Journal of Experimental Psychology: General

VOL. 111, NO. 1 MARCH 1982

Subitizing: An Analysis of Its Component Processes

George Mandler and Billie Jo Shebo
University of California, San Diego

SUMMARY

The term *subitizing* was coined by Kaufman, Lord, Reese, and Volkmann in 1949 to describe the rapid, confident, and accurate report of the numerosity of arrays of elements presented for short durations. They noted that this process, different from counting and estimating, was restricted to arrays with 6 or fewer elements. Ever since the general awareness of some such process in the nineteenth century, the phenomenon has been a benchmark for the limited capacity of human consciousness. Previous research, as well as the data in these experiments, shows that the reaction time function for different arrays (ranging in size from 1 to 15) has a shallow slope for arrays with 1 to 3 elements followed by a straight line slope for arrays of 4 to 6 or 7, at which point the reaction time discontinuity typically occurs; reaction times then stay fairly constant as the array size increases, and the numerosity response becomes much less accurate. When the array is exposed for unlimited presentation time, the reaction time slope is a straight line from size 4 to as large as 30 or 40. In the latter case, subjects are clearly counting the array.

Apart from replicating the previous findings under a variety of conditions, we have shown that subjects can make a very fast "countability" judgment indicating whether or not they could, if requested, give an accurate numerosity response. These judgments are fast and produce a yes response within the subitizing range, and a no response thereafter.

Developmental data had indicated that children count arrays as small as 2 and 3; adults seem to give a more automatic response, shown also in the fast reaction times to those array sizes. The suggestion that this response is an acquired one to certain frequently appearing canonical patterns of two and three events (pairs/lines and triples/triangles) was explored in another experiment in which subjects were given canonical patterns for arrays of up to 10. These data showed that within very few trials the response to these canonical patterns was just as fast and accurate as the response to the smaller (1 to 3) array sizes. Finally, the data demonstrated that the slope for array sizes from 4 to 6 with short exposure time was indistinguishable from the slope for array sizes from 4 to 15 under an unlimited exposure condition.

We concluded that the reaction time function found in subitizing consists of three processes: a response to arrays of 1 to 3 that is fast and accurate and is based on acquired canonical patterns; a response to arrays of 4 to 6 or 7 that is based on mental counting, that is, the counting of arrays that can be kept in consciousness (attention); and an estimating response for arrays larger than 6 that cannot be held in consciousness for mental counting.

Copyright 1982 by the American Psychological Association, Inc. 0096-3445/82/1101-0001$00.75

The postulation of a limited capacity system that governs human attention and consciousness has been central to modern cognitive psychology since Miller's (1956) review of the limitation on human information processing. Among the instances Miller cited was the phenomenon of subitizing, the limitation of the rapid apprehension of numerosity to some 6 or fewer events. In the intervening quarter of a century the limited capacity notion has been extended to a variety of phenomena in human judgment and attention. In our work we have focused on the effect of the limited capacity of consciousness on memory and semantic categorization (Graesser & Mandler, 1978; Mandler, 1967, 1975a). Given some of these extensions, it was surprising that a basic reference phenomenon such as subitizing had never been given a satisfactory explanation and description. The purpose of the present experiments is to provide such an explanation.

The reference experiment for the subitizing phenomenon was an extensive study by Kaufman, Lord, Reese, and Volkmann (1949). They presented arrays of dots, with the number of dots varying from 1 to 210. Presentation time was 200 msec without a mask. When subjects were instructed to strive for accuracy, performance was essentially perfect up to arrays of 5; it was not much different when subjects were given speed instructions.

In their discussion, Kaufman et al. (1949) distinguished among three processes: subitizing, counting, and estimating. They invented the term *subitize* to describe the process, different from estimating, that characterizes the report of numerosity for arrays of 6 or fewer events. *Subitize* is derived from "the classical Latin adjective *subitus,* meaning *sudden,* and the medieval Latin verb *subitare,* meaning *to arrive suddenly*" (p. 520, footnote). Kaufman et al. reserve the process of estimating for arrays of 6 or more events. The process of counting refers to the usual pairing of objects with the incrementing number series. Kaufman et al. noted that estimating and subitizing are similar operations, essentially distinguished only by the array size. Subitizing is a "more accurate, more rapid, and more confident process" than estimating (p. 525). They based their conclusion on the discontinuities (systematic abrupt changes in slope) for both the time and confidence judgments at or near the 6-dot array. They also concluded that there is no such phenomenon as "the immediate cognition of number" (p. 525). Their basic data on reaction times to arrive at numerosity judgments are presented in Figure 1, which shows the results of Kaufman et al., as well as the data of Jensen, Reese, and Reese (1950) with a subject-determined presentation time. The basic phenomenon, repeated by others, is that the time to report the numerosity of arrays increases up to arrays of 6 and then stays essentially flat for the increasing array sizes. If presentation time of the array is terminated only when the subject responds, as in the Jensen et al. study, reaction time is apparently linear with increasing array size.

In summary, Kaufman et al. invoked the term *subitizing* for a process of arriving at judgments of visual numerosity that applies to arrays of 6 or fewer objects. They did not specify what that process might be. Some of the subsequent discussions in the literature might well have paid heed to the coiners of the term. Subitizing is defined by systematic changes in the slope for judgments of numerosity; it is not a concept disputable on theoretical grounds. To assert that people do or do not subitize requires only a demonstration that certain discontinuities can or cannot be demonstrated in the data obtained from numerosity judgments. To determine what processes are involved during subitizing is indeed a theoretical enterprise, which Kaufman et al. touched on only tangentially.

Cattell (1886) concluded the first experimental investigation of the phenomenon with a very modern conclusion:"[W]hen 4 or 5 lines are visible for 0.01 second, their number is usually estimated correctly, which

The research reported here and the preparation of this report were supported by National Science Foundation Grant BNS 79-15336. We want to thank Patricia Haden for experimental assistance. A brief report of the findings was presented at the 1980 meeting of the Psychonomic Society in St. Louis.

Requests for reprints should be sent to George Mandler, Center for Human Information Processing, University of California, San Diego, La Jolla, California 92093.

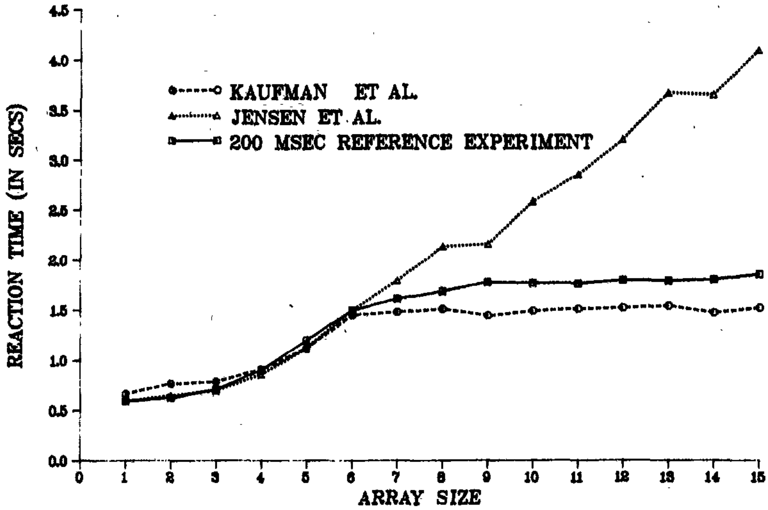

Figure 1. Mean reaction times to report number of items as a function of array size from Kaufman et al. (1949), Jensen et al. (1950), and the 200-msec condition from Experiment 1.

gives the number of simple impressions that can simultaneously be held in consciousness" (p. 124). Similarly, Jevons (1871) talked about counting "by a single mental act" (p. 281), and Wundt (1896) claimed that a maximum of 6 random dots can be "apperceived at one time" (p. 248).

A similar conclusion was reached by Warren (1897), who claimed to have found evidence for "perceptive" counting that proceeds by an apprehension of the group as a whole. He based his conclusion on the very slight slope in reaction times up to sets of 4 combined with the very low (nearly zero) error rates of sets of up to 3. These data were obtained with a 131-msec exposure rate; for unlimited exposure he found a fairly constant slope of 200 msec for sets of 1 to 8. Warren also suggested that "progressive" (incremental) counting has an upper limit of 7 with the short exposure time.[1]

The position held by Cattell, Jevons, Wundt, and Warren led to one of two views on the topic that are held to this day. The first is that the perception of small numbers of objects (variously considered to be from 3 to 6) is a unitary direct perceptual event. For example, Bourdon (1908) claimed that they are sensations of the same order as blue, green, red, round, square. That view has been held over the past century by Bourdon, Klahr (1973a, 1973b), and Taves (1941).

The other, less popular, view claims that since there is an increase in the time of perception with the number of events, the perception of small numbers is continuous with the perception of larger numbers (i.e., counting). However, since the function relating reaction time to display size is usually positively accelerated, the typical resolution has been to fit a logarithmic function to the data. This point of view is typically represented by von Szeliski (1924) and Saltzman and Garner (1948). It is important to note that changes in the slope of the function for accuracy, reaction time, and confidence at and around 6 are usually found only if the display is presented for brief (several hundred millisecond) exposures. These discontinuities, or changes in the slope of the function, do not appear if the exposure time is subject determined. Under those conditions one typically finds a shallow slope for displays of up to about 3 items, followed by a steep, linear continuous slope for displays of from 15 to

[1] In relation to an argument we shall develop later, Warren (1897) noted that familiar patterns facilitated counting, particularly when they were regular; he called this phenomenon "inferential" counting. He concluded that "inference tends to shorten progressive counting and to lengthen perceptive counting, when it takes their place wholly or in part" (p. 590).

30 items. Figure 1 shows this effect for the data by Jensen et al. (1950). Recently it has also been demonstrated for younger subjects. Svenson and Sjöberg (1978), using subject-determined display times, concluded that their data with 7–8-year-old children showed a shallow slope within a subitizing range of 1 to 3, and a counting slope with displays of 5 and more.

The best summary of the accuracy of numerosity judgments has been presented by Woodworth and Schlosberg (1954). Collating data from three different experiments (Fernberger, 1921; Glanville & Dallenbach, 1929; Oberly, 1924), all of which used collections of dots on white stimulus cards with a 100-msec exposure, they found average correct responses of 100% for sets of 2 and 3 dots, 98% for 4, and 38% for 7, followed by a discontinuous drop to 3% for 8 and 0% for 11.

Woodworth and Schlosberg suggested that small numbers can be directly perceived because "two items are seen as a pair, three items a trio, four as a quartet, and so on. . . . [A]nd sometimes you can see a collection as falling apart into a few small *groups* as was suggested . . . [by] Hamilton [1859]" (pp. 94–95). They raise the question of how far and under what conditions one can use the "direct" perception of small groups for the perception of number. As far as the reaction times for the report of numerosity are concerned, Woodworth and Schlosberg conclude from the Kaufman et al. and other studies that "there is a distinct process with an upper limit at about 6 units" (p. 100).

The primary focus of our investigation is the shape of the reaction time function. We have already noted that the typical shape is a slowly accelerating function for display sizes 1 to 3, an essentially linear function for sizes 4 to 6 under experimenter-determined presentation times, and a similar linear function for sizes from 4 up to large displays of 20 to 200 events when presentation time is subject determined.

A somewhat different result was obtained by Atkinson, Campbell, and Francis (1976), who studied numerosity judgments for visual arrays of dots and lines presented for 150 msec. Under some conditions they found essentially a flat slope for arrays of 1 to 4. One possible reason for that finding was that reaction times were obtained by subjects releasing a manual response key while giving a verbal response. Thus, as we also note below, the reaction times might not have been the reaction times of the numerical response as such but rather of the perception of a "countable" array (Atkinson, Note 1).

Klahr generated several related models for the subitizing effect. A representative one (Klahr, 1973a) involves an attempted match between the stimulus elements and a template sequence that starts with a template of 1 and is incremented by single additional elements, that is, from n to $n + 1$. As soon as the stimulus elements match the "pieces" in the template, the subitizing goal is satisfied. The response symbol is then accessed. The discontinuity seen in subitizing is incorporated as one of the assumptions of the model; Klahr assumes that the subitizing process is limited to a stimulus element set of 4 (sometimes 5) or smaller. This model requires a linear constant slope for display sizes 1 to 4 (or 5). However, none of the available results is easily fitted to this procrustean bed, and Allport (1975) has shown how Klahr's attempt (Klahr, 1973b) to do so violates his and other data sets.

Gelman and Gallistel (1978, pp. 65–72) discussed Klahr's claims in detail with respect to the "subitizing" performance of young children. Their review of the literature indicates that children as young as 2–3 years old not only willingly count arrays of 2, 3, and 4 items but do so before they use any perceptual or grouping mechanisms for this kind of task. If counting precedes perceptual apprehension in ontogeny, it is unlikely that the kind of process that Klahr advocates is an invariant characteristic of the human individual.

Gelman and Gallistel (1978) also examined in detail the claim that the human individual has available a mechanism that provides for the "direct perceptual apprehension" of the numerosity of arrays of up to 5 or 6. One of the questions they pose, and one that we will answer indirectly, is, What is it that "all possible configurations of three elements, both linear and nonlinear, have in common that distinguishes them from all possible configurations of two, four, five and six elements?" (p. 222). They note that most of the data on young children (e.g., Gelman & Tucker, 1975) can be accounted for by "the application of a counting procedure following a clear perception of the array" (p. 221).

Gelman and Gallistel considered the claim that children (and adults) might have concepts of twoness and threeness just as they clearly have concepts of treeness and cowness. The fact that we do not clearly understand what featural invariances might underlie such concepts does not invalidate them. What does make them difficult, however, is the fact that apparently a positive slope exists in the subitizing range. There is nothing in the nature of concepts such as threeness (or treeness) that would predict a natural and systematic increase in the time needed to apply them to the respective arrays. Thus, the tendency of small children to count arrays as small as two or three, and the argument just cited, suggests that the subitizing phenomenon (both the discontinuity and the positive slope within the subitizing range) is unlikely to be a function of some built-in mechanism of direct perceptual apprehension of numerosity.[2]

Gelman and Tucker (1975) presented data that at least suggest that children in the 3–5-year age range count arrays of 1 to 5. Accuracy was independent of item heterogeneity (color and shape) and of visual angles from 1.7° to 25.4°. Increases of exposure time from 1 to 60 sec increased accuracy. Gelman and Tucker suggested that these findings are best accounted for by a counting process, particularly since increases in exposure time affected sets as small as 2 and 3 for the youngest children.

Chi and Klahr (1975) reported number-estimating data for adults and for 5–6-year-old children. The task required subjects to report the number of dots on a screen, using subject-terminated, unlimited exposure times. Chi and Klahr claimed different slopes for the 1–3 and 4–10 ranges. However, inspection of their Figure 1 suggests that the 3-dot display for children is indistinguishable from a point predicted from the slope for the 4–10 range. The increment from 1-dot to 2-dot displays is clearly less (about 180 msec) than the slope for 3 to 10 dots (1,049 msec). For adults, both slopes are significantly lower (46 and 307 msec, respectively). In the case of adults, it appears that the 1 to 3 slope is in fact distinguishable from the 4 to 10 slope. In addition, children are also operating at a slower rate for both ranges. These data suggest that adults use a different process for the 1–3 range than for the 4+ range, whereas children operate differently on the 1–2 range than on the 3+ range.

We have discussed the data on children's subitizing performance at some length because the argument we shall present is essentially an ontogenetic one. If it is the case that children count arrays as small as 2 or 3 but that adults apparently do not, then it would be reasonable to assume that adults have developed a novel strategy for dealing with these arrays. Such a strategy would be different from that used for arrays of 4 or more.

Our basic argument is that adults have developed canonical pattern perceptions for arrays of 2 and 3—doublets (perceptually straight lines) and triplets (triangles). These canonical patterns develop slowly during childhood, such that 5-year-olds apparently have adopted the twoness pattern but not yet the triplet triangular pattern. Random presentation of displays of 2 and 3 will always produce the canonical pattern for 2 and frequently the canonical triangle for 3. If there is a canonical pattern for fourness it presumably is a square array, but random generation will not frequently produce such a pattern.

We believe that if canonical patterns are not available, adults count the arrays. If the array can be "held in consciousness," that is, if it has 6 or fewer elements, it can be counted accurately after short presentations of a few hundred milliseconds. For arrays with more than 6 elements, postexposure counting fails and estimating procedures are substituted.

One additional point is worth noting with reference to the Chi and Klahr data dis-

[2] Starkey and Cooper (1980) recently reported that 22-wk.-old infants can discriminate between arrays of 2 and 3 objects but not between arrays of 4 and 6 objects. The implication of this finding for an unlearned apprehension of *number* is not clear.

cussed above. We mentioned that the counting slope for children was nearly three times greater than that for adults. Do children merely count slower than adults, or do they count differently? The former may well be the case, but the latter certainly is. Adults do not typically count by ones; they frequently count by twos, often by chunks of threes, and sometimes even by fours and fives.

Experiment 1

The first study in this series was designed to replicate the existing data from the literature and to calibrate our experimental conditions. Apart from ruling out possible artifacts such as head movements, saccadic movements, and peripheral vision, we wanted to have available a reference experiment against which subsequent experimental manipulations could be compared. Therefore, in addition to protecting the results from artifactual effects, we wanted to compare (a) a range of presentation times and (b) homogeneous and heterogeneous arrays.

Method

Design. All the experiments were conducted on, and all the materials were generated by, a PDP-12 computer. Subjects were presented with arrays of sizes 1 to 20 for four different display durations (100, 200, 400, and 800 msec). Each subject was given two blocks of 800 trials each. The two blocks differed in display time; two subjects each were given each of the four display times in the first block and a different display time for the second block. Thus the 100-msec condition in the second block was preceded by 400- or 800-msec conditions, the 200-msec condition by 100 or 400 msec, the 400-msec condition by 200 or 800 msec, and the 800-msec condition by 100 or 200 msec.

The displays consisted of random presentations of the letters X and O on a cathode-ray tube screen. Each display was either "pure" (all Xs or all Os), or "mixed" (a mixture of Xs and Os). Each block of 800 trials consisted of 40 trials each for each display size; there were 20 pure and 20 mixed cases for each display size from 2 to 20 and 40 single-letter trials for display size 1. The order of presentation of pure and mixed cases was random.

After completing the two blocks, each subject was also given the same display task at a 50-msec exposure rate to determine if extensive practice affected performance on the task at marginal display times. Another four subjects were given the full display task at the 50-msec exposure rate without any prior practice.

The subjects in the full experimental design were also given two control tasks after completion of the 50-msec task. The first (Control A) investigated whether differences in reaction times could be ascribed to differences in retrieval times for number names; numerals from 1 to 20 were presented in random order and subjects responded with their names. The second (Control B) condition was designed to test the voice-activated relay and to determine if pick-up time differed for different numeral names. Subjects were given the numerals 1 to 20 in order and asked to name them as quickly as possible.

Display generation. Pure and mixed trials were generated for two different sets of displays. Each set contained 10 different pure and 10 different mixed cases for each display size. For the pure cases (in which the display consisted entirely of an array of Xs or Os), five trials were generated with all Xs and five with all Os. For the mixed cases, all the possible combinations of Xs and Os for a given display size (from 2 to 20) were generated. For example, display size 3 included displays with one X and two Os and with two Xs and one O. If the number of possible X and O combinations was less than 20, the remaining combinations were selected randomly from the original set. Each trial for every display size was generated by randomly choosing the appropriate number of spaces in a 64-space (8×8) matrix.

Procedure. A total of eight volunteer paid subjects participated in eight sessions, lasting approximately 1 hr. each. Each session was divided into two sections of 100 trials. The subject sat facing the scope where the displays were presented. The subject's head was placed in a chin rest with the display 119.4 cm from the subject's eyes. The matrix underlying the display was a 3.6-cm square, extending a visual angle of 1.7°. Of course, the subject saw only the Xs and Os displayed on the screen and not the underlying matrix. A tone sounded to indicate that the next trial was about to start, and the subject pressed a ready button. A dot was presented for 100 msec in the center of the display and the subject was required to fixate on it. The display appeared on the screen 400 msec after the termination of the dot. The display consisted of 1 to 20 Xs and/or Os. It stayed on the screen for the determined display time, after which it was displaced by a random dot masker covering the entire underlying matrix. The subjects were instructed to respond verbally as quickly and as accurately as possible, giving the total number of items presented. The masker (or the display when the subject responded prior to display termination) was turned off as soon as the subject responded. Each response triggered a voice-activated relay interfaced to the computer. The subject's reaction time was automatically recorded and the experimenter noted the numerical response. For the few occasions when the voice relay responded prior to the subject, the relevant trials were rerun at the end of the session. Approximately once every 20 trials the subject was asked, after giving the numerical response, whether the display was pure or mixed. This manipulation was to assure that subjects attended to the display type.

Reference condition. After the data on these initial conditions had been collected, and it had been decided to use the 200-msec condition as the baseline or reference condition, another five subjects were run on the 200-msec condition with an array range of 1–15. All data presented below for the 200-msec condition are for all nine subjects who had been given the 200-msec presentation time.

Results and Discussion

Figure 2 shows the results for Control Conditions A and B. For Condition A, there is a variable response time to retrieve and pronounce the appropriate response for different numerals. The only important effect is the positive slope for array sizes 4 to 6, which might be reflected in the experimental data. However, the relatively small slope of 58 msec suggests that, at worst, some constant needs to be subtracted from slopes for the array sizes 4 to 6 in the various experimental conditions. Such a procedure would not affect any of the results or conclusions reported. The apparent increase in reaction time from 2 to 3 might be important, if it were not paired with a similar decrement from 3 to 4. These two changes in the control data are apparently not systematically related to any other observations in our or other existing observations. Condition B shows essentially no differential sensitivity of the voice-activated relay to the names of the different numerals.

Figure 3 shows three different analyses for the 200-, 400-, and 800-msec conditions. In all cases here, and in subsequent experiments, the data for the pure and mixed presentation conditions were combined. These two conditions were always indistinguishable, both for the average and the individual data. Subjects responded to the number of elements in the array independent of the nature of those elements, that is, whether the display was pure or mixed.

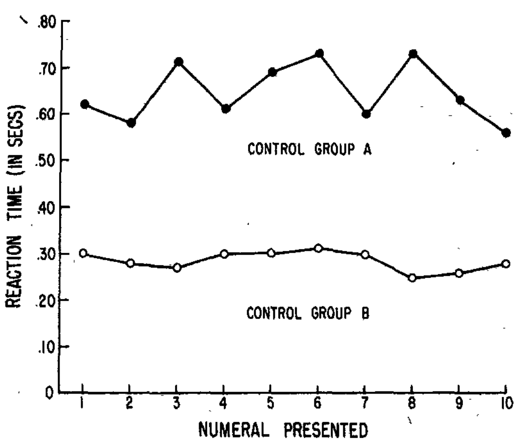

Figure 2. Mean reaction times for the two control conditions (naming numerals presented randomly in Control Condition A, and naming numerals presented sequentially in Control Condition B).

It was our hope that an exposure time of 200 msec or less would provide stable data, thus avoiding the contamination of the numerosity judgments by the occurrence of saccadic movements. The 100-msec condition (shown in Figure 4 and to be discussed shortly) provided unstable data, but exposure times of 200 msec and greater showed stable and comparable results.[3] The reaction times (central panel of Figure 3) show the classical subitizing function for all three exposure durations: a slowly accelerating reaction time function for array sizes 1 to 3, a straight line function for array sizes 4 to 6, and a flat function starting at array sizes 6 to 8. The break is clearly at size 6 for the 200-msec condition and later for the other two. Comparison of the 200-msec condition with the Kaufman et al. (1949) data in Figure 1 shows that, except for a higher asymptote, our data are generally comparable with previous findings in the subitizing range.[4] The top panel of Figure 3 shows the proportion of errors (proportion of incorrect numerical responses). Here the 400- and 800-msec conditions are essentially identical and superior to the 200-msec condition in the 4–10 range. Again, the 200-msec condition commends itself as a baseline condition, since it provides better than 50% accuracy up to array size 6. The mean response (shown in the bottom panel), on the other hand, shows accurate average response for all conditions for the smaller arrays. There-

[3] It was relatively easy to define the stability of an effect on the basis of individual subjects' results. Reaction times in particular provide highly similar results when the functions from individual subjects are inspected. This was the case for the 200-msec presentation time data (but not for 100 msec). Although we present only average data for this and other experiments, in all cases individual subjects' results were inspected and, except for greater variability, were essentially of the same form as the average data.

[4] The possibility exists that the flat portion of the functions (beyond array size 8) is due to an artifact of what the subjects actually report. If regardless of array size subjects report only about 8 items, the flat function would be obtained. However, the function of reaction times plotted against the number of items *reported* by the subjects is indistinguishable from the data in Figure 3.

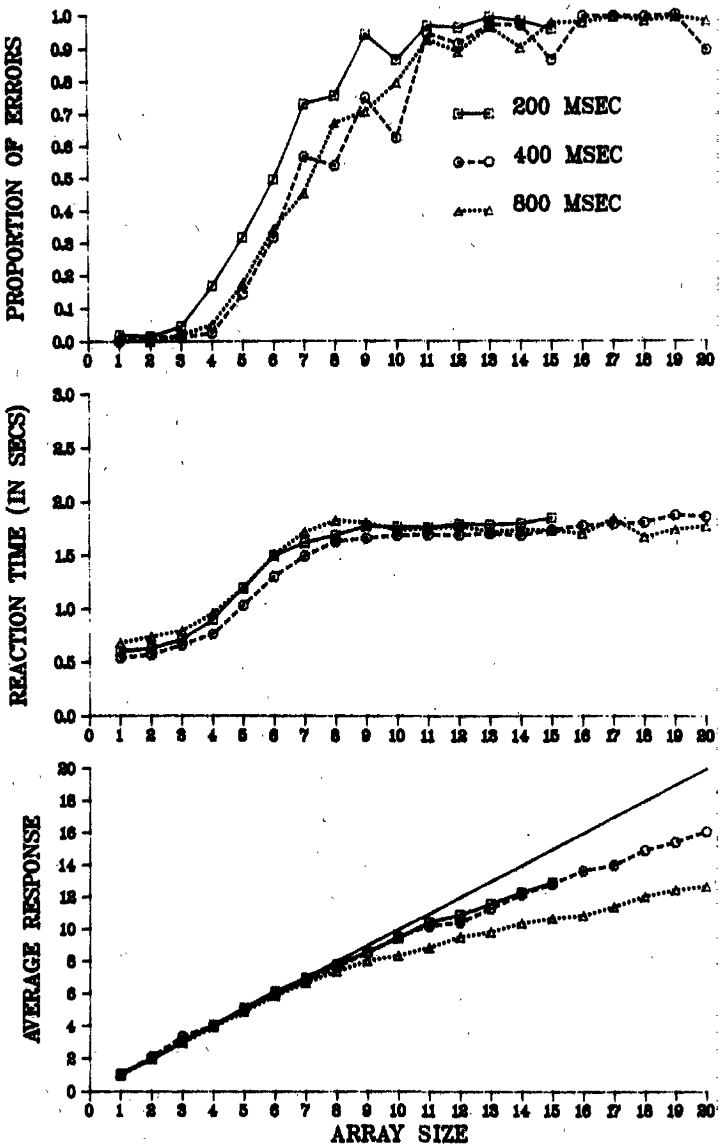

Figure 3. Results for numerosity judgments under three exposure conditions (200, 400, and 800 msec) for proportion of incorrect responses (top panel), reaction times (middle panel), and average number response (bottom panel).

after, the 200- and 400-msec conditions are highly similar and the 800-msec condition is clearly inferior.[5]

These data suggest that for the longer presentation times subjects might attempt to count the larger arrays, particularly in the 800-msec condition. This results in longer reaction times, fewer errors, but also worse average response. Additional saccadic movements during the presentation of the array would also contribute to a better scanning of the array beyond the subitizing range, as well as longer reaction times. Since our con-

[5] The average numerical response is a rather noisy dependent variable and should be used with caution. In some cases the removal of one or two subjects who consistently and grossly overestimate would change the relevant functions. The data are primarily interesting in showing the array sizes for which subjects give correct average responses.

cern is not primarily with a comparison of different presentation times, the major conclusion from these data is that the 200-msec condition provides a good replication of the traditional subitizing results and stable data for experimental manipulation.

Figure 4 shows the data for the 100-msec condition and for practiced and unpracticed subjects at the marginal 50-msec presentation time. These data are compared with the 200-msec baseline data. The 100-msec data show an irregularly increasing function for reaction times, primarily the result of combining data from subjects who performed the task comparably to subjects in the 200-msec condition with data from subjects who essentially performed like the 50-msec presentation time subjects. The 50-msec data show a slowly increasing function with array size, with no apparent abrupt changes in the slope

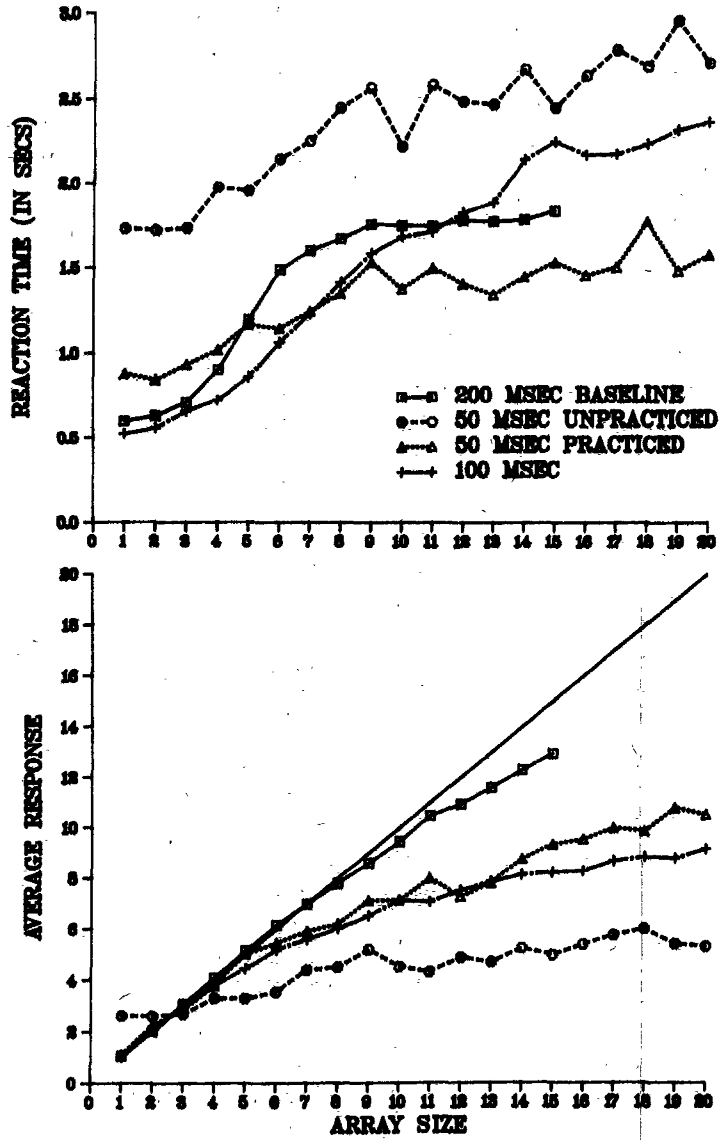

Figure 4. Reaction time and average number response data for the 100-msec presentation time and for practiced and unpracticed subjects with 50-msec presentation times, together with the 200-msec baseline data from Experiment 1.

of the function. If subjects are given prior practice on the task, the practiced 50-msec condition shows the same general function but overall a much faster response time. The bottom panel shows the general inability of subjects to make accurate average responses, even within the subitizing range. The mean response for the 100-msec condition diverges from perfect responding at arrays of 5 and deviates widely from the 200-msec baseline condition. The unpracticed 50-msec condition appears to produce primarily guessing behavior, and the practiced 50-msec condition looks much like the 100-msec condition.

We conclude that a 200-msec presentation time provides stable data that are comparable to previous findings. In addition, presentation times of 100 msec and less produce noisy data and often show no subitizing effects within the boundary conditions of our experimental situation. Finally, the similarity of the reaction time function within the subitizing range for exposure times ranging from 200 to 800 msec suggests some generality of the processes used to estimate numerosity within the 1–6 array size range.

What is the nature of the task for the subject in the subitizing range? Is it possible to make judgments about the countability of an array relatively quickly and accurately? In other words, can and do subjects make a judgment that a particular array can be handled by available processes? If that is the case, then individuals ought to be able to make judgments about the array without actually attempting to give a numerosity response. We address this question in the next experiment.

Experiment 2

This study was concerned with establishing people's ability to perceive an array as cognitively relevant to a particular set of operations. If arrays of 6 or fewer are countable after very short exposure times, the perception of the array should give rise to a particular state of knowledge. People should be able to make a quick response to the array, establishing one of two states: a judgment that the array is "countable," or a judgment that the mental counting process cannot be invoked—that the array cannot be held in consciousness. We assume that if the state is one of countability, this is followed by the numerosity response if a canonical pattern is presented and by counting if it is not. The judgmental process requires the cognition (rather than the perception) of the array, even though such a cognitive state follows a perceptual experience. The argument that an assessment of the array is not merely a response to "perceptual" variables is supported by the fact that the subitizing effect has been found in a wide variety of visual conditions. Consider, for example, the contrast between our conditions and those of Kaufman et al. (1949). In our experiment, subjects placed their heads in a chin rest, they were given a fixation point, the array was presented foveally, and an effective mask was used. Kaufman et al. conducted their experiment in an auditorium. The subject sat facing a screen 9.75 m away on which a space 2.59 m wide was marked off, the array appeared in the "approximate center of this space," and no mask was used. And yet these two conditions produced essentially the same data.

In the experiment to follow, subjects were asked to make a judgment whether, if required, they could give an accurate assessment of the number of items presented.

Method

There were three conditions in this experiment, two at 200-msec exposure time and one at 800 msec. The design and procedure were essentially identical to those of Experiment 1, with the major exception of the instructions to the subjects. We also restricted the range of display sizes to 1–15.

For the two 200-msec conditions, half of the 10 subjects were given four blocks (of 150 trials each) on the numerosity response task, followed by a new countability judgment task; the other 5 subjects started with the new procedure and were given the standard task second. Another 5 subjects were given the new task with a presentation time of 800 msec, with no prior or subsequent numerosity task. All subjects were given a 150-trial practice session for each task.

The new instructions told the subjects to respond "whether or not you could tell the total number of items presented if you were asked to do so. If you feel that you could . . . press the button labelled YES, otherwise press the button labelled NO. Remember that your task is *not* to actually determine how many items are being presented." Additional instructions elaborated slightly on this verbatim version.

Results and Discussion

Subjects' indication that they could, if requested, give the number of dots in the array is similar to some old data of Taves (1941). He presented arrays of 2 to 180 for 200 msec and requested both a number judgment and a confidence rating. Both accurate reports and high confidence ratings were found up to arrays of 6 and then deteriorated rapidly. The yes judgment in this study can, of course, be seen as equivalent to a high confidence rating. Figure 5 shows both the reaction times and proportion of no responses under the three conditions: 200 msec without prior experience (naive), 200 msec with prior experience on the numerosity task (practiced), and 800 msec without prior experience. The reaction times for the two 200-msec conditions produce essentially parallel functions, with the practiced subjects some 50 to 100 msec faster. The function to the left of the break is for yes responses, and the one to the right is for no responses. The break was arbitrarily set at the 50% mark, that is, when half or more of the responses were yes responses, we show the reaction time for yes, and when less than half are yes, we show the reaction time for no.

The functions are much flatter than those for the numerosity judgments. Reaction times increase slightly for the yes responses, but the slopes are very shallow; in comparison with the slopes for the numerosity judgments they are practically flat. The slopes are 14 and 19 msec for arrays of 1 to 4 for the naive and practiced conditions, respectively; the comparable slopes for the numerosity judgments are at least five times larger. In general these responses are as fast as the numerosity judgments in Figure 3 for the easiest of the arrays (1–3). The 800-msec condition produces the same reaction times in the 1–5 range of array sizes and then rises rapidly. However, these average data hide a more interesting finding.

Every one of the five subjects showed an unusually large increase in reaction times within the 5 to 8 range. In each case this rise was well outside the distribution of increments for the other arrays. In Figure 6 we show this finding by plotting reaction times for each of the subjects backward and forward from the array with the largest "jump" (the J array, i.e., the array that showed the largest increment from the preceding array size). If we now inspect the proportion of no responses that are also plotted backward and forward from the J array, we note a similar increase in the percentage of no responses. However, this increase occurs with the array size that is the next larger array to the J array. These data suggest that when the array size first exceeds attentional capacity, subjects respond to the uncertainty generated by this array by spending more time on processing it. When such additional temporal attention does not provide any clearer perception of the array, the result is the decision that they cannot, in fact, give an accurate response to the numerosity question—and the proportion of no responses increases.

The bottom half of Figure 5 shows the proportion of no responses for the three groups as a function of array size. The 800-msec and the 200-msec practiced groups show parallel functions, suggesting primarily a shift in response criterion, whereas the 200-msec naive group falls between the other two. Another way of describing these data, given the instructions to the subjects, is in terms of the degree of confidence (or optimism) that the subjects might give to a numerosity judgment. The least confident are the experienced subjects who show not only the highest proportions of no responses but also a sharp discontinuity between arrays of 4 and 5, which suggests that their experience with the numerosity task has informed them that they can be accurate for the lower numbers but that from 5 onward they are more likely to fail. However, regardless of experience, 50% or more of all judgments are no responses once the array size exceeds 7.

We noted earlier that Atkinson et al. (1976) obtained manual reaction times together with verbal numerosity judgments. The manual response of their subjects may well have been an inadvertent precursor of the method used in this experiment. Subjects may have released the response key when the array was clearly countable, with the verbal response occurring sometime later. Such a conjecture would not only explain the flat reaction times they obtained but also provide

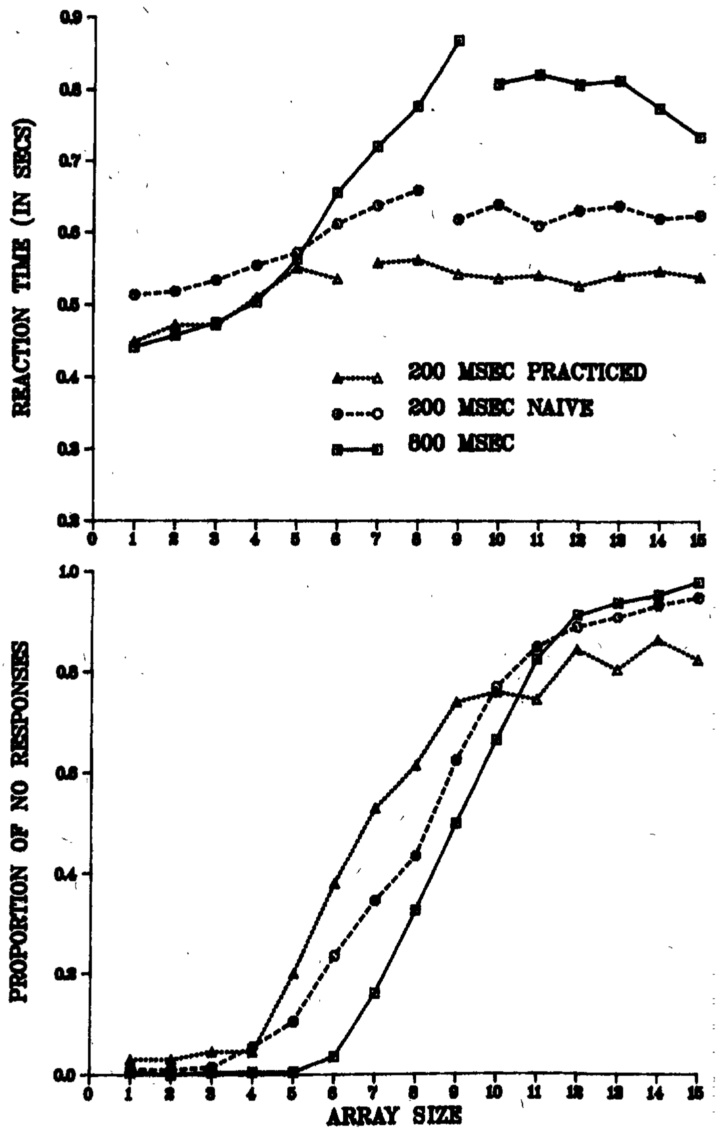

Figure 5. Reaction times and proportion of no responses in Experiment 2. (The reaction times are shown for yes responses to the left of the break in the function and for no responses to the right of the break. The break was determined as the point where yes responses constituted less than 50% of all responses. "Naive" refers to subjects who had no prior experience with the numerosity task; "practiced" subjects had been given the numerosity task previously.)

additional evidence for the argument presented here. It is also interesting that in one experiment in which subjects were required to respond with numerical responses for arrays of 1 to 4 and with "more than 4" after that, reaction times were flat at about 350–400 msec for all array sizes from 1 to 14 except for size 5, which produced a mean reaction time of more than 500 msec.

The results of Experiment 2 provide a positive answer to the question about the cognition of the arrays. Subjects can and do make relatively quick judgments about the "countability" of the arrays. These judg-

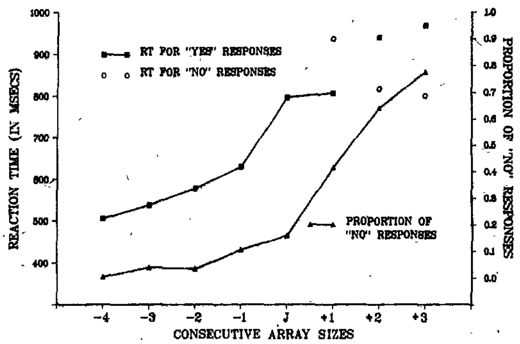

Figure 6. Reaction times and proportion of no responses for the subjects in the 800-msec condition of Experiment 2. (The data are plotted by determining for each subject the J array, the array at which the largest increase in reaction times occurred. Smaller and larger arrays were then plotted backward and forward for each subject and averaged over subjects. The data on proportion of no responses are also plotted in this fashion, using the J definition supplied by the reaction time data. The open circles show reaction times for the no responses beginning with the array at which the no responses represent more than 40% of all responses.)

ments parallel the actual numerosity response of Experiment 1. The range of yes responses to the question of countability is the same as the range of accurate numerosity responses. These countability judgments are as fast as the numerosity responses to the smallest arrays. In other words, in making the countability judgments, subjects do not count the number of dots, they respond to the whole array of several dots as quickly as they respond to "canonical" displays of one or two dots with a number judgment. It might be argued that in this task subjects respond to the arrays with a yes response if there are few dots, and with a no response otherwise. Such an argument begs the question, since it is exactly the distinction between "small" and "large" arrays that lies at the basis of the countability judgment. Furthermore, subjects' (and experimenters') reports indicate that the arrays within the subitizing range convey an intuitive sense of clarity; they "look countable," whereas the larger arrays do not.

In the countability task, prior experience with the *number* judgment task decreases reaction times and confidence about being able to make the number judgments. This also suggests that subjects are reacting to the countability of the arrays. Additional display time for the arrays does increase subjects' confidence that they will be able to make the numerosity judgments, though it does not in fact increase their ability to do so.

Experiment 3

One of the factors that might determine whether subjects engage in canonical pattern perception, linear counting of a mental representation, or estimating of the array is the range of array sizes. If the presentation rate is experimenter-determined and if the range is relatively large (e.g., from 1 to 20), then the likelihood that any particular array will make it possible for the individual to respond by mental counting or perception of some canonical form of twoness or threeness is obviously low. If the range is small (e.g., from 1 to 6), then most if not all of the arrays will permit mental counting as well as canonical pattern responses. Thus, a smaller range of array sizes would encourage a strategy that produces a greater readiness for counting and canonical response. In the first part of Experiment 3, the array range was restricted to 1–6 with the expectation that the counting strategy would be encouraged and that subjects would be more accurate and faster, at least for arrays of 4 to 6. Subjects would be more accurate because they would be likely to use counting rather than estimating for these arrays, and they would be faster because the necessity of deciding whether to count or estimate would be reduced.

Experiment 3A

Method

Five subjects were each given 480 trials, 80 for each array size. The array size varied from 1 to 6. Display duration was 200 msec. Stimuli were generated as in Experiment 1, that is, both pure and mixed arrays were used.

Results and Discussion

Figure 7 shows both the reaction time and accuracy results for the restricted array

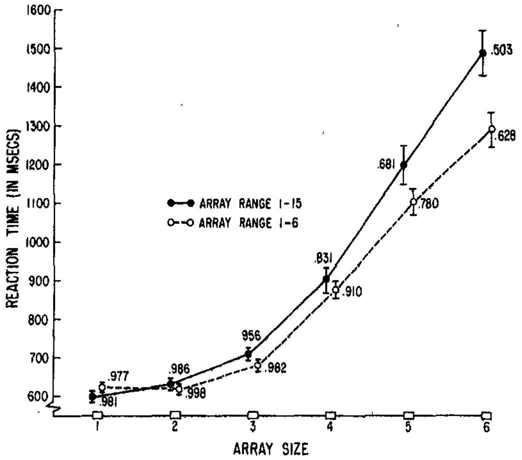

Figure 7. Reaction times for Experiment 3A, in which only arrays of 1 to 6 were presented, compared with the same range of array sizes from the 200-msec baseline condition of Experiment 1. (The proportion next to each data point shows probability of correct numerosity response. The limit bars indicate the 95% confidence interval.)

range, as well as the reference data for the 200-msec condition from Experiment 1. The slopes for arrays of 4 to 6 diverge, and the probability of accurate responses increases by .081, .099, and .125 for array sizes 4, 5, and 6, respectively. The reaction time differences at array sizes 5 and 6 are statistically significant. Thus, the prediction that restricting array ranges will improve accuracy and reduce reaction times holds for the counting range. At the 1–3 range there appears to be some nonsignificant improvement.

If the reduction of the array range to 1–6 increases the probability that counting and canonical strategies will predominate, then a further reduction of the array range to 1–3 should produce the predicted effect of greater accuracy and shorter reaction times for those arrays. This was done in Experiment 3B.

Experiment 3B

Method

Ten subjects were each given two blocks of 90 trials, preceded by 15 trials of practice. The array size varied from 1 to 3 with 30 arrays of each size per block. Stimuli were generated as in Experiment 1, but only pure cases (O stimuli only) were used.

Results and Discussion

Figure 8 shows the comparison of the data from the present experiment with the 200-msec reference condition from Experiment 1. Only the pure case presentations were used from Experiment 1. As in Experiment 3A the predicted result emerged. Both accuracy and reaction times improve with the restricted array range. Array sizes 2 and 3 show an improvement of .018 and .029, respectively, in the probability of a correct response, and there is a significant decrease in reaction time for array size 3.

Manipulation of the range of array sizes in these two experiments provides evidence for the processes that we have postulated to operate in the subitizing range. It is assumed that mental counting is more likely when no or few arrays in the experiment require estimating processes, resulting in faster and more accurate responding. When the range encourages only canonical responding, accuracy increases even more, and reaction times approach the shallow slope of the simple yes/no response in Experiment 2. The latter data suggest that the canonical re-

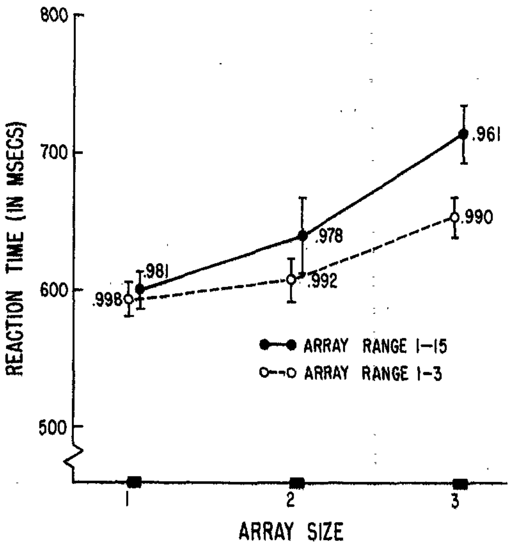

Figure 8. Reaction times for Experiment 3B, in which only arrays of 1 to 3 were presented, compared with the same range of array sizes from the 200-msec baseline condition of Experiment 1. (The proportion next to each data point shows probability of correct numerosity response. The limit bars indicate the 95% confidence interval.)

sponse is made as quickly as the more global response that merely indicates cognitive clarity of the display.

Although these data are consistent with our theoretical arguments, they are also consistent with the well known empirical generalization that a reduction in response alternatives also reduces reaction times. We do not wish to pit these two explanations against each other but rather to present these data as consistent with the theoretical processes we have postulated.

In this and the preceding experiments all the arrays were randomly generated. For array sizes 1 and 2 such a process does, of course, guarantee a "canonical" pattern of points (unity) and straight lines (pairs). However, for arrays of 3 or more, no single canonical pattern is guaranteed. Even for arrays of 3 the random generation will sometimes produce patterns clearly deviant from a canonical triangle or an obvious triple. Consistent responding in the canonical mode is possible only if the same canonical pattern is assured for every array that is displayed. We produced such a condition in Experiment 4.

Experiment 4

We now arrive at a central argument of our studies, the possibility that people can respond quickly and accurately to the canonical form or pattern of arrays, particularly in the lower range of array sizes. Subjects were given both random displays, as in the previous experiments, and canonical displays, that is, displays in which the pattern of any given array size was consistent.

Method

Design and stimulus generation. Display sizes in this experiment were restricted to the 1–10 range. For half of the displays the arrangement was random (as in Experiment 1), for the other half the display was an instance of the canonical form. Choosing particular canonical forms is somewhat problematical, particularly for displays of 5 or more. Common experience suggests fairly unequivocal choices for displays of 3 and 4 (i.e., an equilateral triangle and a square). For sizes 5 to 10 we chose, after some informal testing, the patterns shown in Figure 9. There do not seem to be unequivocal patterns for these numbers of events in our culture; patterns are needed that at least do not contradict cultural expectations, that are symmetrical, and that may be

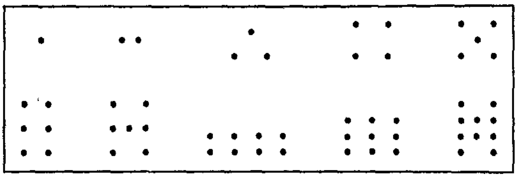

Figure 9. The 10 canonical patterns used in Experiment 4.

discriminated from one another with relative ease. Since subjects are expected to acquire the relation between a particular canonical form and its numerosity, the justification of these particular patterns must be found in the data they generate. For each canonical pattern five trials were generated. These instances were constant in form, but varied in size and position in the 8 × 8 matrix and only Os were used. The five instances were selected by first generating all possible combinations of pattern size and position and then randomly selecting five of these instances. Displays in either the canonical or random condition in which two adjacent spaces of the underlying matrix were occupied by Os were eliminated. Each array size consisted of five different randomly generated patterns. A block consisted of 100 trials, 50 random patterns and 50 canonical patterns, with 10 instances (5 random and 5 canonical) of each of the 10 array sizes.

Procedure. Five subjects continued with this task until their performance over three consecutive blocks showed no improvement in reaction time. As a result, one subject completed 9 blocks, one 12, two 14, and one 21. Each experimental session took approximately 1 hr., with two or three 100-trial blocks. There were no practice trials. The instructions were identical to those in Experiment 1, and subjects were not informed about the presence of random and canonical patterns.

Results and Discussion

The data for Experiment 4 are shown in Figure 10 for both the random and the canonical patterns, segregated by successive blocks of 50 trials. It is evident that reaction times for the canonical patterns are much faster than for the random patterns, and in neither group does there seem to be much improvement after the first 50 trials. More important, the reaction times are flat for canonical patterns of 1 to 5 by the middle 50 trials. All the responses take about 600 msec, and there is no difference between the response to the canonical patterns for array sizes 1 to 3 and the random display for these same sizes. Similarly, performance is errorless by the middle 50 trials for canonical patterns of 1 to 5.

Figure 11 shows the mean reaction times for the first five trials of canonical and ran-

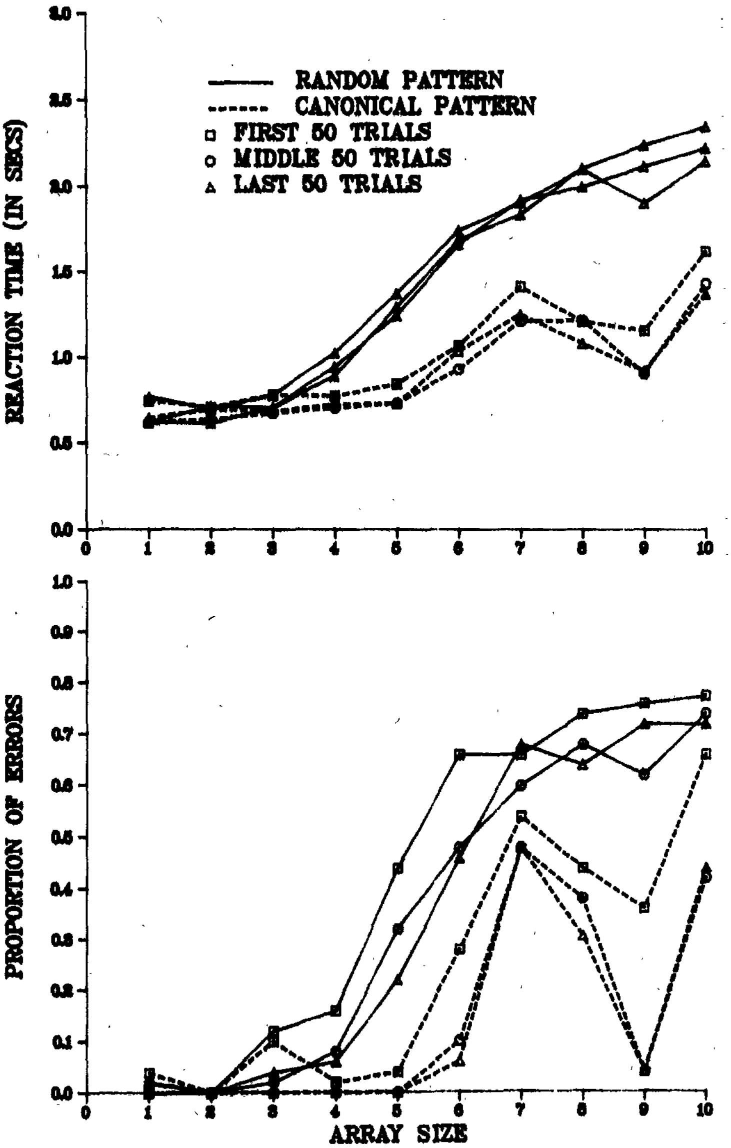

Figure 10. Mean reaction times and proportion of errors for responses to canonical and random patterns in Experiment 4, segregated by first, middle, and last 50 trials.

dom patterns. Improvement occurs very early in the task; within five presentations subjects show superior performance (fewer errors and faster reaction times) starting with array size 4.

It is important to note that the response speed to the canonical patterns of 3 to 5 is identical to the response speed to random displays of 1 and 2 in the noncanonical experiments. Thus, canonical patterns take the same time for processing as do naturally occurring canonical patterns (displays of 1 and 2, and sometimes 3, events). The data on canonical patterns for arrays of 6 to 10 largely reflect the problem of generating perceptually distinct patterns for those arrays. This is particularly evident for the canonical patterns for 7 and 10 (cf. Figure 9), which subjects repeatedly confused.

There is one finding about the random arrays that needs mentioning. In contrast to all the previous experiments (see Figures 1,

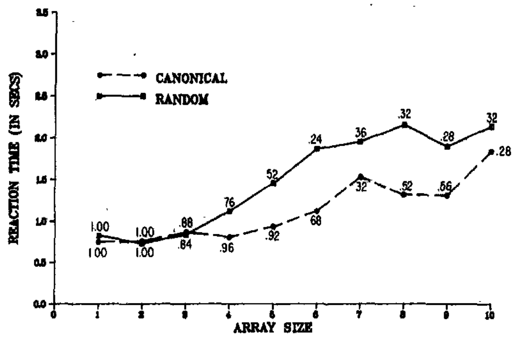

Figure 11. Mean reaction times for the first five trials of each array size for both random and canonical displays in Experiment 4. (Numbers next to data points show probability of correct numerical response.)

3, 7, and 8) there is not even a slight increase in reaction time for arrays of 3. This must be due to the context of the canonical patterns, which might well induce a more systematic approach to triplets than is the case when only random arrays are present. As a result, practically any configuration of a triplet is recognized and accepted as a 3.

The data of this experiment are consistent with the notion that number estimation based on a perceptual pattern mechanism can be acquired by adults (as well as children). This is the claim made by Gelman and Gallistel (1978) for the developmental data and presented here for the adult performance. Children, particularly very young children, do not have available the canonical representation of numberness. Adults have acquired the concept of oneness, twoness, and threeness associated with single events, straight lines (pairs), and triangles. As a result, children's slope for the 3–10 range represents straightforward counting (cf. Chi & Klahr, 1975), but at the age of 6 they already have some notion of the twoness concept of the pair and therefore show the shallow slope for the 1–2 range. Very young children still count in the 2–3 range (Gelman & Tucker, 1975), and adults apparently also count in the 4+ range when random presentations of dots do not provide an adequate canonical representation of fourness, fiveness, and so on.

Experiment 5

If we are correct in asserting that subjects use canonical patterns of oneness, twoness, and threeness for arrays of 1 to 3 and engage in mental counting for larger arrays, then the processes of "counting" should be continuous from arrays of 4 through those well beyond the subitizing range. Arrays of 4 to 20, for example, should show a single linear slope under conditions in which the counting process is "environmental" rather than "mental."

Method

Design. The experimental design was identical to that used in Experiment 1, with arrays of 1 to 20 and all conditions the same. The only difference was that display time was controlled by the subject rather than the experimenter.

Procedure. Five subjects participated in this experiment. Each subject ran through 10 blocks of 80 trials each. Each block contained 4 trials for each display size (1–20), with 2 of the 4 trials pure (all Xs or Os) and the other 2 trials mixed. There were six 1-hr. sessions; the first was a practice session, the other five each contained two blocks of 80 trials. Pure and mixed trials were selected and assigned as in Experiment 1. Ready signals and dot fixation points were the same as in Experiment 1. The display terminated as soon as the subject responded and was replaced by a random dot masker that lasted for 400 msec. The instructions again told the subjects to respond as quickly and accurately as possible.

Results and Discussion

Figure 12 shows the results for Experiment 5 compared with the 200-msec reference condition from Experiment 1. These data are consistent with the data from Jensen et al. (1950) shown in Figure 1. The data show a counting procedure, with a straight line slope of 382 msec for arrays of 4 to 20. The response to arrays of 1 to 3 is clearly not part of the same counting mechanism; subjects here, as well as in the restricted exposure experiments, respond quickly to the canonical pattern of these small arrays. The proportion of errors shows a slight slope with increasing array sizes, but even with arrays as large as 20, over 80% of the numerical responses are correct.

General Discussion

The important challenge posed by Kaufman et al. (1949) some 30 years ago was to find some reasonable interpretation of the subitizing function: What happens during the determination of numerosity for arrays

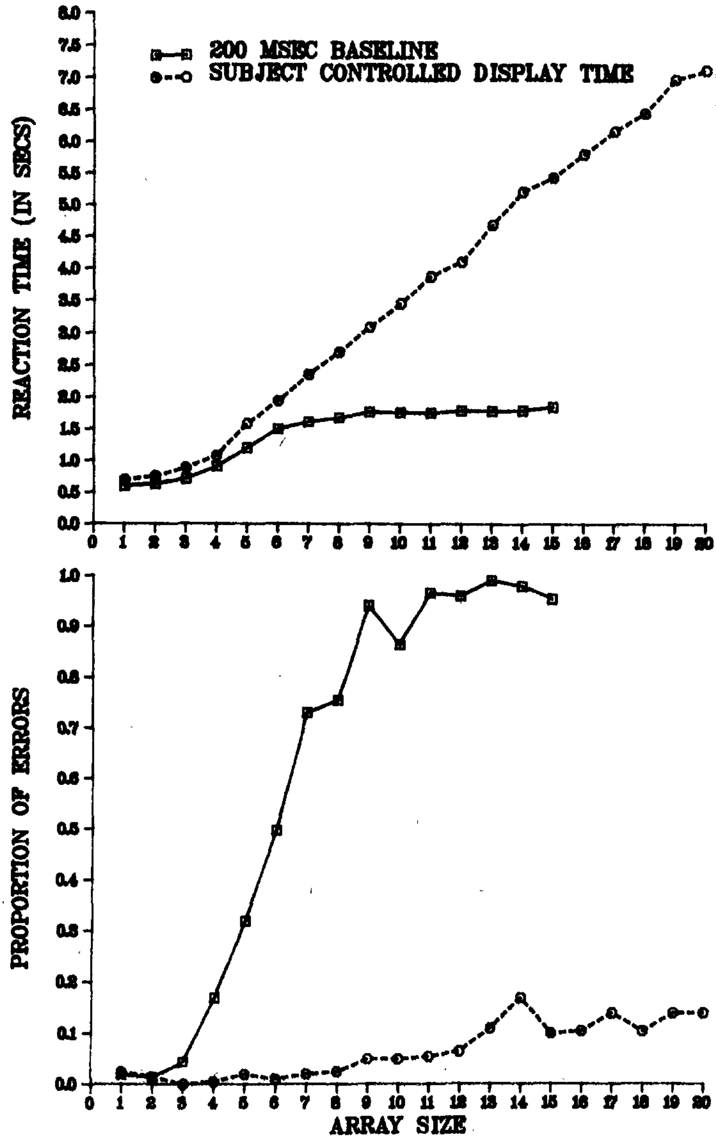

Figure 12. Reaction time and proportion of errors for the subject-determined presentation time data from Experiment 5. (For comparison, the data from the 200-msec baseline experiment are also shown.)

of 6 and fewer? The limitation of human consciousness in the apperception of discrete events has been known at least since the early 18th century (see Hamilton, 1859). However, the very distinctive function relating reaction time to array size that Kaufman et al. discovered, and that has remained remarkably robust over the years, goes beyond that limitation and permits us to ask what kind of processing is likely to characterize numerosity judgments.

We have demonstrated two quite different mechanisms to explain the subitizing phenomenon by dividing the positively accelerated slope of reaction times into a combination of acquired canonical patterns and counting of the array held in consciousness. We ascribe the discontinuity between these two "subitizing" processes and the estimating used for larger arrays to the more general limitation of human attention or consciousness. Only up to 6 or 7 discrete events can be held in consciousness and counted. Larger arrays cannot be so maintained and the accuracy of counting suffers; estimation is substituted for it.

The shallow, or often flat, slope of reaction times for arrays of 1 to 3 items seems to be explained by acquired canonical patterns. As we noted earlier, a similar suggestion was made by Woodworth and Schlosberg in 1954. The fact that canonical patterns for larger arrays could be acquired by our subjects and were responded to as quickly and as accurately as the smaller arrays supports this argument.

Our results also support in large part Warren's (1897) conclusions, discussed earlier. However, we have suggested that instead of perceptive counting and inferential counting being available in addition to progressive (incremental) counting, the apparent perceptual apprehension of numerosity is probably also a case of inference. Patterns of small numbers are familiar and responded to "inferentially"; people "know" that a triangular pattern represents 3 and respond accordingly.

The notion of "mental counting," the application of a simple counting mechanism to a representation of the array that is held in consciousness, deserves more detailed examination. If the counting mechanism operates on the conscious array in the same way that it operates on the physically presented array, then the slope for reaction times in arriving at numerosity judgments should be similar under the two conditions.

The slopes, intercepts, and variance accounted for are shown in Table 1 for several different experiments. The major comparisons we are interested in concern the slopes for arrays of 4 to 6 under "subitizing" conditions (brief array exposures) and under unlimited time conditions (when subjects can count the physically present array). For the Kaufman et al. (1949) and Jensen et al. (1950) experiments the two slopes differ less than 10%. However, it is difficult to arrive at a definitive conclusion because these two experiments compare reaction times for primarily correct responses (in Jensen et al.) and a mixture of correct and incorrect responses (in Kaufman et al.). The next row in Table 1 (Experiment 1) shows the slope for our 200-msec reference condition, which seems to be comparable to the previous data. However, the slope for the unlimited time condition in Experiment 5 is significantly larger. As we have just indicated, most responses under these conditions are correct responses, which is not true of the 200-msec reference condition. The latter includes incorrect fast responses, reflecting the usual speed/accuracy trade-off. The last row of the table shows the slope for the 200-msec reference condition but only for the correct responses. Under these conditions the slope for the arrays of 4 to 6 increases from 296 to 347 msec. The latter value is within 10% of the slope value for Experiment 5 (382 msec). It seems reasonable to conclude that—in the context of our experimental conditions—subjects count items either mentally or environmentally at a rate of about 350 to 380 msec per item.

We noted earlier that the counting slope for young children is two to three times larger than the slope for adults, and that the difference is probably due to the fact that adults learn to count by twos and threes, rather than by ones. The postulation of canonical patterns for arrays of 1, 2, and 3 suggests a source for the adult counting strategy that also is phenomenally and intuitively apposite. It is also consistent with

Table 1
Linear Regression Analysis for Selected Experiments and Set Sizes

Experiment	Set sizes	Slope	Intercept	Variance accounted for
Kaufman et al. (1949)	4–6	270	−190	.984
Jensen et al. (1950)	4–15	290	−289	.992
Experiment 1	4–6	296	−282	.999
Experiment 5	4–20	382	−348	.998
Experiment 1 (correct responses only)	4–6	347	−550	.990

Beckwith and Restle's (1966) conclusion that "in counting, the set of objects is grouped perceptually" (p. 443). However, if arrays of 6, for example, are counted by the use of doublets and triplets, then the limitation of consciousness is overestimated by a factor of two or three. In fact, one would have to conclude that only three chunks are held in consciousness, a conclusion also suggested by Broadbent (1973). This particular speculation must be left open, since the limitation to three events would be difficult in light of other data (see Mandler, 1975a; Miller, 1956).

The limitation to some 6 visual events that can be accommodated in a single attentional episode has also been supported by some recent experiments of Bartram (1978). In Bartram's experiments subjects saw randomly placed disks on a 4 × 5 grid and were later asked to reproduce the pattern. Bartram concludes that "spatial discontinuities in the distribution of attention . . . are regarded as 'defining' chunks of stimulus elements" (p. 324). These perceptual chunks showed that "seven disks provide an upper limit on the chunk size" (p. 350). Bartram also suggests that studies that have concluded that individuals take in only one chunk per glance are characterized by strategies that take advantage of previously known, familiar, and meaningful patterns (see, e.g., Chase & Simon, 1973; Reitman, 1976). The ability of subjects in our Experiment 4 to accommodate more than 6 items in canonical patterns is clearly an instance of such a strategy.

The model presented here also sheds light on some cross-cultural data on numerosity judgments in the subitizing range presented by Cole, Gay, and Glick (1968). Comparing results obtained with American subjects and subjects from the Kpelle tribe of Liberia, these authors found no difference in the probability of correct responses for arrays of 3 but superior performance by the American subjects beyond that array size. Patterned arrays produced greater improvement for the Americans than for the Kpelle. Since the longest presentation time in these experiments was 100 msec and no data were collected on arrays of 1 and 2, no definite conclusion can be reached, but it is likely that the difference between the two groups is one of counting strategies. American subjects count by twos and threes, and the Kpelle do not. As a result, the American subjects are better able to process the larger array sizes. These differences in cultural approaches to counting are not relevant to the small array sizes (1–3).

Some of the problems that can now be approached within this experimental paradigm concern the nature of the counting process in particular and the use of canonical patterns in numerosity judgments in general. If, as seems likely, adults first develop simple canonical perceptions for twoness and threeness and then apply these schemas to the counting of large arrays, it is of some interest to investigate further the conditions under which such counting strategies are used. Not only is it possible to structure arrays so as to facilitate such strategies, but it should also be possible to teach the use of larger canonical patterns for even more efficient approaches to the numerosity task. Whether people can use such patterns as chunks that are processed in parallel is of both practical and theoretical interest. The theoretical problems to be approached concern the nature of parallel processing on the one hand and a better understanding of the limitation of the attentional/conscious capacity system on the other.

Having started this discussion with a review of counting and numerosity judgments from a developmental point of view, we can also use the results of these experiments to study the growth of canonical perception and of counting. Now that we seem to understand the subitizing strategies of adults, we can ask how these strategies develop in the child.

We have rejected the notion that some automatic apperception of numerosity exists, but we seem to have substituted for it the automatic apperception of countability. The data in Experiment 2 certainly suggest some such conclusion. However, countability is a derivative state; what is more likely is that the individual recognizes that the representation of the array "fits into" the range of events that can be accommodated within the conscious state. If consciousness is seen as the "state of a structure" (Mandler, 1975b), then the limited capacity of consciousness refers to the number of events (or chunks) that may be in that state at any one point in time. "Consciousness" then refers to a mechanism that permits the subsequent operation of certain processes—such as counting, judging, or comparing—on the events that are in the "conscious" state.

Are we dealing with a limited capacity to perceive or visualize some limited number of visual events or with a more general mechanism that limits attention or consciousness to a limited set of events? There is some evidence that the visualization mechanism is of a more general character. As a result of a series of interference studies, Phillips and Christie (1977) concluded that the visualization of simple patterns requires "general purpose resources" (p. 649). The imperviousness of the subitizing phenomenon to a wide variation of "perceptual" variables such as masking, foveal projection, and presentation time also suggests that we are not dealing with a simple visual short-term storage system.

Thus, one of the next steps in the use of the subitizing paradigm will be to study the effect of visual, auditory, and more general cognitive interference tasks on the cognition of numerosity. The distinction between the canonical pattern process and mental counting will make it possible to study interference and, indirectly, the use of general-purpose resources as they affect and are affected by these two processes separately.

Reference Note

1. Atkinson, J. Personal communication, August 1980.

References

Allport, D. A. The state of cognitive psychology. *Quarterly Journal of Experimental Psychology,* 1975, *27,* 141–152.

Atkinson, J., Campbell, F. W., & Francis, M. R. The magic number 4 +/− 0: A new look at visual numerosity judgements. *Perception,* 1976, *5,* 327–334.

Bartram, D. J. Post-iconic visual storage: Chunking in the reproduction of briefly displayed visual patterns. *Cognitive Psychology,* 1978, *10,* 324–355.

Beckwith, M., & Restle, F. Process of enumeration. *Psychological Review,* 1966, *73,* 437–444.

Bourdon, B. Sur les temps nécessaire pour nominer les nombres. *Revue Philosophique,* 1908, *65,* 426–431.

Broadbent, D. E. *In defense of empirical psychology.* London: Methuen, 1973.

Cattell, J. M. Ueber die Trägheit der Netzhaut und des Sehcentrums. *Philosophische Studien,* 1886, *3,* 94–127.

Chase, W. G., & Simon, H. A. The mind's eye in chess. In W. G. Chase (Ed.), *Visual information processing.* New York: Academic Press, 1973.

Chi, M. T. H., & Klahr, D. Span and rate of apprehension in children and adults. *Journal of Experimental Child Psychology,* 1975, *19,* 434–439.

Cole, M., Gay, J., & Glick, J. A cross-cultural investigation of information processing. *International Journal of Psychology,* 1968, *3,* 93–102.

Fernberger, S. W. A preliminary study of the range of visual apprehensions. *American Journal of Psychology,* 1921, *32,* 121–133.

Gelman, R., & Gallistel, C. R. *The child's understanding of number.* Cambridge: Harvard University Press, 1978.

Gelman, R., & Tucker, M. F. Further investigations of the young child's conception of number. *Child Development,* 1975, *46,* 167–175.

Glanville, A. D., & Dallenbach, K. M. The range of attention. *American Journal of Psychology,* 1929, *41,* 207–236.

Graesser, A. C. II, & Mandler, G. Limited processing capacity constrains the storage of unrelated sets of words and retrieval from natural categories. *Journal of Experimental Psychology: Human Learning and Memory,* 1978, *4,* 86–100.

Hamilton, W. *Lectures on metaphysics and logic* (Vol. 1). Edinburgh: Blackwood, 1859.

Jensen, E. M., Reese, E. P., & Reese, T. W. The subitizing and counting of visually presented fields of dots. *Journal of Psychology,* 1950, *30,* 363–392.

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. An information processing approach to the study of cognitive development. In A. D. Pick (Ed.), *Minnesota Symposia on Child Psychology* (Vol. 7). Minneapolis: University of Minnesota Press, 1973. (a)

Klahr, D. A production system for counting, subitizing and adding. In W. G. Chase (Ed.), *Visual information processing.* New York: Academic Press, 1973. (b)

Mandler, G. Organization and memory. In K. W. Spence & J. T. Spence (Eds.), *The psychology of learning and motivation: Advances in research and theory* (Vol. 1). New York: Academic Press, 1967.

Mandler, G. Memory storage and retrieval: Some limits on the reach of attention and consciousness. In P. M. A. Rabbit & S. Dornic (Eds.), *Attention and performance V.* London: Academic Press, 1975. (a)

Mandler, G. *Mind and emotion.* New York: Wiley, 1975. (b)

Miller, G. A. The magical number seven, plus or minus two: Some limits on our capacity for processing information. *Psychological Review,* 1956, *63,* 81–97.

Oberly, H. S. The range for visual attention, cognition and apprehensions. *American Journal of Psychology,* 1924, *35,* 332–352.

Phillips, W. A., & Christie, D. F. M. Interference with visualization. *Quarterly Journal of Experimental Psychology,* 1977, *29,* 637–650.

Reitman, J. Skilled perception in Go: Deducing memory structures from inter-response times. *Cognitive Psychology,* 1976, *8,* 336–356.

Saltzman, I. J., & Garner, W. R. Reaction time as a measure of span of attention. *Journal of Psychology,* 1948, *25,* 227–241.

Starkey, P., & Cooper, R. G., Jr. Perception of numbers by human infants. *Science,* 1980, *210,* 1033–1035.

Svenson, O., & Sjöberg, K. Subitizing and counting processes in young children. *Scandinavian Journal of Psychology,* 1978, *19,* 247–250.

Taves, E. H. Two mechanisms for the perception of visual numerousness. *Archives of Psychology,* 1941, *37,* 1–47.

von Szeliski, V. Relation between the quantity perceived and the time of perception. *Journal of Experimental Psychology,* 1924, *7,* 135–147.

Warren, H. C. Studies from the Princeton Psychological Laboratory, VI–VII (VI. The reaction time of counting). *Psychological Review,* 1897, *6,* 569–591.

Woodworth, R. S., & Schlosberg, H. *Experimental psychology.* New York: Holt, 1954.

Wundt, W. *Grundriss der Psychologie.* Leipzig: Engelmann, 1896.

Received November 7, 1980 ■