A conceptual category effect in visual search: O as letter or as digit*

JOHN JONIDES and HENRY GLEITMAN
University of Pennsylvania, Philadelphia, Pennsylvania 19104

Evidence is presented for a processing mechanism in visual recognition that depends upon how the stimulus array is conceptually categorized rather than upon its physical characteristics. Ss had to detect a letter or digit target in a field of letters or digits. When target and field were of the same category, reaction time increased with display size. When target and field category differed, reaction times were independent of display size. This category effect held even for the ambiguous target character O that yielded reaction time functions appropriate to how it was specified prior to presentation: as "zero" or as "ō."

Many tasks require an S to discriminate a particular stimulus item from other items among which it happens to occur. There are indications that such discriminations may be based not only upon characteristics inherent in the stimulus (e.g., square vs circle) but also upon the conceptual category to which the stimulus belongs (e.g., letter vs digit). The most relevant evidence for category-based discrimination comes from experiments on visual search.

In a visual search task, S looks for a specified target among a number of other items ("field items") and signals its presence or absence. Several studies that utilized this paradigm have focused upon the relation between reaction time (RT) and the number of items (n) in the display. When the target and field items are of the same conceptual category (e.g., an "A" amidst a number of other letters), RT has generally proved to be a linearly increasing function of n (Atkinson, Holmgren, & Juola, 1969; Nickerson, 1966; Sternberg, 1967). This effect has often been interpreted in terms of a serial search process in which display items are sequentially compared to an internal representation of the target.[1] Another result has been reported, however, for tasks in which target and field items belong to such different categories as letter vs digit (e.g., an "A" amidst several digits). Here the function relating RT to n has sometimes been found to be flat, suggesting a parallel processing mechanism in which all display items are dealt with simultaneously (Egeth, Jonides, & Wall, 1972). A related result was obtained by Brand (1971), whose Ss were generally faster when scanning down a long array of letters for a digit target than for a letter target. Similar effects were also reported by Posner (1970) and by Sperling, Budiansky, Spivak, & Johnson (1971) for still different experimental paradigms.

This pattern of results suggests that category discrimination is in some sense "easier" than character identification, a hypothesis proposed by Posner (1970) and Brand (1971) among others. But it is not clear that the "category effect" obtained in these various experiments was really based on the conceptual differences among the relevant stimuli. It may have been due simply to a physical feature difference between members of the categories employed in these studies. There is good evidence that RT functions in visual search are flat when target and field items are highly discriminable along such physical dimensions as shape, as demonstrated by Donderi and Case (1970) and Donderi and Zelnicker (1969) with a same-different paradigm. There is also evidence that Ss trained to discriminate among alphanumeric characters on the basis of appropriate distinctive features eventually attain flat RT functions (Yonas, 1969; using a card-sorting task, a similar effect has been obtained by Rabbitt, 1967). This distinction between letters and digits may be based on a similar physical difference (e.g., symmetry about some axis is more prevalent among letters than among digits; on many type faces, letters tend to be wider than digits). In short, the category effect may simply be an artifact of the stimulus materials.

To firmly establish the phenomenon as a result of conceptual categorization, one must demonstrate that the effect will still occur when all physical stimulus differences are completely controlled. To this end, the very same physical target stimulus, the symbol O, was utilized under two different conceptual sets. This target was presented as a digit ("zero") or as a letter (the vowel "ō") amidst field items, all of which were either letters or digits. To insure the appropriate mental set, targets that were unambiguous exemplars of the appropriate categories were used in addition.

*Supported in part by a Sigma Xi grant-in-aid of research (to J.J.) and in part by Grant NIH 20041-02 from PHS-NIH. We thank James Johnston, James McClelland, Jacob Nachmias, and Paul Rozin for helpful comments and criticisms. Requests for reprints should be addressed to: Department of Psychology, 3813 Walnut Street, University of Pennsylvania, Philadelphia, Pennsylvania 19104.

METHOD

Subjects

Thirty-six male and 12 female students of the University of Pennsylvania served as volunteer Ss. Each was paid for participation in a 40-min session.

Design

The Ss were assigned to eight groups of six each according to a 2 by 2 by 2 factorial design. The first factor was field category (of which all Ss were informed when the session began): for half of the Ss, the field items were letters; for the other half, they were digits. The second factor was type of response: half of the Ss ("presence responders") were required to press a telegraph key if the target was present but to make no response if it was absent; the other half of the Ss ("absence responders") were to press the key only when the target was absent. The last factor concerned the relation between target and field category. For half of the Ss, both targets and field items belonged to the same category (letter among letters or digit among digits); for the other half, targets and field items were of different categories.

All Ss were presented with three display sizes, 2, 4, and 6. On half of the trials, the target was present; on the other half, it was absent. The levels of all within-S variables were randomly distributed over trials.

Stimulus Materials

The stimuli were Letraset black uppercase letters and numerals (Alternate Gothic, No. 2, 12-point). For half of the Ss, the targets were the three uppercase letters A, Z, and O (specified as the vowel "ō"); for the other half, they were the three digits 2, 4, and O (physically identical to the O employed in the letter-target set, but here specified as "zero"). Each target was used equally often within a session. The field items were randomly chosen either from a subset of uppercase letters (all letters excluding A, B, D, G, M, O, Q, W, and Z) or from the subset of digits 1, 3, 5, 6, 8, and 9. Both target and field items (.43 deg in height) were located around the circumference of an imaginary circle of 3.4 deg in diam whose center coincided with a preexposure fixation point. Each array was flashed for 150 msec in a two-field mirror tachistoscope (Polymetric

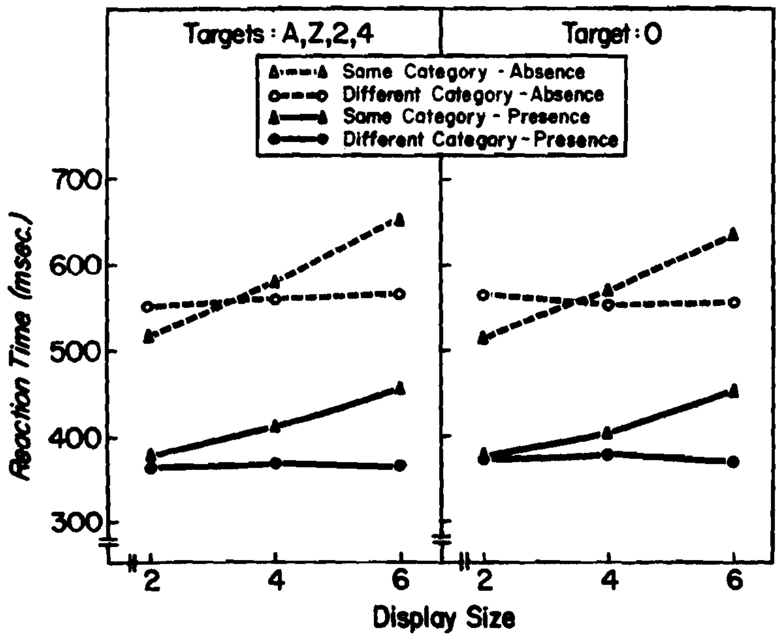

Fig. 1. Mean reaction time as a function of display size (n) for unambiguous and ambiguous targets when target and field are of the same and of different categories.

Company) at a viewing distance of 40.6 cm.

The stimulus items were located on the display cards so as to keep the overall visual angle constant. If there were only two items, these were placed at diametrically opposed locations on the imaginary circumference. If four or six items were present, two (including the target if a target was present) were placed at opposing loci and the remaining were located randomly. Further details concerning stimulus construction can be found in Egeth et al (1972).

Procedure

Each S received 18 practice trials followed by 108 test trials presented in six blocks of 18 trials each. If S made an error, he received an extra trial immediately thereafter from which data were not collected. Subsequently, the card on which an error was made was presented again.

Before each trial, S was told the target name (e.g., "A," "two," "zero"). He was then given a verbal ready signal, after which he could initiate the trial by depressing the footswitch. Four-tenths of a second later, the stimulus card flashed on. The interval between trials was approximately 10 sec. Midway through each session, the S was allowed a short break.

The Ss were instructed to perform accurately but, within that limitation, to respond as quickly as possible. They received no feedback concerning either speed or accuracy during the experimental session.

RESULTS

Reaction Times: Same vs Different Category

While the critical comparisons concern the processing curves for the O targets, the others—A and Z for four groups, 2 and 4 for the other four—were included as unambiguous instances of the relevant categories in order to maximize the likelihood of inducing the appropriate mental set. The left-hand panel of Fig. 1 presents the results for unambiguous targets (A, Z and 2, 4). The panel displays mean reaction times as a function of n for presence and absence responders and for targets that are of the same category as field items and for those that are different. The mean slopes per character (calculated on the basis of individual slope estimates) are 25.9 and 1.3 msec for same and different categories, respectively. The slope is significantly different from zero when the categories are the same but not when they are different ($t = 7.99$, $df = 23$, $p < .001$; $t = .70$, $df = 23$, $p > .20$). An analysis of variance of the individual slope estimates yields a highly significant difference between the same and different category conditions ($F = 46.9$, $df = 1/44$, $p < .001$). These results replicate and amplify those obtained by Egeth et al (1972, Experiment 3), but with one difference: while they found flat functions for presence responders and increasing functions for absence responders, the present study yields flat functions for both. The reasons for this discrepancy are not immediately apparent.

The critical findings concern reaction times when the target was O, which are displayed in the right-hand panel of Fig. 1. The two panels seem virtually identical. In particular, there is no doubt that the effect of category is found even here: the mean slope is 24.0 msec per character when the specified category was the same and −2.4 msec when it was different. The first slope is significantly different from zero ($t = 5.87$, $df = 23$, $p < .001$), the second is not ($t = .85$, $df = 23$, $p > .20$), and the difference in slope estimates for same and different category conditions is highly significant ($F = 28.1$, $df = 1/44$, $p < .001$). This result is especially impressive considering that when the target was O, the identical stimulus cards were employed for both the same category and the different category conditions.

A final point concerns the difference between RTs in the same and different category conditions at $n = 2$. Although Fig. 1 suggests that at this value there is a crossover of the curves, this effect does not begin to reach significance ($F < 1$). On the other hand, if the curves are extrapolated to $n = 0$, the resulting intercepts are significantly greater for the different than for the same category curves ($F = 16.1$, $df = 1/44$, $p < .001$).

Reaction Times: Presence vs Absence Responders

Inspection of Fig. 1 reveals that the factor of presence vs absence response had a marked effect on the absolute level of the RT functions. It is apparent that this effect is almost entirely due to a difference in intercept values (extrapolated to $n = 0$), as indicated by a hugely significant F value of 89.6 ($df = 1/44$, $p < .001$). The difference in slopes is small enough to contribute little to the overall difference in absolute RT level; analysis of variance of individual slope estimates yields only a marginal effect of presence vs absence responders ($F = 3.8$, $df = 1/44$, $p > .05$).[2] A similar difference in intercepts has been obtained in several previous studies (Atkinson et al, 1969; Egeth et al, 1972; Nickerson, 1966). It has typically been attributed to response factors (e.g., response threshold, response incompatibility) that have little to do with the comparison stage of the recognition process. Although the differences obtained in the present study are somewhat larger than those found previously, they are probably produced by similar factors. The present study used a one-key response which may well have accentuated the response incompatibility for absence responders. This is corroborated by the remarks of several such Ss, who said that they found it difficult to restrain themselves from pressing the key when a target was present.

Errors

The overall error rate for errors of both omission and commission was 3.6%. A breakdown of commission error rates by various experimental conditions is displayed in Table 1.

It is difficult to evaluate to what extent the obtained slope differences of the RT functions should be attributed to a tradeoff between speed and accuracy. In part, this difficulty is simply due to the paucity of errors. In part, it is due to the absence of an appropriate model of the speed-accuracy tradeoff. However, the grossest assumptions of such a model might be expected to predict an interaction between n and the category factor for errors, mirroring the one obtained for reaction time. While Table 1 shows some evidence for such an interaction, this trend did not reach significance ($F = 2.6$, $df = 2/88$, $p > .10$). It is worth noting that even this insubstantial relationship is largely attributable to the aberrance of one cell (absence responders: same category, $n = 2$). Under the circumstances, there appears to be little support for a tradeoff hypothesis.

Table 1
Mean Percentage of Errors of Commission for Presence and Absence Responders by Display Size and Category Condition*

	N = 2	N = 4	N = 6
Presence Responders			
Same Category	0.5 ± .06	2.1 ± .10	2.8 ± .18
Different Category	1.5 ± .09	3.8 ± .18	5.7 ± .26
Absence Responders			
Same Category	5.2 ± .17	2.1 ± .06	4.0 ± .16
Different Category	2.6 ± .12	3.1 ± .12	4.3 ± .07

Error rates are displayed with their 95% confidence intervals. Since omission errors were small in number (14% of all errors) and since the majority of even these were contributed by only a few Ss, they are not presented here. Further, they have no immediately clear bearing on the issue of a speed-accuracy trade-off.

DISCUSSION

The results clearly demonstrate that the category effect is not an artifact of a simple physical difference between the target stimulus and the field items among which it happens to occur. It evidently does not matter whether the targets are unambiguous or ambiguous category instances *physically*; what matters is how they are specified to the S.

What mechanisms could possibly account for this phenomenon? One might propose an interpretation which assumes that the recognition system has limited processing capacity and operates in parallel as long as this capacity is not overstrained.[3] The category effect follows if classifying a stimulus as a member of a superordinate category places a lesser load upon the system's capacity limit than identifying it. In contrast, identification is assumed to tax the processing system beyond the limit up to which it can operate in parallel. As a consequence, categorization will lead to flat RT functions, while identification will generate increasing ones.

Of course, the problem now becomes one of specifying why categorization is in this sense easier than identification. At present, one can only offer speculations. Two possibilities merit consideration. One alternative hinges on the notion of different levels of perceptual processing, developed by Neisser (1967) among others. The initial assumption is that the category membership of a stimulus is defined by fewer features than is its identity. If target and field are of different categories, then each display item need only be processed to the point where its category membership can be determined. If, however, the categories for field and target are identical, feature extraction must proceed to a greater depth. Since, by hypothesis, category determination requires less feature extraction per item than does identification, it presumably will place a smaller load on the system's capacity. In consequence, several items can be processed simultaneously, yielding reaction times that are independent of n.

Another possibility is that perceptual analysis proceeds to the same depth in both categorizing and identification. In this view, the difference occurs after all features have been extracted, when the perceptually analyzed items are encoded. Posner (1970) has recently suggested that access to certain category tags does not require the intermediate link of the particular character name: both character label and category tag can be evoked directly by the percept. By analogy, the category effect in visual search may follow on the assumption that processing effort (and its resulting capacity load) increases with the number of internal encoding labels to which the stimuli are mapped; the categorizer has to select between only two categories, while the identifier must select from many characters. To determine whether these two alternatives are genuinely different requires procedures whereby depth of perceptual processing can be independently assessed.

One might argue that both of these alternatives make a prediction that is contrary to the obtained results: the absolute level of RTs for the different condition should never exceed that of the same condition. In fact, the functions violate this prediction when extrapolated to $n = 0$. To maintain either of the two alternatives just considered, one might assume, not implausibly, that categorization adds a constant increment to reaction time. This increment is added to those processing stages that are *not* involved with the comparison of the (mental) target with the display items and is thus reflected in an elevation of the intercept.

A final point: Whatever its underlying mechanism, the category effect implies a flexible processing repertory, whether at the level of feature extraction or of encoding. The present results show that the choice of processing strategy depends upon the situational context. It is not too surprising that there is some evidence for individual differences in this choice as well (Brand, 1971).

REFERENCES

ATKINSON, R. C., HOLMGREN, J. E., & JUOLA, J. F. Processing time influenced by the number of elements in a visual display. Perception & Psychophysics, 1969, 6, 321-326.

BRAND, J. Classification without identification in visual search. Quarterly Journal of Experimental Psychology, 1971, 23, 178-186.

CORCORAN, D. W. J. *Pattern recognition.* Baltimore: Penguin, 1971.

DONDERI, D. C., & CASE, B. Parallel visual processing: Constant same-different latency with two to fourteen shapes. Perception & Psychophysics, 1970, 8, 373-375.

DONDERI, D. C., & ZELNICKER, D. Parallel processing in visual same-different decisions. Perception & Psychophysics, 1969, 5, 197-200.

EGETH, H., JONIDES, J., & WALL, S. Parallel processing of multielement arrays. Cognitive Psychology, 1972, in press.

NEISSER, U. *Cognitive psychology.* New York: Appleton-Century-Crofts, 1967.

NICKERSON, R. S. Response times with a memory-dependent decision task. Journal of Experimental Psychology, 1966, 72, 761-769.

POSNER, M. I. On the relationship between letter names and superordinate categories. Quarterly Journal of Experimental

Psychology, 1970, 22, 279-287.
RABBITT, P. M. Learning to ignore irrelevant information. American Journal of Psychology, 1967, 80, 1-13.
SPERLING, G., BUDIANSKY, J., SPIVAK, J. G., & JOHNSON, M. C. Extremely rapid visual search: The maximum rate of scanning letters for the presence of a numeral. Science, 1971, 174, 307-311.
STERNBERG, S. Scanning a persisting visual image versus a memorized list. Paper presented at the annual meeting of the Eastern Psychological Association, 1967.
TOWNSEND, J. T. A note on the identifiability of parallel and serial processes. Perception & Psychophysics, 1971, 10, 161-163.
YONAS, A. The acquisition of information-processing strategies in a time-dependent task. Unpublished doctoral dissertation, Cornell University, 1969.

NOTES

1. Increasing functions of RT with n do not necessarily require an interpretation in terms of a serial process. A parallel process, for instance, might be plausible if one assumes a system of limited capacity in which less processing "energy" is devoted to the examination of each item as the number of items increases; this would result in a corresponding increase in RT (Corcoran, 1971; Townsend, 1971).

2. It is worth noting that while this slope difference is minimal for the different category condition (—1.2 vs 1.5 msec per character for presence vs absence responders), it is substantial for the same category condition as indicated by slopes per character of 19.8 and 30.8 msec for presence and absence responders, respectively. Several authors (e.g., Sternberg, 1967) have discussed the implication of slope differences between RT curves for presence and absence responses. If a presence RT function has a slope one-half that of its corresponding absence function, this is considered evidence for a serial search mechanism terminating when the target is found. Conversely, if both functions are found to have identical slopes, an exhaustive serial search is implicated for both presence and absence responders. Evidence in support of both models has been presented for visual search tasks (Atkinson et al, 1969; Nickerson, 1966; Sternberg, 1967). Although the present slope differences fail to support either model unequivocally, the issue of the self-termination or exhaustiveness of the search for the same category condition is not of critical relevance to the present discussion.

3. One important feature of such an interpretation concerns processing characteristics when the capacity limit of the system is not exceeded. It is assumed that there is an optimal amount of "processing energy" for the analysis of any stimulus element. Once this optimum is reached, further increments in available processing energy will not decrease the time required to examine an element. This assumption is necessary to justify flat functions of RT with n within the capacity limit of the system.

(Received for publication July 3, 1972; revision received August 14, 1972.)