of stimuli with the larger overall visual angle and of those with the smaller visual angle were presented in separate blocks in a different order on each day. At the beginning of each block, there were practice trials in which each figure was shown at least once. After the practice period, there followed 72 experimental trials in which each figure appeared six times in a randomized order.

RESULTS

Error rates were low, amounting to 1.9% of total responses, and analysis revealed no significant patterns of error. Figure 2 summarizes the RT results. A log transformation normalized the residuals, and a repeated measures analysis of variance with three factors was undertaken on the transformed scores. The factors were visual angle (8° or 12° overall), target level (global, intermediate, or local), and target letter (H or S).

There was clearly no performance advantage for the intermediate level; indeed the results were in the opposite direction, with a significant difference between target levels [$F(2,360) = 5.00, p < .01$]. The global level elicited longer RTs for the 12° overall angle as opposed to the 8° angle, whereas the opposite was true of the intermediate level, producing a highly significant interaction between angle and level [$F(2,360) = 31.33, p < .001$], while the overall advantage for 8° figures was not significant.

RTs were faster to H targets than to S targets [$F(1,360) = 4.54, p < .05$].

Several subjects mentioned what Ward (1982) called "the level readiness effect"; it was easier to detect a tar-

get when it appeared at the same level as it did in the preceding figure. One subject commented on a difficulty in refocusing from the largest level to the smallest level and vice versa. Two subjects described being aware of the identity of the target letter before they could locate it: "I knew there was an H there but could not see it"; and, "I knew there was an S somewhere but I had to look again to locate it."

DISCUSSION

The slower RTs for the intermediate level are not predicted by any of the theoretical accounts of the processing of global and local levels of visual information. This is not to say that our results cannot be accounted for—it is rather that we have shown that the multilevel letter paradigm may not be a good test for order of processing. The speed with which letters are recognized is affected by too many factors—for example, expectation, familiarity of the font, clarity of definition, confusability of pairs of letters in an experiment, and so on. And what is true of the multilevel paradigm holds at least as strongly for two levels, where the extent of such effects can be less clearly estimated.

The fact that the middle level was processed slowest in our experiment can be explained more easily by its physical characteristics than by its position in the hierarchy of levels, and this restricts the extent to which we should generalize from the letter paradigm to other situations involving the processing of global and local visual information.

Martin (1979) varied the number of local elements that made up the larger letter and found that when there were many local elements (and the global letter was thus well defined), global discrimination was faster than local. For stimuli with few local elements, the result was reversed. Ward (1982) claimed that features that are brighter, that stand out more from the background, that are nearer the fovea, or that are better formed are recognized more rapidly than others.

Our finding that 12° global letters were identified more slowly than 8° letters may be related to the greater spread of the 5x5 matrix forming the large letters and their greater average distance from the fovea. In addition, the larger figures had less white space around them and so may have appeared less unitary. Related factors can explain the difference in RT between H and S targets. In the set of letters that has been used for studies of this type, the S formed by a 5x5 matrix has an unfamiliar box-like shape with three pronounced horizontal segments. That alone may be enough to delay recognition, but we also suspect that there is greater interference for that shape of S than there is for

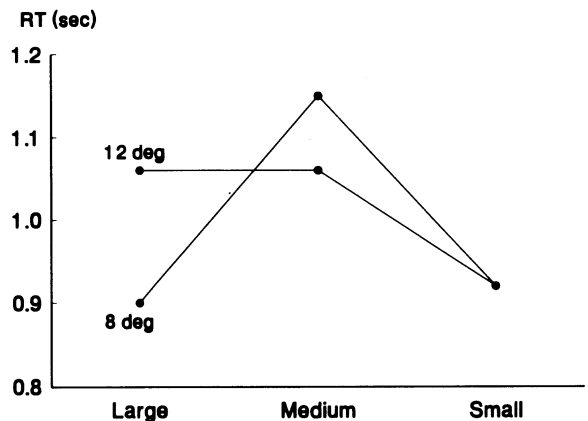


Figure 2. Geometric mean response times for the decision that an S or an H had been present in a three-level figure as a function of the visual angle of the whole figure and the size of letter (large, medium, or small) within the figure at which the target letter occurred.

H from the letters A and E, which have two and three horizontal segments, respectively.

We also believe that factors such as these are mainly responsible for the slower recognition of the middle size of letter in our experiment. In a three-level figure, the middle level is embedded within the display, having neither the advantage of being a unitary figure surrounded by white space (as with the largest letter) nor that of being made up of a continuous solid line (as with the smallest letter). The middle level is also subject to interference from two adjacent levels in the hierarchy, whereas the other two levels are not. Thus our result may be partly explained in terms of greater overall interference for the middle level. But the overall implication is clear: our results directly contradict the prediction of an advantage for elements of intermediate size.

There is no simple ordering of response time as a function of level, because the smallest and largest letters are identified faster than the intermediate one. Also, 2 subjects mentioned being aware of the identity of the target letter before being able to locate it. Thus to the extent that our results do relate to the order of processing of visual information they do not support *any* model predicated on a systematic processing order for levels. Instead, they support accounts involving a parallel extraction of features, with identification speed being based on the quality of the information available. That is essentially the conclusion also reached by Grice, Canham, and Boroughs (1983), who interpreted their data from experiments with two levels of letter size as evidence for a

continuous-growth model of information processing rather than a sequence of stages.

REFERENCES

- GRICE, G. R., CANHAM, L., & BOROUGHS, J. M. (1983). Forest before trees? It depends where you look. *Perception & Psychophysics*, **33**, 121-128.
- KINCHLA, R. A., & WOLFE, J. M. (1979). The order of visual processing: "Top-down," "bottom-up," or "middle-out." *Perception & Psychophysics*, **25**, 225-231.
- LAMB, M. R., & ROBERTSON, L. C. (1989). Do response time advantage and interference reflect the order of processing of global- and local-level information? *Perception & Psychophysics*, **46**, 254-258.
- MARTIN, M. (1979). Local and global processing: The role of sparsity. *Memory & Cognition*, **7**, 476-484.
- NAVON, D. (1977). Forest before trees: The precedence of global features in visual perception. *Cognitive Psychology*, **9**, 353-383.
- WARD, L. M. (1982). Determinants of attention to local and global features of visual forms. *Journal of Experimental Psychology: Human Perception & Performance*, **8**, 562-581.

(Manuscript received February 5, 1992.)