

Unconscious perception of meaning: A failure to replicate

KAREN A. NOLAN and ALFONSO CARAMAZZA
Johns Hopkins University, Baltimore, Maryland 21218

An unsuccessful attempt was made to substantiate Marcel's (in press) claim that semantic information can be extracted from visually presented words and can affect responses to subsequent stimuli under conditions that prevented identification of the stimulus word and even awareness of its presence. The results reported here are consistent with traditional information processing models of reading that hold that retrieval of a semantic representation for a visually presented word requires prior computation of a graphemic code for the word and that semantic information cannot become available in the absence of corresponding visual/graphemic information.

Two recent studies claim to demonstrate that semantic information may be extracted from a visually presented word under conditions of pattern masking that prevent identification of, and even awareness of the presence of, the masked word (Allport, 1977; Marcel, in press; Marcel & Patterson, 1978). That such effects are reminiscent of the now notorious reports of subliminal perception that emerged in the late 1950s and early 1960s (see Dixon, 1971, for a review) might alone warrant closer examination of the procedures by which they were obtained. In addition, substantiation of these reports of "unconscious reading" may necessitate the reevaluation of some of the most basic assumptions of the information processing approach to cognitive psychology.

Traditionally, it has been assumed that complex cognitive processes, such as reading, are dependent on a series of discrete, sequential stages such that information obtained in one stage provides necessary input for the next stage of processing (Sternberg, 1969). For example, identification and comprehension of a written word is generally assumed to involve computation of some kind of graphemic code for the visual stimulus (Baron, 1975; Forster, 1978). Although some models permit components of the information processing system to operate continuously, and hence do not require complete identification of all the component letters prior to lexical access (McClelland, 1979), such models assume that a complete graphemic representation becomes available when a semantic representation for a word is accessed. Since pattern masking involves dichoptic presentation of a mask that is similar in component features to the target,

it is assumed to be central masking (Turvey, 1973) and, therefore, to terminate perceptual processing of the target, only the information that has been extracted prior to mask presentation should be available for report. If Allport's (1977) and Marcel's (in press) results are accepted, then, we must revise either our views of how words are identified and comprehended or our notions regarding the effects of central masking. If information processing does indeed proceed in sequential stages, then perhaps masking does not terminate processing but only renders the outcome of early processing stages (e.g., visual analysis) inaccessible while sparing the records of later stages.

In one of these studies, conducted by Allport (1977), arrays of two to four words were displayed tachistoscopically and subjects were required to report as many words as possible from each array. Of the words incorrectly reported, 6%-9% were semantically related to the presented word. The occurrence of these semantic errors has been interpreted to indicate that semantic information can be extracted from a visually presented word and can affect conscious reports in the absence of sufficient visual information for correct identification. Allport has used these data to support a model of reading that holds that the mechanism that is disrupted by pattern masking normally functions to suppress semantic errors in reading and that disruption of this mechanism is the cause of deep dyslexia, a syndrome characterized by the production of semantic errors in oral reading. However, two other studies have been reported that indicate that the 6%-9% semantic error rate reported by Allport may have arisen by chance (Ellis & Marshall, 1978) or may be accounted for by a guessing strategy (Williams & Parkin, 1980) based on overall knowledge of the target population (i.e., concrete nouns). In a third study (Fowler, Wolford, Slade, & Tassinari, 1981), subjects were required to guess the identity of pattern-masked stimuli and the target-error pairs obtained from each subject were later presented to judges who were

The research reported in this paper was supported by a Biomedical Research Support Grant 5 S07-RR-07041-14 from the Johns Hopkins University to the second author and by National Institute of Health Grant NS 14099 to the Johns Hopkins University. Inquiries and requests for reprints should be addressed to Karen A. Nolan, Department of Psychology, University of Pennsylvania, Philadelphia, Pennsylvania 19104.

asked to rate the semantic relatedness of the pairs from each protocol. Although a stimulus was presented on only half of the trials, judges were unable to distinguish target-error word pairs in which the target had actually been presented to the subject from those in which no target had been presented. The results of these studies suggest, therefore, that the rate of production of semantic errors in the identification of pattern-masked words reported by Allport is not a reliable effect. The original study from which Allport's arose was conducted by Marcel (in press) but has not been generally available until recently. Even if Allport's results are dismissed as spurious, Marcel's study deserves consideration in its own right, especially since it could provide the basis for even stronger claims regarding the automaticity of retrieval of semantic representations for visually presented words.

On each trial in Marcel's (in press) experiment, either a word or a blank field was displayed and followed by a pattern mask. For each subject, the longest stimulus-onset asynchrony (SOA) at which the subject began to make errors in judging whether a word preceded the pattern mask was determined and was used as the starting SOA. For the experimental trials, six different SOAs were used, ranging from the individually determined starting SOA to 20 msec below that SOA. Marcel found that, as SOA was decreased, presence/absence judgments were the first to reach chance level (defined by Marcel as 60% correct). Furthermore, at SOAs at which subjects' judgments regarding which of two words was more similar to the presented word in terms of graphic similarity first fell to chance, most (14/17) subjects were able to choose which of two words was more similar to the presented word in terms of semantic characteristics with significantly greater accuracy. Thus, the order in which different kinds of information about the presented words became unavailable (first simply presence-absence, then graphic characteristics, and finally, meaning) is precisely the order in which these types of information are generally presumed to become available and the reverse of the order in which they are predicted to become unavailable by information processing models. Therefore, these results, if accepted, require drastic revision of currently held views of the reading process.

However, there are several flaws in Marcel's (in press) study, not the least of which is his arbitrary definition of chance performance as 60% correct. Since a word was presented on half of the trials for which detection decisions were required and since judgments regarding graphic and semantic similarity were obtained in the form of a two-alternative forced-choice task, chance performance in this study should have been defined as 50% correct for each of the three judgments. This issue is important because Marcel's results are controversial only if they provide evidence that semantic information may be available without comparable graphic informa-

tion. If they indicate that semantic information can be obtained (or guessed at) from partial graphic information, then they are consistent with at least some of the current information processing models of reading.

Another serious flaw in Marcel's (in press) study concerns the repeated presentation of the stimulus words. A set of 240 words was used (for each subject) for a total of 720 experimental trials. Half of these words were used for detection trials; the entire set was used for 240 trials each of semantic and graphic similarity judgments. Thus, half of the words in this set must have been presented to each subject (at least) three times and half of the words presented twice. Other studies have shown that a stimulus is responded to more quickly when it is repeated, even when several stimuli intervene between the presentations of the repeated stimulus (Hyman, 1953) and that this effect is unaltered by increases in time intervals or the insertion of an extraneous task between successive stimuli (Keele, 1969). Another of Marcel's (in press) own studies showed that, while sequential repetition had no effect on the probability of detection, the effect of semantic association on lexical decision reaction times was enhanced by stimulus repetition. Repetition might have affected graphic and semantic similarity judgments even more directly; if target words were presented more than once, the alternatives for the similarity judgments for those words might also have been presented more than once. If the first occurrence of a pair of alternatives was one of the earlier trials in the experiment, SOA would be relatively long and the subject would be likely to have sufficient information to select the correct alternative. Thus, the subject might be biased to choose the same, incidentally correct alternative on the second presentation of the pair, even though he or she had no information regarding the stimulus word. While the discrepancy between semantic and graphic similarity judgments at the shortest SOAs would only be expected to arise if the rate of repetition differed for these alternatives, repetition of alternatives could easily account for the superiority of similarity over presence-absence judgments.

In addition, since subjects were required to make only one of the three judgments on any given trial, it is possible that the superior performance on semantic similarity judgments was due to subjects' actually seeing the target word on more of those trials than on the other types of trials. It is also possible that an unintended bias was introduced in the selection of alternatives for the graphic and semantic similarity judgments. Fowler et al. (1981), who succeeded in replicating Marcel's (in press) results in two separate experiments, conducted a third experiment in which they presented their alternatives to subjects who were asked to select which of the two words they thought had more words "like it" in terms of the appropriate dimension of similarity. Fowler et al. found that subjects chose the "correct" alternative

more often for semantic pairs than for graphic pairs, and they concluded that the similarity between the results of this control experiment and those of the masking experiments was sufficient to warrant caution in attributing the results of the masking experiment to information extracted from the stimulus word.

In view of these and other potential problems in Marcel's (in press) procedure (e.g., the use of precuing), and since this, as yet unpublished, experiment has been so widely cited, we have attempted to replicate his findings while employing what we believe to be stricter controls. In order to eliminate any potential effects of repetition, no stimulus word, and no pair of alternatives, was presented to a given subject more than once. Furthermore, in order to diminish the strangeness of requiring subjects to make similarity judgments for stimuli of which they had no subjective awareness, and in order that we might examine responses to similarity alternatives on trials on which no stimulus was presented for indications of response bias, subjects were required to make both a detection and a similarity judgment on every trial.

METHOD

Subjects

Twenty-four subjects, students at the Johns Hopkins University, were pseudorandomly assigned to two groups, visual or semantic, so that, in addition to a detection judgment, each subject was required to make only one type of similarity judgment on each trial. This eliminated the need for precuing before each trial, and similarity judgments and detection performance could be compared within trials.

Stimuli

Correct alternatives and distractors for the similarity judgments were chosen from separate ratings of visual and semantic similarity completed by two different groups of 15 subjects each. Words for which one alternative received a high similarity rating (mean ≥ 5 on a 7-point scale) on each of the ratings and for which another alternative on each of the rating scales received low similarity ratings (mean ≤ 3) were retained for use as stimuli. The same set of words was used as stimuli for both visual and semantic groups, but each group received different alternative pairs.

Procedure

There were 12 blocks of 16 trials each; the first 4 blocks were considered practice blocks, with SOAs of 100.2, 83.5, 66.8, and 50.1 msec, respectively. On half of the trials within each block, a word was presented, and on half of the trials, a blank field.

On each trial, either a single word, printed in uppercase letters, or a blank field was displayed on the viewing screen of a PET computer, followed by a pattern mask at the SOA for that block of trials. The screen was then cleared and the question "WAS A WORD PRESENTED?" was printed in the center of the screen, with the alternatives "YES" and "NO" printed below it, one to the right and one to the left of center. (For all judgments, placement of the alternatives was pseudorandomly counterbalanced so that the correct alternative appeared an equal number of times in each position, with the restriction that the correct alternative could not appear in the same position on more than three consecutive trials.) Subjects responded by pressing the button (with the dominant hand) that corresponded

in position to the chosen alternative. As soon as the subject responded to the detection question, the screen was again cleared and the second question "WHICH IS MORE SIMILAR VISUALLY [or IN MEANING]?" and its two alternatives were similarly displayed and remained on the screen until the subject responded.

RESULTS AND DISCUSSION

Inherent in an experiment such as this is the assumption that subjects may have more information about the stimulus word than can be inferred from their detection responses. That is, when a subject responds "NO" to the detection question, we cannot presume that the visual system registered nothing at all prior to the mask, but only that the information available to the subject is insufficient to induce a willingness to make a positive response. It is therefore reasonable to analyze detection responses in terms of the theory of signal detection (Green & Swets, 1966; Swets, Tanner, & Birdsall, 1961), which permits calculation of two measures, d' and β . The first is a measure of the observer's sensitivity or ability to discriminate between stimulus and no-stimulus conditions, whereas β represents the observer's criterion for responding. For trials at the shortest SOA, both measures were calculated for 19 of the 24 subjects. For the other five subjects, d' could not be calculated because they were, in essence, able to discriminate no-stimulus trials perfectly, making no "false alarm" errors. Detection performance was then compared with accuracy of similarity judgments for the remaining subjects, using chi square. The two subjects who performed significantly better on the similarity judgments ($\chi^2 = 6.3$ and 5.0 ; $p < .02$ and $p < .05$) were both in the semantic group. However, both of these subjects were above chance on the detection task (63% and 55% correct). Thus, the within-subjects analysis lends no support to Marcel's (in press) claims.

When the data were averaged over subjects (including only those 9 of the 19 subjects for whom d' could be calculated and who also met Marcel's (in press) criterion of 60% or less accuracy on the detection task at the shortest SOA), we found that similarity judgments were slightly, but not significantly, better for subjects in the visual group than for those in the semantic group (60% correct vs. 58% correct). Furthermore, the difference in accuracy on similarity judgments was in the same direction as the difference in detection accuracy for the two groups (57% correct vs. 55% correct). Nor was either group significantly better at making similarity judgments than at detecting stimuli. Finally, mean accuracy of similarity judgments made on no-stimulus trials (50% correct) did not differ from chance, indicating that no bias was present.

While our failure to replicate the results of Marcel's (in press) experiment may be suggestive of flaws in the original design and/or analysis, the discrepancy between

the results of the two experiments might simply be due to a Type I error, rejection of a true null hypothesis, in Marcel's experiment. In any case, the results reported here further undermine the plausibility of the claim that either graphic or semantic information may be available in the absence of awareness of the presentation of a stimulus word. Furthermore, we have been unable to obtain any evidence to support the notion that semantic information may be available in the absence of corresponding graphic information.

Traditional information processing models of reading assume that both extraction and encoding of information from the stimulus proceed in a series of discrete stages from the level of visual analysis through graphemic, and possibly phonological (Rubenstein, Lewis, & Rubenstein, 1971; Spoehr & Smith, 1975), analysis to the point at which the stimulus can be identified with a stored lexical entry from which a semantic representation can be retrieved. Since encoding of information at a given level of analysis is initiated prior to encoding of information at subsequent levels of analysis, termination of processing by pattern masking should result in lack of information that can be obtained only in later stages of processing. Even models that assume that the components of an information processing system operate continuously, such as McClelland's (1979) cascade model, would predict that, if processing is interrupted, more information should be available to lower level components of the system (e.g., visual analysis) than to higher levels of analysis (e.g., semantic analysis), although higher level analyses may be initiated prior to completion of lower level analyses. The results reported here are entirely consistent with both the discrete stage models and the cascade model. Replication of Marcel's (in press) results, on the other hand, would have necessitated drastic revision of commonly held views of the reading process and/or of the effects of central masking.

On the basis of Allport's (1977) and Marcel's (in press) reports, some theorists (Newcombe & Marshall, 1980) have assumed that encoded visual information normally serves as a record that is used to constrain reading responses to the presented word and, thereby, to suppress the production of semantic paralexias. Their interpretation of deep dyslexia was, therefore, that the brain damage that gave rise to this syndrome prevented use of this visual record, either by eliminating the record itself or by disrupting the operation of the checking mechanism that employed this record. In view of the results reported here, postulation of such an independent checking mechanism is unjustified, and reconsideration of information processing models in general is, at best, premature.

REFERENCES

- ALLPORT, D. A. On knowing the meaning of words we are unable to report: The effects of visual masking. In S. Dornic (Ed.), *Attention and performance VI*. Hillsdale, N.J.: Erlbaum, 1977.
- BARON, J. Successive stages in word recognition. In P. M. A. Rabbitt & S. Dornic (Eds.), *Attention and performance V*. New York: Academic Press, 1975.
- DIXON, N. F. *Subliminal perception: The nature of a controversy*. New York: McGraw-Hill, 1971.
- ELLIS, A. W., & MARSHALL, J. C. Semantic errors or statistical flukes? A note on Allport's "On knowing the meaning of words we are unable to report." *Quarterly Journal of Experimental Psychology*, 1978, **30**, 569-575.
- FORSTER, K. I. Accessing the internal lexicon. In E. W. Walker (Ed.), *Explorations in the biology of language*. Montgomery, Vt: Bradford Books, 1978.
- FOWLER, C. A., WOLFORD, G., SLADE, R., & TASSINARY, L. Lexical access with and without awareness. *Journal of Experimental Psychology: General*, 1981, **110**, 341-362.
- GREEN, D. M., & SWETS, J. A. *Signal detection theory and psychophysics*. New York: Wiley, 1966.
- HYMAN, R. Stimulus information as a determinant of reaction time. *Journal of Experimental Psychology*, 1953, **45**, 188-196.
- KEELE, S. W. Repetition effect: A memory dependent process. *Journal of Experimental Psychology*, 1969, **80**, 243-248.
- MARCEL, A. J. Conscious and unconscious perception: Visual masking, word recognition, and an approach to consciousness. *Cognitive Psychology*, in press.
- MARCEL, A. J., & PATTERSON, K. E. Word recognition and production: Reciprocity of clinical and normal studies. In J. Requin (Ed.), *Attention and performance VII*. Hillsdale, N.J.: Erlbaum, 1978.
- MCCLELLAND, J. L. On the time relations of mental processes: An examination of systems of processes in cascade. *Psychological Review*, 1979, **86**, 287-330.
- NEWCOMBE, F., & MARSHALL, J. C. Transcoding and lexical stabilization in deep dyslexia. In M. Coltheart, K. Patterson, & J. C. Marshall (Eds.), *Deep dyslexia*. London: Routledge & Kegan Paul, 1980.
- RUBENSTEIN, H., LEWIS, S. S., & RUBENSTEIN, M. A. Evidence for phonemic recoding in visual word recognition. *Journal of Verbal Learning and Verbal Behavior*, 1971, **10**, 645-657.
- SPOEHR, K. T., & SMITH, E. E. The role of orthographic and phonotactic rules in perceiving letter patterns. *Journal of Experimental Psychology: Human Perception and Performance*, 1975, **1**, 21-34.
- STERNBERG, S. The discovery of processing stages: Extensions of Donder's method. In W. G. Koster (Ed.), *Attention and performance II*. Amsterdam: North-Holland, 1969. (Reprinted from *Acta Psychologica*, 1969, **30**, 276-315.)
- SWETS, J. A., TANNER, W. P., & BIRDSALL, T. G. Decision processes in perception. *Psychological Review*, 1961, **68**, 301-340.
- TURVEY, M. T. On peripheral and central processes in vision: Inferences from an information-processing analysis of masking with patterned stimuli. *Psychological Review*, 1973, **80**, 1-52.
- WILLIAMS, P. C., & PARKIN, A. J. On knowing the meaning of words we are unable to report—Confirmation of a guessing explanation. *Quarterly Journal of Experimental Psychology*, 1980, **32**, 101-107.

(Received for publication June 4, 1982.)