Span of attention, backward masking, and reaction time

TADASU OYAMA
University of Tokyo, Tokyo, Japan

TADASHI KIKUCHI
Industrial Products Research Institute, Ibaraki, Japan

and

SHIGERU ICHIHARA
Tokyo Metropolitan University, Tokyo, Japan

A test pattern consisting of 0 to 15 dots and a following random dot masking pattern were presented for 5 msec each with SOAs varying between 30 and 200 msec. The subject was asked to report the perceived number of dots in the test pattern as soon as possible and to assign a confidence rating to each report. The span of attention (upper limit for 50% correct numerosity judgments) increased from 2.4 to 9.5 as the SOA increased. Backward masking reduced the reported number of dots from the actual number in the test pattern, especially with small SOAs. Reaction time increased linearly at a low rate (approximately 40 msec/dot) up to 4 dots in the test pattern and then increased linearly at a high rate (approximately 370 msec/dot) as the *reported*, or *perceived*, number of dots increased. The two different branches of the reaction time curve were considered to represent two separate processes, *subitizing* and *counting*, as suggested by Klahr (1973), who found similar dual increase rates as a function of the actual number of dots. These findings, as well as causal inference based on partial correlations and path analysis, indicated that the reported (perceived) number of dots and confidence rating were both determined by the number of stimulus dots and the SOA and that the reaction time was determined by the so-determined perceived number of dots and level of confidence. A multistage model is proposed.

The terms "span of attention" and "span of perception" have been used to refer to the maximum number of simple objects correctly perceived in a single brief presentation (Graham, 1951; Woodworth & Schlosberg, 1954). Since Jevons' (1871) pioneer work, many experimental studies have been conducted on this subject. The most systematic study in this traditional paradigm was reported by Hunter and Sigler (1940), who found reciprocity between the time and intensity needed to produce the same span of attention in tachistoscopic presentations of stimulus patterns for short exposure durations.

Recently, this old problem has been studied from a new, information-processing point of view.

Averbach (1963) and Lorinstein and Haber (1975) utilized a backward masking technique to assure that processing of the stimulus pattern did not continue internally after the end of exposure. Klahr (1973) measured the reaction time of numerosity judgments for 1 to 20 dots and found two different rates of increase in reaction time: one rate for numbers of stimulus dots up to 4 or 5 and one for greater numbers. He considered these two different rates to represent two different processes—*subitizing* and *counting*—as named by Kaufman, Lord, Reese, and Volkman (1949).

In the present study, an attempt was made to combine two techniques, backward masking and reaction time measurement, in one experiment and to differentiate the processing of the internal image left by the stimulus pattern from the processing of the real stimulus pattern. This was done by keeping the exposure duration very brief and constant, as had been done in the experiments of Kaufman et al. and Lorinstein and Haber. In Averbach's and Klahr's experiments, the exposure duration had been varied with the onset time of mask exposure or reaction occurrence.

METHOD

Apparatus

A stimulus display consisting of a 16 by 16 matrix of light-emitting diodes (LEDs) was positioned 115 cm from the subject, who sat in a chair in a dark booth. The LEDs (Toshiba TLR 103) emitted red light (approximately 40 cd/m^2 at 700 nm), were 5.1 mm in diameter, and were separated by 6.5 mm between centers. The entire matrix of LEDs was 10 × 10 cm in size and 5 × 5 deg in visual angle. The center of the matrix was placed level with the eyes of the subject whose head was held in a fixed position by a chin- and headrest. The LEDs were controlled by a minicomputer (JEC-980 A). A voice key was connected to a chronometer, placed on the experimenter's table outside the booth, to trigger reaction time measurements.

Stimulus Patterns

The test stimuli were 150 random patterns, each of which included from 1 to 15 dots (10 different patterns for each of 15 numerosities). Each test pattern, sampled randomly from the central 10 by 10 matrix of the stimulus display by the controlling computer program, consisted of an array of simultaneously flashed LEDs. Individual LEDs, to be flashed during the subsequent masking pattern, were excluded by an exclusion procedure, so that the same LED was never used in a single trial in both test and masking patterns. Furthermore, LEDs that were adjacent to each other horizontally, vertically, or obliquely were also excluded from the same test pattern. Five different masking patterns were used. Each consisted of 128 LEDs sampled randomly from the entire stimulus display. Four LEDs placed symmetrically 1 cm outside of the corners of the display were always lit dimly to indicate the region in which the test pattern would be presented.

Temporal Conditions

The exposure durations of both the test and the masking patterns were always 5 msec. The stimulus onset asynchronies (SOAs) between the two could be any of the following 12 values: 30, 40, 50, 60, 70, 80, 90, 100, 125, 150, 175, or 200 msec. Two control conditions were added to these 12 masking conditions. In the first control condition (no-mask condition), the test pattern was presented for 5 msec but no masking pattern followed it. In the second control condition (continuous presentation condition), the test pattern was continuously presented for 6 sec, also without any masking pattern.

Subjects

Three well-trained subjects (the second and third authors and an undergraduate student) served in all of the 10 sessions. Each subject had normal visual acuity.

Procedure

Each session was preceded by a 5-min dark-adaptation period. In each session, each test pattern, with varying dot numerosity, was presented once for each of the 12 masking conditions and the 2 control conditions. A blank trial (no test stimulus) was also inserted into each of the 12 masking conditions and the 2 control conditions. Thus, 224 trials [(15 different dot numerosities × 14 masking and control conditions) + (1 blank trial × 14 conditions)] were presented in random order in each session, which lasted about 1 h and 30 min. The same procedure was repeated in 10 sessions, except that a different order of conditions and different combinations of test patterns and SOAs were used in each session.

On each trial, following a warning signal (3-sec intermittent presentation of a buzzer sound), a test pattern (or blank stimulus) and a masking pattern (except in the control conditions) were presented at a particular SOA. The subject was asked to report the perceived number of dots as quickly as possible. The reaction time was measured with the voice key. The subject was also asked to assign a confidence rating to each report, on a 5-point scale, in which "1" meant "the least confident" and "5" indicated "the most confident."

RESULTS AND DISCUSSION

Span of Attention

The "span of attention" has traditionally been defined as the number of dots that can be reported correctly in 50% of presented trials. We obtained the span of attention for each SOA condition and for the no-mask condition from the pooled data of the three subjects, using the method of linear interpolation. Filled circles in Figure 1 indicate the obtained span of attention as a function of SOA. The span of attention was very small for the short SOAs (e.g., 2.4 for the 30-msec SOA) but increased rapidly at first as the SOA increased, to reach the level of 7 or 8 in the SOAs of 80 to 100 msec. Then the rate of increase gradually reduced and almost leveled off when the SOA reached 200 msec. At this point, the span of attention was 9.5 and quite close to that (10.2) measured for the no-mask condition. Individual results also showed the same trend.

The large span of attention obtained in the no-mask condition indicated that a 5-msec exposure was sufficiently long to get normal spans in the present experimental condition (light dots in dark). However, Lorinstein and Haber (1975) found that a 20-msec exposure was needed to produce a span of attention comparable to those obtained in the present study in their experimental condition (black dots on white background). The difference may be caused by the contrast reduction in the brief exposure in the black-on-white condition.

Figure 1 is very similar to Figure 3 in Averbach's (1963) report. It should be noticed, however, that Averbach plotted the span of attention as a function of the *exposure duration*, while Figure 1 in the
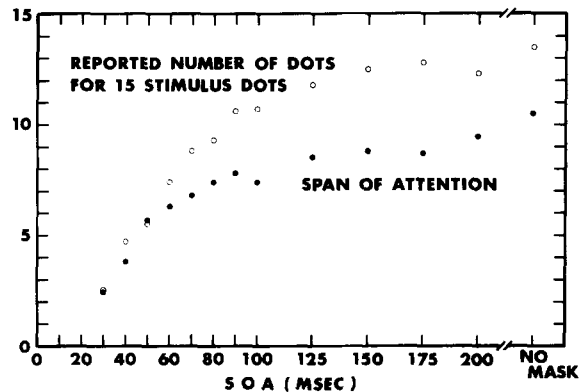

Figure 1. The span of attention (filled circles) and the reported number of dots for the 15 stimulus dots (open circles) as functions of the SOA between the test pattern and the masking pattern.

present paper shows the span as a function of SOA. In Averbach's experiment, he exposed the test pattern for 40 to 600-msec, until the onset of the masking pattern. In the experiment we reported here, we always presented the test pattern for only 5 msec regardless of the SOA condition. The fact that very similar results were obtained in both studies, in spite of large differences in exposure duration, means that the information given by the brief exposure of the test pattern is maintained internally by the subject for a considerable period of time in some form. This is another way of specifying what has been called the visual image (Haber & Hershenson, 1973) or the short-term visual memory (Phillips, 1974). Information stored in this form can be processed in the same way as the information given by a real stimulus.

Reported Number of Dots

Figure 2 shows the mean reported number of dots as a function of the actual number of stimulus dots. The various symbols in this figure represent the different values of SOA. To avoid pointless complication of the figure, only the results obtained in 7 (30-, 40-, 50-, 60-, 80-, 100-, and 200-msec SOAs) of the 12 masking conditions and the no-mask conditions are shown in this figure and in the following figures. Similar results were obtained under the other conditions not plotted here. When no masking pattern was presented, the mean reported numbers of dots were nearly the same as the actual numbers of dots when the test patterns included fewer than 9 dots. However, a slight tendency toward under-

estimation of the number of dots can be found for the test patterns with more than 10 dots. When the masking pattern was presented, this tendency toward underestimation of the number of dots became stronger. As the SOA decreased, underestimation began to occur at a smaller number of stimulus dots and the amount of underestimation became greater. Similar results were reported by Lorinstein and Haber (1975), although they found earlier starts of underestimation and greater amounts of underestimation for corresponding SOAs.

Inspection of Figure 2 suggests that each SOA condition had an asymptotic value toward which the reported number of dots converged; this asymptotic value increased as the SOA increased. The asymptotic value seemed to be 2 or 3 for the 30-msec SOA, 4 or 5 for the 40-msec SOA, about 6 for the 50-msec SOA, and so on. To clarify this tendency, the mean reported number of dots for the test patterns with 15 dots (the largest number of stimulus dots used in the present study) are plotted with open circles as a function of the SOA in Figure 1. These results produce a curve very similar, in shape, to that of span of attention (plotted with filled circles). The curve for the 15 dots rose rapidly at first, on the same line with the span of attention until 50 msec, then continued the rapid rise to reach the level of 10 in 100 msec, and finally gradually decreased the rate of rise to level off at about 200 msec. The rate of rise in the first 100 msec was about 8 msec per dot. Sperling (1969) found a rate of 10 msec per letter in a similar masking experiment. Lorinstein and Haber (1975) reported a smaller scanning time of dots (4 msec per dot) when they observed the initial slope of the curve of the medial number of reported dots as a function of the SOA. The results of these experiments indicate that masking limits the maximum number of items that can be scanned and that the maximum number reported increases nearly linearly for the first several items at the rate of 4 to 10 msec/item.

Reaction Time

Figure 3 shows the geometrical mean of reaction time as a function of the actual number of dots. When no masking pattern was presented, as shown by two broken lines, the reaction time increased linearly at a low rate (42 msec per dot) for 1 to 4 dots and then increased linearly again at a higher rate (367 msec per dot) for 5 to 15 dots. These two different rates were similar to those reported by Klahr (1973) (66 msec per dot for 1 to 5 dots and 268 msec per dot for 6 to 10 dots), to those found by Aoki (1975) (42 msec per dot for 1 to 4 dots and 297 msec per dot for 5 to 15 dots), and to those obtained in the continuous presentation condition in the present experiment (40 msec per dot for 1
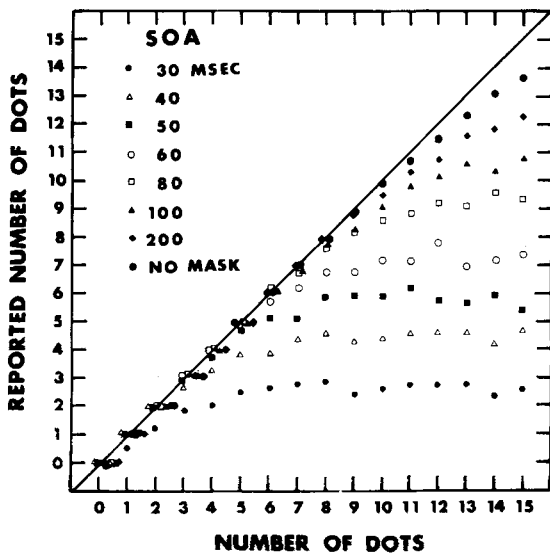

Figure 2. The reported number of dots as a function of the actual number of stimulus dots. Different symbols represent different conditions of SOA.
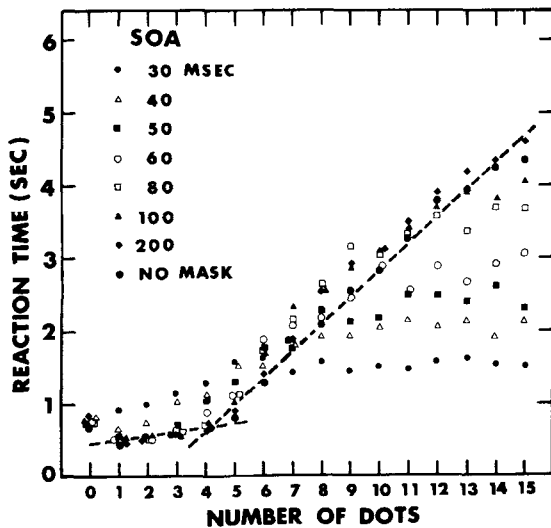
Figure 3. The reaction time as a function of the actual number of stimulus dots. Different symbols represent different conditions of SOA.

to 4 dots and 312 msec per dot for 5 to 15 dots). These two distinct rates are considered to represent two separate mechanisms, *subitizing* and *counting*, as suggested by Klahr (1973). The upper limit of stimulus number to be subitized was found at 4 in the present study, as in the study of Atkinson, Campbell, and Francis (1976).

The mean reaction time to the blank (zero-dot) pattern was 723 msec, a value much longer than the estimated reaction time to the zero-dot pattern (423 msec), that is, the intercept of the straight line fitted to the obtained reaction times to the 1- to 4-dot patterns. The difference would correspond to the time needed to confirm that no dot was present.

The mean reaction time to the 10-dot test pattern, which nearly corresponded to the span of attention in the no-mask condition, was 2,827 msec. The difference between this value and the mean reaction time to the 1-dot pattern (469 msec), the shortest mean reaction time, was 2,358 msec. According to the subtraction method originated by Donders (1969), this difference can be considered to represent the time needed to process the additional 9 dots in the former pattern. Another estimate of the time needed to process 10 dots comes from the rates of increase of reaction time as a function of stimulus dots mentioned previously. According to this idea, the time needed to subitize 4 dots and count 6 dots should be $(42 \times 4) + (367 \times 6) = 2{,}370$ msec. In the no-mask condition in this study, the test pattern was presented for only 5 msec, and in the continuous presentation condition, it was presented for 6 sec. In Klahr's (1973) experiment, the test pattern was presented until the subject's reaction was obtained. In the no-mask condition, the subject

cannot use the real stimulus pattern during processing. This means that some internally stored information (some kind of a visual image or short-term visual memory) left by the test pattern was still available for counting for more than 2 sec in 50% of the trials. However, the small difference between the no-mask condition and the continuous presentation condition in the rate of counting suggests some superiority of the real stimulus over the visual image.

When the masking pattern was presented, the results were quite different from those for the no-mask condition, especially for short SOAs. Reaction time did not increase to any major degree even for larger numbers of stimulus dots. It seemed to level off at shorter times. The masking pattern did not make reaction time longer; on the contrary, it made reaction time *shorter*! An inspection of the raw data indicated a tendency for shorter times to occur when the subjects reported smaller numbers of dots, irrespective of the actual number of stimulus dots. This finding led us to re-analyze the same data in a different way, as shown in Figure 4. In this figure, reaction time is replotted as a function of the *reported* number of dots rather than the actual number of stimulus dots. In this figure, nearly all the plotted points fall in a narrow band, unlike the widely scattered area shown in Figure 3. The reaction times found in short-SOA conditions are longer than those in the no-mask condition in Figure 4. These longer reaction times might be caused by a guessing process which might occur with unclear perception. To test this assumption, in Figure 5 we plotted only those data to which the subjects assigned high confidence ratings (4 or 5). In this figure, it is shown
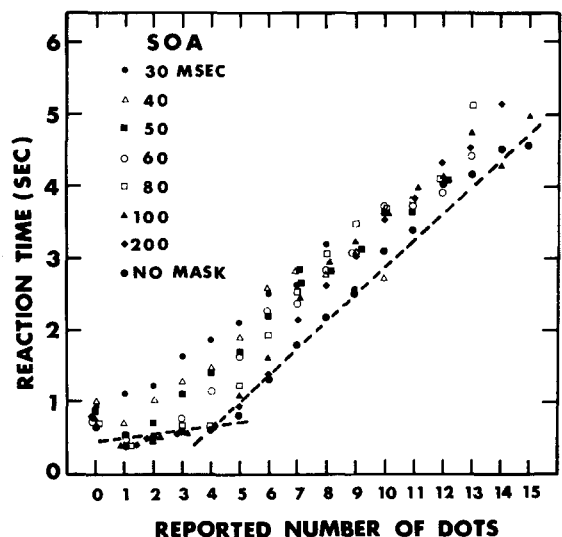

Figure 4. The reaction time as a function of the *reported* number of dots. Different symbols represent different conditions of SOA.
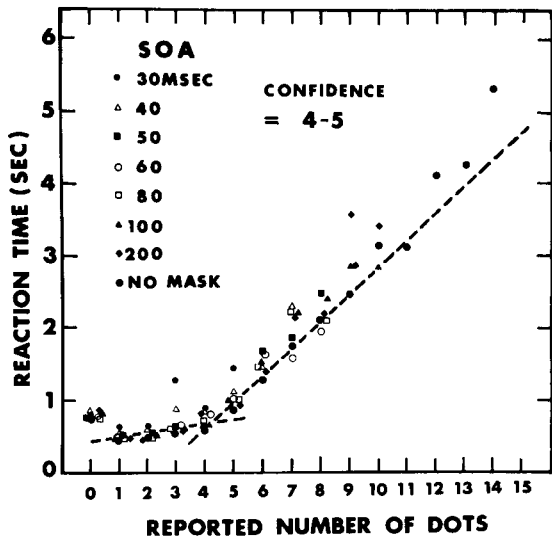
Figure 5. The reaction time as a function of the *reported* number of dots: A plot of those data to which the subjects assigned high confidence ratings (4 or 5). Different symbols represent different conditions of SOA.

more clearly that nearly all the plotted points fall on the same lines, even though the numbers of points decrease for large numbers of dots (only a few high confidence ratings were obtained with large numbers of dots reported). In Figure 5, we can observe two straight lines with different slopes corresponding to subitizing and counting, in all masking conditions at all the various SOAs as well as in the no-mask condition, when the reaction times are plotted against the *reported* number of dots.

The results of these analyses strongly suggest that reaction time is determined by the reported or perceived number of dots rather than by the actual number of stimulus dots. The smaller the reported or perceived number of dots, the shorter the reaction time. The perceived number of dots was reduced by backward masking, which occurred in an early stage of visual processing, and both the subitizing and counting processes occurred in a postmasking stage. The first stage may correspond to the level of visual information storage (Sperling, 1960) or the icon (Neisser, 1967), in which masking occurs and which persists only very briefly, and the second may correspond to the level of visual image (Haber & Hershenson, 1973) or to short-term visual memory (Phillips, 1974), which is not susceptible to masking and which persists longer than the first stage. According to the results of the present experiment, the rate of the information transfer from the first stage to the second is estimated as about 8 msec/dot, at least in the first 10 dots, and the information in the second stage is considered to be available for (subitizing and) counting, at least for 2 sec in 50% of the occasions.

Inference of Causal Relations

An inference method for the determination of causal relations by means of partial correlations was developed by Blalock (1962) and Simon (1954) and has been applied to perceptual problems by Coren and Ward (1979), Oyama (1974a, 1974b, 1977), and van der Meer (1979). According to this model, if the effect of an experimental variable (X) on a dependent variable (Y) is shown by experimental data or by a significant simple correlation between X and Y, and the partial correlation between these two variables with a third variable (Z) held constant ($r_{XY \cdot Z}$) is nearly zero, the effect of X on Y is considered to be mediated by Z.

Figure 6 shows the results of the application of this method in the present experiment. The second- or third-order partial correlations between every combination of two variables are obtained from the individual data of the three subjects. The solid lines in this figure indicate combinations of variables between which large and significant partial correlations were found, while dashed lines represent combinations of variables between which only small insignificant partial correlations were found. According to the Simon-Blalock theory, solid lines represent direct causal relations, while dashed lines represent indirect relations. As the lines and partial correlations in this figure indicate, the reported or perceived number of dots (n) increases with an increase in the actual number of stimulus dots (N) and an increase in the stimulus onset asynchrony (t), while the confidence rating (C) in-
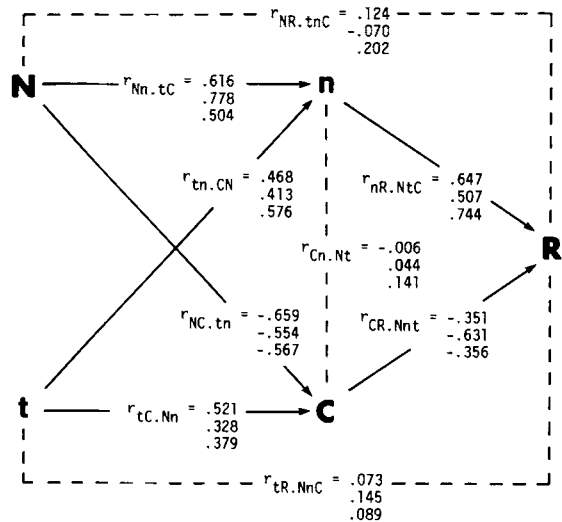

Figure 6. The results of causal inference by means of partial correlations. N, t, n, C, and R represent the actual number of stimulus dots, the SOA, the reported number of dots, the confidence rating, and the reaction time, respectively. Three values between every combination of variables represent the second- or third-order partial correlations obtained from the individual data of the three subjects. The solid lines represent direct causal relations, and the dashed lines, indirect relations.

creases with a decrease in N and an increase in t. Finally, the reaction time, R, increases with an increase in n and a decrease in C. The two stimulus variables (N and t) and reaction time are related only indirectly. The path analysis technique (Wright, 1960) was also applied to the experimental data. The squared multiple correlations, R², obtained from the data of individual subjects by applying the same model as that shown by the solid lines in Figure 6, are fairly large (.8422, .8559, and .8477) and practically identical to those obtained using a more complex model that included the relations indicated both by the solid and the dashed lines in Figure 6 (.8456, .8601, and .8646). The results of this analysis also supported the tentative conclusion that subitizing and counting are processes that occur at a postmasking stage or stages.

Tentative Multistage Model

Figure 7 presents a multistage model to account for our experimental results. The first stage is VIS, or iconic storage, which persists for only a fraction of a second. The information given by the test pattern is initially stored at this stage, and then subsequently transferred to the second stage, which may be called visual image or short-term visual memory, at a rate of about 8 msec/dot, until presentation of the masking pattern. Information about the test dots, which has not been transferred to the second stage prior to the masking pattern, is thereafter mixed with the masking pattern dots and cannot be discriminated. That is what we mean by masking! A visual image is not susceptible to visual masking and continues to exist for longer than 2 sec, as shown by reaction time to the 10-dot pattern, which corresponds to the span of attention. During the period of persistence of the visual image, the subject is able to subitize the number of dots at a rate of approximately 40 msec/dot and to count the dots in this visual image at a rate of approximately 370 msec/dot, even after the test stimulus has disappeared. The reported or perceived number of dots

represents the number of dots stored in the visual image. The reaction time is determined by the number of dots in the visual image, rather than the number in the stimulus pattern. Subitizing and counting represent two kinds of information transfer from the visual image to the third stage, short-term memory or short-term verbal memory (Murray & Newman, 1973). Finally, the response is produced from the verbalized information stored in this third stage. Reaction time corresponds to the total processing time plus the simple reaction time.

CONCLUSIONS

(1) The "span of attention" increased as the SOA between the test pattern and masking pattern increased, even when the exposure duration was kept at 5 msec. The span of attention increased rapidly at first, reaching a level of 7 or 8 at approximately 100-msec SOA, and then gradually leveled off at 9 or 10 when the SOA reached 200 msec.

(2) When no mask was presented, the mean reported number of dots was almost the same as the actual number of stimulus dots, but a slight tendency toward underestimation appeared for the stimulus patterns that consisted of more than 10 dots. When the mask was presented, the tendency toward underestimation also occurred for smaller numbers of stimulus dots. The reported number of dots then leveled off, even when the number of stimulus dots increased. The asymptotic value for the reported number of dots was reduced as the SOA became shorter.

(3) When no mask was presented, the reaction time increased linearly at a rate of approximately 40 msec/dot for 1 to 4 stimulus dots and then increased linearly again at a higher rate of approximately 370 msec/dot for stimuli consisting of 5 to 15 dots. These two different rates were considered to represent two separate processes, subitizing and counting, as already suggested by Klahr (1973). When the mask was presented, reaction time was generally decreased. This decrease in reaction time was considered to be caused by the decrease in the perceived number of dots effected by backward masking. Reaction time increased as the *reported* number of dots increased, at two constant rates, which should represent the subitizing and counting processes. This tendency became clearer when only the data obtained with high confidence ratings were plotted.

(4) The methods of causal inference by means of partial correlations and of path analysis were applied to the obtained data. They indicated that both the reported or perceived number of dots and the confidence rating were determined independently by the two stimulus variables, the actual number of stimulus dots and the SOA between the stimulus pattern and the masking pattern. They also indicated that the reaction
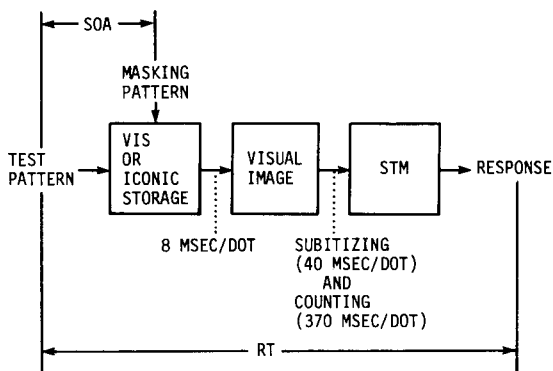

Figure 7. The proposed multistage model to account for the observed relations.

time was determined by the perceived number of dots and the confidence level. There was no direct relation between the stimulus variables and the reaction time.

(5) A multistage model explains the observed relations. The first stage is VIS, or iconic storage, in which the information given by the test pattern is initially stored for a very brief moment and then is transferred to the second stage at a rate of about 8 msec/dot. When information about the test and masking patterns is mixed before the former is transferred, masking occurs. The second stage is visual image or short-term visual memory, which exists for more than 2 sec and is not susceptible to masking. During this time, the subject can subitize or count the number of dots stored at this stage. The subitizing and counting represent two kinds of information transfer from the second stage to the third stage, short-term memory or short-term verbal memory. The response is produced from the verbalized information stored at this third stage.

REFERENCES

Aoki, T. [A study on enumeration processes. Essays in celebration of the 10th anniversary of Aichi Prefectural University]. Nagoya: Aichi Prefectural University, 1975. (In Japanese)

Atkinson, J., Campbell, F. W., & Francis, M. R. The magic number 4±0: A new look at visual numerosity judgments. Perception, 1976, 5, 327-334.

Averbach, E. The span of apprehension as a function of exposure duration. Journal of Verbal Learning and Verbal Behavior, 1963, 2, 60-64.

Blalock, H. M. Four-variable causal models and partial correlations. American Journal of Sociology, 1962, 68, 182-194.

Coren, S., & Ward, L. M. Levels of processing in visual illusions: The combination and interaction of distortion-producing mechanisms. Journal of Experimental Psychology: Human Perception and Performance, 1979, 5, 324-335.

Donders, F. C. On the speed of mental processes. Acta Psychologica, 1969, 30, 412-431. (Translated by W. G. Koster from Dutch original paper in 1869.)

Graham, C. H. Visual perception. In S. S. Stevens (Ed.), Handbook of experimental psychology. New York: Wiley, 1951.

Haber, R. N., & Hershenson, M. The psychology of visual perception. New York: Holt, Rinehart & Winston, 1973.

Hunter, W. S., & Sigler, M. The span of visual discrimination as a function of time and intensity of stimulation. Journal of Experimental Psychology, 1940, 26, 160-179.

Jevons, W. S. The power of numerical discrimination. Nature, 1871, 3, 281-282.

Kaufman, E. L., Lord, M. W., Reese, T. W., & Volkmann, J. The discrimination of visual number. American Journal of Psychology, 1949, 62, 498-525.

Klahr, D. Quantification processes. In W. G. Chase (Ed.), Visual information processing. New York: Academic Press, 1973.

Lorinstein, B., & Haber, R. N. Perceived numerosity: An information processing analysis. Canadian Journal of Psychology, 1975, 29, 224-236.

Murray, D. J., & Newman, F. M. Visual and verbal coding in short-term memory. Journal of Experimental Psychology, 1973, 100, 58-62.

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

Oyama, T. Inference of causal relations in perception of space and motion. Psychologia (Kyoto), 1974, 17, 166-178. (a)

Oyama, T. Perceived size and perceived distance in stereoscopic vision and an analysis of their causal relations. Perception & Psychophysics, 1974, 16, 175-181. (b)

Oyama, T. Analysis of causal relations in the perceptual constancies. In W. Epstein (Ed.), Stability and constancy in visual perception. New York: Wiley, 1977.

Phillips, W. A. On the distinction between sensory storage and short-term visual memory. Perception & Psychophysics, 1974, 16, 283-290.

Simon, H. A. Spurious correlation: A causal interpretation. Journal of the American Statistical Association, 1954, 49, 467-479.

Sperling, G. The information available in brief visual presentations. Psychological Monographs, 1960, 74, 1-29.

Sperling, G. A model for visual memory tasks. In R. N. Haber (Ed.), Information-processing approaches to visual perception. New York: Holt, Rinehart & Winston, 1969.

Van der Meer, H. C. Interrelation of the effects of binocular disparity and perspective cues on judgments of depth and height. Perception & Psychophysics, 1979, 26, 481-488.

Woodworth, R. S., & Schlosberg, H. Experimental psychology (Rev. ed.). New York: Holt, Rinehart & Winston, 1954.

Wright, S. Path coefficients and path regressions: Alternate or complementary? Biometrics, 1960, 16, 189-202.