Perception & Psychophysics
1973, Vol. 14, No. 2, 293-297

Metacontrast and brightness discrimination*

IRA H. BERNSTEIN†, JANET D. PROCTOR, ROBERT W. PROCTOR
University of Texas, Arlington, Texas 76010

and

DONALD L. SCHURMAN
Emory University, Atlanta, Georgia 30322

An attempt was made to obtain U-shaped masking functions in two metacontrast experiments. Trained Ss judged whether a square test stimulus (TS) was bright or dim. The TS was presented alone or in conjunction with an adjacent pair of square masking stimuli (MS) whose energy equaled the bright TS. The stimulus onset asynchronies (SOA) ranged from 0 to 125 msec. The task minimized the role of apparent movement cues as a reliable basis for judgment. Similar studies have employed TS plus MS vs MS alone as the alternatives, allowing apparent movement to be a cue. Brightness accuracy was a U-shaped function of SOA. This finding is consistent with neural-net models (Weisstein, 1968). However, analysis of Ss' response bias suggested an alternative explanation involving the MS as a comparison stimulus at short SOA. It was concluded that U-shaped masking functions are also consistent with theories based upon independent component processes, e.g., Schurman and Eriksen (1970) and Uttal (1970).

Metacontrast refers to a performance decrement or a decrease in phenomenonal brightness of a test stimulus (TS) produced by an adjacent but nonoverlapping masking stimulus (MS) presented concurrently or shortly after TS. The term "masking function" describes the relationship between the response measure and the stimulus onset asynchrony (SOA) separating TS onset from MS onset. These masking functions typically are at a minimum at 0-msec SOA (concurrent TS-MS onset) and increase monotonically to the control level (TS alone) between 100- and 200-msec SOA. However, U-shaped functions with minima at 25- to 50-msec SOA have also been reported. The conditions giving rise to this phenomenon are quite controversial. Equal TS and MS energies seem a necessary (Weisstein, 1968) but not sufficient condition (Schurman, 1972; Schiller & Smith, 1966) for U-shaped masking.

Schurman (1972) has described two groups of investigators who have been in conflict over U-shaped masking. One group is represented by Weisstein (1968) and Haber (1970), among others. They tend to define perception in phenomenological terms and to use brightness matching and magnitude estimation measures in their most recent metacontrast studies. Conceptually, they view metacontrast as a unitary phenomenon. Weisstein's (1968) elegant neural-net model, based upon principles of lateral inhibition, typifies this approach. She attempts to demonstrate how differences between U-shaped and monotonic masking functions are the outcome of changes in latencies of inhibitory and excitatory neurons in a five-element neural net. These latencies, in turn, are a function of stimulus energy.

Eriksen (1966; Eriksen, Becker, & Hoffman, 1970; Eriksen & Collins, 1965; Eriksen, Collins, & Greenspon, 1967; Eriksen & Marshall, 1969; Schurman & Eriksen, 1970) and Uttal (1970, 1971) reject, with varying emphases, the above assertions. They tend to eschew neural models and are positivists in the sense of Garner, Hake, and Eriksen (1956) in their definition of perception. Consequently, they typically study recognition accuracy. They view metacontrast, in general, and U-shaped masking functions, in particular, as a composite of temporal integration, apparent movement, and other effects in addition to lateral inhibition, i.e., they are *independent component* theorists.

Schurman (1972) pointed out that this dispute has led both parties to cite the same fundamental set of data (Schiller & Smith, 1966) in support of their diverse positions. The latter employed both a brightness matching task, in which they obtained U-shaped masking, and a detection task with TS + MS and MS alone as the stimulus alternatives, in which they obtained monotonic masking.

The present paper is most directly concerned with masking functions obtained from brightness accuracy data and, indirectly, with relations among brightness accuracy, phenomenal brightness, and recognition accuracy measures. Kahneman (1968) noted a difficulty with Schiller and Smith's (1966) efforts in this direction. He pointed out that an "apparent movement"[1] cue is present in their detection paradigm for TS + MS presented around 50-msec SOA that is absent with MS alone. Even though TS is essentially suppressed at this SOA, the movement cue provides for an easy discrimination between the two alternatives. There is

---

*Supported by funds provided by the Liberal Arts Organized Research Funds, the Graduate School of the University of Texas at Arlington, and the Genco Abbandando Foundation. A portion of these data were presented at the 1973 Convention of the Midwestern Psychological Association, Chicago, Illinois.

†Requests for reprints should be addressed to Ira H. Bernstein, Department of Psychology, University of Texas, Arlington, Texas 76010.

little apparent movement at shorter SOA. Hence, discrimination may be poor through loss of this movement information, even though the TS is not suppressed. Thus, in Kahneman's (1968) terms, the "criterion content" (brightness for very short or long SOA, movement for intermediate SOA) and the SOA manipulation proper are confounded in the above detection paradigm.

However, there is clear need for a brightness accuracy paradigm that provides data comparable to phenomenological data. The differences cited above among investigators active in the area of metacontrast has left several alternatives open to resolve the conflict over U-shaped masking. One is that phenomenal methods confound sensory and response bias effects in the signal detection sense (Green & Swets, 1966). If this is so, U-shaped masking could be a response bias artifact. Alternatively, different sensory processes might be involved in phenomenal brightness and brightness accuracy. Masking functions for the former could "legitimately" be U-shaped, i.e., not a response bias artifact, yet the latter masking functions could be monotonic. Another possibility is that both types of brightness measures are mediated by a process capable of producing U-shaped functions and recognition by a different process. Of course, phenomenal brightness, brightness accuracy, and recognition accuracy may well share common mechanisms which, under suitable circumstances, can provide U-shaped functions. Spencer and Shuntich (1970) have provided evidence for U-shaped masking in recognition using a nonmetacontrast paradigm.

We were able to obtain U-shaped brightness during pilot research. Following the Garner et al (1956) tradition, we did not accept this as evidence for U-shaped masking. Instead, we attempted to devise a task that was more suitable than Schiller and Smith's (1966) detection paradigm yet would minimize response-criterion effects upon performance measures. The task was to discriminate between two TS differing in luminance. The apparent movement effects for both TS were very similar. Hence, apparent movement could not be a reliable basis for judgment. Two studies were run. The SOA was blocked over trials in the first study and varied randomly over trials in the second. In both studies, the TS differed in luminance by .2 log units. Because the studies are quite similar otherwise, they will be reported concurrently rather than successively for ease of exposition.

As we shall note below, the purpose of this study was *not* to test alternative theories of metacontrast. Rather, it was to see if stable U-shaped curves could be obtained for brightness accuracy once apparent movement cues were controlled.

METHOD

Subjects

Three advanced undergraduates served in both experiments. A fourth S participated in the first experiment but not in the second and was replaced by a fifth S, so that a total of four Ss participated in each experiment. All Ss were paid for their participation. The Ss were given extensive practice at the task and had normal or corrected to normal eyesight. They were theoretically unsophisticated.

Apparatus and Stimuli

Stimuli were presented by transillumination on a Scientific Prototype tachistoscope (Model GB). Sylvania bulbs (F8TS/D) were used. Viewing was monocular through a 2-mm artificial pupil used to eliminate binocular parallax that might have arisen from mirror misalignment. A McBeth illuminometer was used to verify luminances, and a Tektronix oscilloscope (Model 564) was used to verify timing accuracy. A telegraph key was used for the S to self-initiate trials upon command from E that the necessary adjustments had been made. The E and S sat in separate rooms.

A small circle that remained on continually was used for fixation. It was contained in the blank field of the tachistoscope and was illuminated by a separate tungsten source and light pipe, which did not affect the other contents of the field. The blank field could, thus, be used as a separate exposure field. The fixation circle was sufficiently dim so as to be barely visible to a light-adapted S. It became clear with dark adaptation and, for this reason, served as a control over adaptation effects.

Exposure Fields I and II contained the two stimulus alternatives, the dim TS and the bright TS. The location of each TS with respect to the two fields was alternated midway through each experiment to eliminate coloration and other cues specific to a given field from serving as a cue. Each square appeared .33 deg of an arc above fixation. The MS was presented in the blank field and consisted of a pair of squares whose sides were adjacent to the TS. The length of each side of TS and MS was .67 deg of visual angle.

The TS and MS durations were both 50 msec. A single timer controlled both TS. The E selected the TS alternative by means of a switch that enabled one of the sets of bulbs to fire. The luminance of the bright TS and the MS was 63 cd/m^2 in both experiments. The dim TS was 40 cd/m^2 in both experiments. In both cases, the luminance of the dim TS was controlled by a .2 neutral density filter (Kodak), as noted above. The dim TS luminance was chosen on the basis of pilot data, which suggested that accuracy of intensity discrimination would be about 90%-95% for TS alone.

Procedure

The Ss were run for several practice sessions until their performance had apparently stabilized. Each S was then run in five 1-h sessions in both experiments. A total of 168 trials were run per session. The two TS occurred equally often. In the first experiment, 24 trials per session were run at each of seven stimulus conditions: no MS and SOA of 0, 25, 50, 75, 100, and 125 msec. Blocks of each stimulus condition were randomized between sessions and over Ss. A like distribution of trials was used within each session for the second experiment, except that stimulus condition was varied on a trial by trial basis. Periodic rest breaks were given in both experiments.

The E began each trial with appropriate dial settings and switch adjustments, including dummy settings to avoid giving information to S. The E then signaled S to initiate the trial when the fixation circle appeared as bright and clear as possible. The S's start key initiated presentation of the stimulus. The S then responded with a 6-point confidence rating (CR). A "1" on this scale denoted that the S was very sure that the bright TS had been presented and a "6" denoted similar confidence that the dim TS had been presented.

Standard CR techniques were used to generate receiver operating characteristic (ROC) curves at each data point (stimulus condition × S) by treating the bright TS as "signal plus noise" and the dim TS as "noise." Various sensitivity indices were computed. These generally agreed closely. The data to be reported are in terms of the area under the ROC curve (see Pollack & Hseih, 1969 for a discussion of the properties of this measure), to be consistent with our earlier work (Bernstein, Amundson, & Schurman, in press). Various bias measures were obtained by disregarding the confidence ratings, i.e., using the criterion separating the "3" and "4" categories. The data reported are in terms of $-\ln$ of Luce's (1963) b measure, defined as:

$$b = \sqrt{\frac{\text{miss rate} \times \text{correct rejection rate}}{\text{hit rate} \times \text{false alarm rate}}}$$

Positive values of $-\ln(b)$ denote a tendency to report stimuli as bright, regardless of accuracy. Other bias measures gave similar results to those reported.

RESULTS

The main results are reported in Fig. 1 in terms of the area under the CR-inferred ROC curves at each SOA for each of the experiments. These data are clearly U-shaped and describe the performance of three of four Ss in the first experiment and all four Ss in the second experiment. The deviant S showed constant performance (±.01) at SOA of 0, 25, and 50 msec. His masking function was slightly U-shaped when analyzed in terms of Luce's (1963) sensitivity measure, η. Individual declines in performance between 0- and 25-msec SOA ranged from .11 to .35 across the two experiments. The difference in minima for the two experiments, however, is probably not meaningful. Two of the three Ss who manifested a U-shaped function in the first experiment reached minimum performance at 25-msec SOA, but the third S reached minimum performance at 50-msec SOA. Conversely, two of the Ss in the second experiment reached minimum performance at 25-msec SOA and two at 50-msec SOA. The lower performance in the second experiment was consistent.

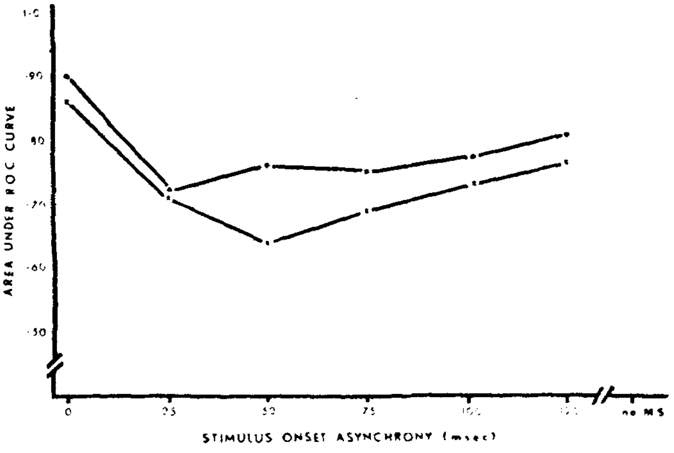

Fig. 1. Area under the CR-inferred ROC curves obtained from the first experiment (filled circles) and the second experiment (crosses) as a function of SOA.

Statistical confirmation of the U-shaped function was accomplished in the following manner. First, the area measures were subjected to a stimulus conditions by Ss analysis of variance (ANOVA), separately for each experiment. The main effect of stimulus conditions was highly significant [$F(6,18) = 5.06$ and 12.61 for the first and second experiments, respectively, $p < .01$]. The stimulus condition effect was then partitioned into two orthogonal components: the difference between the no-MS condition and the 6-SOA, and variation among the 6-SOA. The latter effect was then subjected to a trend analysis. The percentage of main effect variation accounted for these components and the associated significance levels are reported in Table 1. Also reported are the error mean squares and the stimulus condition effect as a percentage of total variance ($SS_{\text{stimuli}}/SS_{\text{tot}}$).

As can be seen from Table 1, there is a highly reliable quadratic component in both experiments that is the largest source of variation in the SOA effect. The average of all TS + MS conditions is, as one would expect, below the no-MS control. However, performance at 0-msec SOA is as good as it is without the MS. The cubic trends

Table 1
Percentage of Main Effect Variance of Area and $-\ln(b)$ Measures Accounted for by the Mask/No-Mask Difference, Variance Among SOA, Orthogonal Polynomial Components, Error Mean Square, and Ratio of Main Effect to Total Sum of Squares

		Measure			
		Area		$-\ln(b)$	
Effect	df	Experiment I	Experiment II	Experiment I	Experiment II
Mask/No Mask	1	42**	27**	73	13**
SOA	5	58**	73**	27	87**
Linear	1	2	6**	4	12**
Quadratic	1	38**	52**	18	15**
Cubic	1	10**	10**	0	47**
Residual	2	8	5	5	13**
MS Error	18	46a	21a	34b	8b
$SS_{\text{Stimuli}}/SS_{\text{Tot}}$		52**	36**	15	72**

a × 10000 *b × 100* **p < .05* ***p < .01*

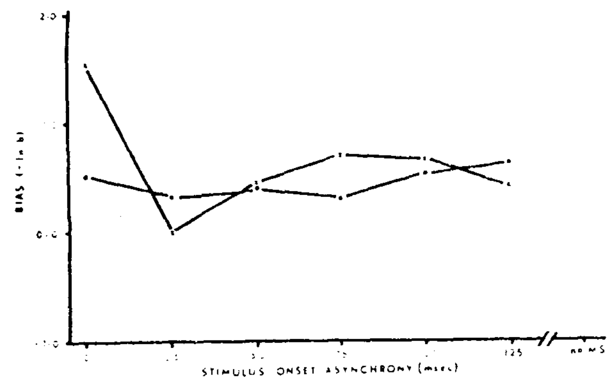

Fig. 2. Bias in terms of −ln(b) obtained from the first experiment (filled circles) and the second experiment (crosses) as a function of SOA.

in both experiments, though significant, do not contribute much variance. The same holds for the linear component in the second experiment. Presence of these odd polynomial effects is in part due to not having reached asymptote at the longest SOA.

Figure 2 contains the criterion, −ln(b) measures, similarly plotted for both experiments as a function of SOA. Table 1 contains the results of the ANOVAs conducted upon these measures. As can be seen, −ln(b) remained constant in the first experiment [$F(6,18) < 1.0$]. This reflects a tendency for Ss to vary their criteria to keep the percentage of "bright" ("1," "2," and "3") responses constant at each SOA despite the variation in average brightness. Thus, the data represent "local" biases established at each SOA.

In contrast, −ln(b) varied significantly but complexly in the second experiment [$F(6,18) = 13.23$, $p < .01$]. This seems to arise because the random variation of SOA prevents Ss from adjusting their criterion to each SOA. Though the cubic component is strongest in the analysis, the main difference is between 0 msec and all other points. This accounts for 73% of the stimulus condition effect [$F(1,18) = 58.36$, $p < .01$]. A second orthogonal effect found in this post hoc analysis was between 25-msec SOA and all other points, excluding 0-msec SOA. This accounts for 13% of the variance [$F(1,18) = 9.61$, $p < .01$]. Finally, the bias for TS alone is below that for 50-125 msec, accounting for 11% of the variance [$F(1,18) = 8.41$, $p < .01$]. The overall shape of this function suggests that the decline in accuracy at 25-msec SOA arises from an increase in miss rate (decrease in hit rate) rather than an increase in false alarm rate (decrease in correct rejection rate).

DISCUSSION

Our results indicate that U-shaped masking occurs for brightness accuracy when apparent movement cues are minimized. This occurred in two experiments that differed procedurally and that produced different response strategies, as seen in the bias data. Hence, the present results and those for phenomenal brightness seem similar.

The present data are consistent with neural-net models (Weisstein, 1968). Her model is deterministic, as it defines a single brightness output to the stimulus energies impinging upon the net rather than a probability distribution of brightnesses. These probability distributions, or some equivalent, are necessary to produce ROC curves for a quantitative comparison of her model and our results. Her model qualitatively accounts for the drop in performance with random SOA through confusions between outputs for a bright TS at one SOA and a dim TS at another. It is not clear how the model would be quantitatively adjusted for SOA variation nor how it would account for the response bias data.

However, it is not necessary to suppose such a model to account for the data. A U-shaped masking function for brightness accuracy is consistent with independent component theories proposed by Schurman and Eriksen (1970) and Uttal (1970). In such models "lateral inhibition" is used only in a descriptive sense (the MS is lateral to the TS and inhibits it), not in a neural sense.

To consider how the latter theories would handle the data, it may be noted that, in a typical metacontrast study employing equal TS-MS energies, the S has a fundamentally different set of cues at short SOA as opposed to longer SOA or no MS. The cues in the former case are furnished by the MS as a comparative or reference stimulus. At short SOA, the S can respond "Yes, the bright TS occurred" when TS and MS are equally bright and "No, the dim TS occurred" when they are not. At longer SOA, providing the asynchrony is detected, Ss should disregard TS and use an absolute judgment strategy to avoid comparison of a fading TS trace with an intact MS trace. Moreover, accuracy of comparative judgments is well known to be better than accuracy of absolute judgments. Thus, both the beneficial effects of comparative judgments and the interfering effects of metacontrast could diminish monotonically with SOA. Providing the effects are not of the same magnitude and do not occur at the same rate, U-shaped masking would result.[2]

Schurman (1972) has made a related point. He has noted the minimum point on U-shaped masking functions and the just noticeable difference for temporal asynchrony judgments (Schmidt & Kristofferson, 1963) typically are both around 25-50 msec. Between 25 and 50 msec, absolute and comparative judgment strategies are likely to overlap. The S would fail to discriminate the asynchrony and employ a comparative judgment. Because the TS trace would have faded, both TS would appear dim, accounting for the bias shift toward a higher miss rate at 25 msec in the second experiment. The bias data from the first experiment are not pertinent, for reasons previously noted.

We do not claim that the above interpretation accounts for all instances of U-shaped masking. It does not seem pertinent to the pattern mask data of Spencer and Shuntich (1970). However, it can be easily modified to account for some pattern recognition data as well as brightness data. In particular, the ring MS in letter recognition judgments can allow comparative judgments of the O vs D TS used by Weisstein and Haber (1965), who reported a U-shaped masking function for pattern recognition. In their task, the contours of the MS and the O are of similar shape. The dissimilarity of the left half of the D could contrast effectively to boost performance above that attributable to metacontrast alone.

On the other hand, Eriksen and Collins (1965) used A, T, and U as the TS. Although the bottom of the latter could be compared with the ring TS, only one alternative would be eliminated with a mismatch. In other words, a comparative judgment strategy would be less informative. Similarly, Eriksen, Becker, and Hoffmen (1970) used the same stimuli as did Weisstein and Haber (1965) in one experiment and H and K in a second experiment. Their composite masking function in the first experiment was flat to 40 msec, with wide individual differences. This suggests strategy differences pertinent to our discussion. No evidence was found for U-shaped masking in the second study where there would be little benefit to the MS as a comparative stimulus.

In conclusion, MS may or may not serve as a comparative stimulus in a pattern recognition task, depending upon the contour relations between TS and MS. It is most capable of serving this role in an equal energy brightness task, whether phenomenal judgments or accuracy is measured and either square-square or disk-ring configurations are the stimuli. As performance changes due to differences in absolute vs comparative judgment strategies are entirely consistent with independent component theories, U-shaped masking functions obtained under these conditions cannot be used per se to decide between lateral inhibition and independent component theories.

REFERENCES

Bernstein, I. H., Amundson, V. E., & Schurman, D. L. Metacontrast inferred from reaction time and verbal report: Replication and comments on the Fehrer-Biederman experiment. Journal of Experimental Psychology, 1973, in press.

Eriksen, C. W. Temporal luminance summation effects in backward and forward masking. Perception & Psychophysics, 1966, 1, 87-92.

Eriksen, C. W., Becker, B. B., & Hoffman, J. C. Safari to masking land: A hunt for the elusive U. Perception & Psychophysics, 1970, 8, 245-250.

Eriksen, C. W., & Collins, J. F. A reinterpretation of one form of backward and forward masking in visual perception. Journal of Experimental Psychology, 1965, 70, 343-351.

Eriksen, C. W., Collins, J. F., & Greenspon, T. S. An analysis of certain factors responsible for nonmonotonic backward masking functions. Journal of Experimental Psychology, 1967, 75, 500-507.

Eriksen, C. W., & Marshall, P. H. Failure to replicate a reported U-shaped visual masking function. Psychonomic Science, 1969, 15, 195-196.

Garner, W. R., Hake, H. W., & Eriksen, C. W. Operationism and the concept of perception. Psychological Review, 1956, 63, 317-329.

Green, D. M., & Swets, J. A. *Signal detection theory and psychophysics.* New York: Wiley, 1966.

Haber, R. N. Note on how to choose a noise mask. Psychological Bulletin, 1970, 74, 373-376.

Kahneman, D. Method, findings, and theory in studies of visual masking. Psychological Bulletin, 1968, 69, 404-425.

Luce, R. D. Detection and recognition. In R. D. Luce, R. R. Bush, and E. Galanter (Eds.), *Handbook of mathematical psychology.* Vol. 1. New York: Wiley, 1963.

Pollack, I., & Hseih, R. Sampling variability of the area under the ROC curve and d'e. Psychological Bulletin, 1969, 71, 161-173.

Schiller, P. H., & Smith, M. C. Detection in metacontrast. Journal of Experimental Psychology, 1966, 71, 32-39.

Schmidt, M. W., & Kristofferson, A. B. Discrimination of successiveness: A test of a model of attention. Science, 1963, 139, 112-113.

Schurman, D. L. Predictive validity of a Rashevsky-Landahl neural net: Text of a model of masking for form. Perception & Psychophysics, 1972, 12, 183-186.

Schurman, D. L., & Eriksen, C. W. Summation and interaction of successive masking stimuli in visual perception. American Journal of Psychology, 1970, 82, 320-332.

Spencer, T. J., & Shuntich, R. Evidence for an interruption theory of backward masking. Journal of Experimental Psychology, 1970, 85, 198-203.

Uttal, W. R. On the physiological basis of masking with dotted visual noise. Perception & Psychophysics, 1970, 7, 321-327.

Uttal, W. R. The psychobiological silly season—or—What happens when neurophysiological data become psychological theories. Journal of General Psychology, 1971, 84, 151-166.

Weisstein, N. A Rashevsky-Landahl neural net: Simulation of metacontrast. Psychological Review, 1968, 75, 494-521.

Weisstein, N., & Haber, R. N. A U-shaped backward masking function in vision. Psychonomic Science, 1965, 2, 75-76.

NOTES

1. Kahneman properly contrasts the apparent movement found in the classical phi phenomenon and that found in metacontrast, pointing out that no suppression occurs in the former configuration. Since the distinction is not necessary to our discussion, we will refer to both cases as apparent movement, without quotes.

2. An alternative that is not inconsistent with our general position is that the effect of presenting stimuli of nearly equal energy in approximate synchrony is better described by a term other than "metacontrast." The presentation of the bright TS + MS (but not the dim TS + MS) combination at 0-msec SOA tended to produce a bar of homogeneous brightness. Hence, metacontrast and paracontrast effects may be discontinuous near 0-msec SOA. Our main point, however, is that any such effects need to be considered relative to changes in performance due to the difference between comparative and absolute judgments.

(Received for publication March 12, 1973;
revision received May 14, 1973.)