

- McClelland, J. L., & Rumelhart, D. E. (1985). Distributed memory and the representation of general and specific information. *Journal of Experimental Psychology: General*, *114*, 159-188.
- Newell, A., & Simon, H. A. (1972). *Human problem solving*. Englewood Cliffs, NJ: Prentice-Hall.
- Nissen, M. J., & Bullemer, P. (1987). Attentional requirements of learning: Evidence from performance measures. *Cognitive Psychology*, *19*, 1-32.
- Reber, A. S. (1967). Implicit learning of artificial grammars. *Journal of Verbal Learning and Verbal Behavior*, *6*, 855-863.
- Reber, A. S. (1989). Implicit learning and tacit knowledge. *Journal of Experimental Psychology: General*, *118*, 219-235.
- Reber, A. S., Allen, R., & Regan, S. (1985). Syntactical learning and judgment, still unconscious and still abstract: Comment on Dulaney, Carlson and Dewey. *Journal of Experimental Psychology: General*, *114*, 17-24, 1985.
- Remington, R. J. (1969). Analysis of sequential effects in choice reaction times. *Journal of Experimental Psychology*, *82*, 250-257.
- Rumelhart, D. E., Hinton, G., & Williams, R. J. (1986). Learning internal representations by error propagation. In D. E. Rumelhart & J. L. McClelland (Eds.), *Parallel distributed processing, 1: Foundations* (pp. 318-362). Cambridge, MA: MIT Press.
- Schacter, D. L. (1987). Implicit memory: History and current status. *Journal of Experimental Psychology: Learning, Memory and Cognition*, *13*, 501-518.
- Servan-Schreiber, D., Cleeremans, A., & McClelland, J. L. (1988). *Encoding sequential structure in simple recurrent networks* (Tech. Rep. CMU-CS-88-183). Department of Computer Science, Carnegie Mellon University, Pittsburgh, PA.
- Servan-Schreiber, E., & Anderson, J. R. (1990). Learning artificial grammars with competitive chunking. *Journal of Experimental Psychology: Learning, Memory and Cognition*, *16*, 592-608.
- Willingham, D. B., Nissen, M. J., & Bullemer, P. (1989). On the development of procedural knowledge. *Journal of Experimental Psychology: Learning, Memory and Cognition*, *15*, 1047-1060.

6 Subitizing: The Preverbal Counting Process

C. R. Gallistel and Rochel Gelman
University of California, Los Angeles

Many account for the capacity of infants, young children, and animals to discriminate small numerosities by an appeal to subitizing (Chi & Klahr, 1975; Davis & Pérusse, 1988; von Glasserfeld, 1982; Klahr & Wallace, 1973; Rumelhart, Savage-Rumbaugh, & Hegel, 1987; Shipley & Shepperson, 1990). Despite the central role attributed to this process, there has never been an explicitly formulated model of it. Kaufman, Lord, Reese, and Volkman (1949) coined the term *subitizing* for the process used by adults to give rapid numerosity judgments for small arrays of simultaneously presented dots.

In the recent literature, the definitive experimental paper on subitizing is by Mandler and Shebo (1982). Their results, which confirm and extend results obtained in experimental investigations dating back to the turn of the century, are the starting point for the model we present here for the subitizing process. Additional sources of our model include the work of Meek and Church (1983), who have developed and experimentally tested a model of the counting mechanism in animals, and research showing that human adult numerical competence rests in part on conversions back and forth between verbal and written representations of numerosity (numerlogs or numerographs) and a preverbal representation of numerosity in which numerosity is represented by statistically ill defined values on a mental number line (Dehaene, 1989; Dehaene, Dupoux, & Mehler, in press; Hinrichs, Yurko, & Hu, 1981; Holyoak, 1978; Holyoak & Mah, 1982).

Our hypothesis is that preverbal representatives of numerosity (preverbal numerons) are magnitudes generated by the counting mechanism proposed by Meek and Church (1983). The magnitude (point on the mental number line) of a numeron generated by this counting process is assumed to vary from one correct count of the same numerosity to the next. Thus, the values assigned to a given

numerosity form a numeron probability density distribution on the mental number line. In accord with the animal counting model and the experimental findings on the representation of numerosity by animals, we assume that the variance of the numeron probability density function is a scalar function of the numerosity: the greater the numerosity, the greater the variance in the distribution of numerons representing that numerosity, hence, the greater the likelihood of confounding a numerosity with adjacent numerosities.

THE REACTION TIME FUNCTION

Mandler and Shebo's data, in common with all the previously published reaction time data, suggest that the process by which adults generate the numerlogs for arrays of one to six items involves a serially incrementing mechanism because there is a sizable and statistically significant increment in the reaction time for each additional item in the array whose numerosity is to be estimated (Figure 6.1). The reaction time data also suggest, however, that the process used in this range differs in some way from the process used with more numerous arrays. This is because the increments in mean reaction time, as array size increases from one to two (30 msec), two to three (80 msec), and three to four (200 msec), are smaller than the 300–325 msec/item increment seen in the *reaction-time method* with arrays greater than 6, where subvocal verbal counting is clearly being used. In the reaction-time method, the display of the array terminates with the subject's response, whereas with the *tachistoscopic method*, the presentation time of the array is fixed, usually at 200 msec. Another way to code the difference between these methods is to consider who controls the offset of the display; the subject (in the reaction-time method) or the experimenter (in the tachistoscopic method).

It has frequently been proposed that subitizing (Davis & Pérusse, 1988; Klahr & Wallace, 1973; Strauss & Curtis, 1984; Woodworth & Schlosberg, 1954) is a process analogous to those by which we perceptually categorize things like cows and trees. Mandler and Shebo did two control experiments that make this idea less plausible. First, they showed that the reaction time to name the arabic numerals "1, 2, 3, and 4" did not increase systematically. There is, of course, no

FIG 6.1. The time to state the numerosity of an array of randomly arranged dots or circles as a function of the numerosity of the array in four different experiments. The panels on the left are from "reaction-time" experiments in which the array is displayed until the subject responds. These panels have the same horizontal and vertical scales. The panels on the right are from tachistoscopic experiments in which the array is only displayed for 200 msec (400 msec less than the time it takes to react to a one-item array). The upper two panels have the same horizontal and vertical scales. A. Data from Jensen, et al. (1950).

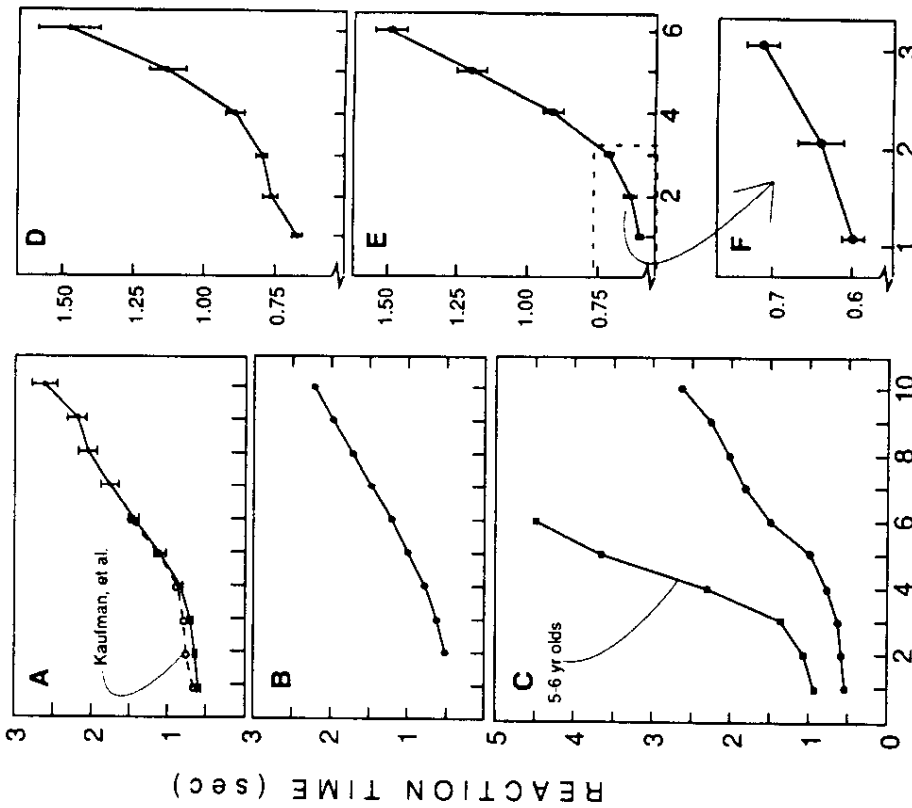


FIG 6.1. (Continued)

Points are means of the median reaction times of the 5 subjects. Vertical bars show ± 1 standard error of the mean. The superimposed open circles connected by a dashed line are the data from the tachistoscopic experiment in Panel D. The asymptotic slope of this function is sustained up to arrays of 30, the largest tested. B. Data from Saltzman and Garner (1948). They did not use arrays of only one. C. Data from Chi and Klahr (1975) for adults (lower curve) and 5-6-year-old children. D. Tachistoscopic data from Kaufman, et al. (1949), using same arrays as Jensen, et al. in their "reaction-time" experiment. Vertical bars are ± 1 standard error of the mean ($n = 5$). E. Data from Mandler and Shebo's (1982) tachistoscopic study. Vertical bars show 95% confidence interval. F. Larger scale display of Mandler and Shebo's data for the range 1-3. Vertical bars show 95% confidence levels.

reason why it should, because the recognition of these numerals is presumably mediated by a pattern-perception process. If the recognition of the oneness, twoness, threeness and fourness of random arrays were also mediated by a pattern perception process, then there would also be no reason to expect the systematic and very sizable increments in reaction time seen in the range from 1 to 4 (Figure 6.1, panels D and E). Sizable increments are seen over this range even in highly practiced subjects and even when the only arrays being shown are arrays in the range from one to three (see Mandler & Shebo, 1982, Figure 8). It has consistently been found that it takes practiced adult subjects about 300 msec longer to respond "four" to an array of four than to respond "one" to an array of one. This increment in reaction time within the subitizing range is equal to half the total latency to respond to a 1-item array (Jensen, Reese, & Reese, 1950; Kaufman et al., 1949; Mandler & Shebo, 1982).

Mandler and Shebo (1982) also showed that when they used 'canonical' arrays in which the arrangement of the dots composing a given numerosity was always the same, thereby making a strategy based on pattern recognition plausible, the reaction time function was flat over the range from one to five. Again, it is not surprising that the reaction time to recognize and numerically name distinctive spatial patterns does not increase systematically as a function of the numerosity of the dots composing the patterns, but this finding underscores the significance of the finding that when the spatial relations among the dots composing an array vary randomly from trial to trial, the reaction time function does increase systematically as a function of the number of dots.¹ It implies that the process whereby the subjects generate the numerals corresponding to these randomly arrayed sets is not a pattern-recognition process. The pattern recognition hypothesis for subitizing in infants has also been rejected by Cooper (1984) because the infants discriminate the numerosity of linear arrays of two and three.

It is often claimed that there is a discontinuity in the reaction-time data or in their slope somewhere in the range of numerosities between 3 and 8. Where in this range the discontinuity is imagined to be located has varied substantially from author to author. In the data from the tachistoscopic method (used by Kaufman, et al., 1949, and Mandler & Shebo, 1982, among others), the discontinuity in the slope of the reaction function is clear, as is its probable origin. In both of these tachistoscopic studies (where the display time is set at a constant

value for varying set sizes), the reaction-time function accelerates steadily as the numerosity of the arrays increases from 1 to 6 (each successive unit increment in numerosity produces a greater increment in reaction time) and then it levels off abruptly to a reaction time of about 1.5 sec for all numerosities greater than 6 (even though, interestingly, the number reported increases systematically with numerosity up to numerosities at least as large as 200). When the display-offset time is controlled by the subject in the reaction-time method, there is the same acceleration over the range from one to six (Figure 6.1, panels A-C), but there is no leveling off. The leveling off of the reaction-time function in the tachistoscopic paradigm is probably due to fading of the iconic image of the array, which puts an end to the operation of both verbal and preverbal serially incrementing (counting) processes. The use of fixed-time tachistoscopic presentations forces this leveling off. It is presumably of no significance for models of the process by which the mind generates verbal number judgments for arrays of six or less.

For the range 1-6, the reaction-time method and the tachistoscopic method yield superimposable reaction-time functions (Figure 6.1, panel A). This suggests that the processes that generate the appropriate numeral are independent of the mode of presentation of the array over this range. There is an appreciable decline in absolute accuracy over this range. The Mandler and Shebo (1982) subjects gave the correct number 97.7% of the time on trials with an array of numerosity one but only 50.3% of the time on trials with numerosity six.

When the display of the array is sustained until the subject responds, there is no discontinuity in the slope of the reaction-time function (Figure 6.1, panels A-C). Chi and Klahr's (1975) efforts to discover a discontinuity by fitting straight lines to different subsets of the data has been trenchantly criticized by Allport (1975). Jensen et al. (1950) show more sensitivity to the problem of determining whether there is a discontinuity, and, if so, where it is. Although they claim to find a slope discontinuity between 6 and 7, what the data from all the studies show is a slope that steadily increases up to an asymptotic value somewhat greater than 300 msec/item (Figure 6.1, panels D and E). The steadiness of this acceleration is most apparent in the Mandler and Shebo data, which have the smallest uncertainty (Figure 6.1, panels E and F, vertical bars give 95% confidence intervals). The slope is within 10% of its asymptotic value when the numerosity exceeds 4.

THE ANIMAL COUNTING MECHANISM

The common laboratory animals (rats and pigeons) discriminate simultaneously and successively presented sets (of diverse composition) on the basis of their numerosity, for numerosities at least as high as 50 (Gallistel, 1990, Chapter 10). The upper limit on the animal capacity to discriminate numerosities has not been experimentally probed. Animals discriminate between the numerosities of simul-

¹The data plotted in Figures 10 & 11 of Mandler and Shebo (1982) come from the experiment where subjects were shown both random and canonical dot patterns. In this case the RT function for the random displays of 1, 2, and 3 is flat. Some might conclude that this result is inconsistent with the remaining data in Mandler and Shebo, their explanations, and our review of the literature. When we asked Mandler if this was the case, he noted that these data come from an experiment where subjects were told to look for canonical patterns and where a within subjects' design was used. He suggests that subjects took time to assimilate non-canonical displays in this range to their canonical representations of 2 and 3, a proposal, he noted, that explains why the RT function is elevated in this particular study. (Mandler, personal communication, April, 1990.)

taneously and sequentially presented stimuli or responses regardless of how widely separated the items are in time and space. Given the tempo-spatial range of stimulus presentations that support discriminations based on numerosity and the large numerosities that may be discriminated (e.g., 45 pecks versus 50 pecks—Rilling & McDiarmid, 1965; Rilling, 1967), it is difficult to envision any mechanism for deriving the mental representatives of numerosity (numerons) that underlie these discriminations other than a mechanism that sequentially passes through an ordered series of states, the last of which represents the cardinal numerosity of the set. Animals count sets of heterogeneous items as readily as sets of homogeneous items and they transfer the discrimination immediately to sets composed of stimuli that were not in the training set (Capaldi & Miller, 1988; Fernandes & Church, 1982). They also appear to perform addition and subtraction operations with representatives of numerosity (numerons), including numerals (Boysen & Berntson, 1989).

A finding from this literature that is important for our model of the subitizing process is the imprecision with which animals represent even small numerosities. Figure 6.2 gives data from a number-discrimination task in which the animal had to press a lever some fixed number of times in order silently to arm a food-dispenser. When the dispenser was armed and the rat's head interrupted a beam in front of the feeding cup, a pellet was dispensed. If, however, the rat interrupted the beam before the requisite number of presses had been made, it incurred a penalty whose severity was one of the experimental variables. Figure 6.2 plots the probability of the rat's interrupting a sequence of lever presses to try the feeder as a function of the number of presses it had made since the last interruption, when the penalty for premature interruptions was a 10-second time out. The between-curve variable is the number of presses required to arm the

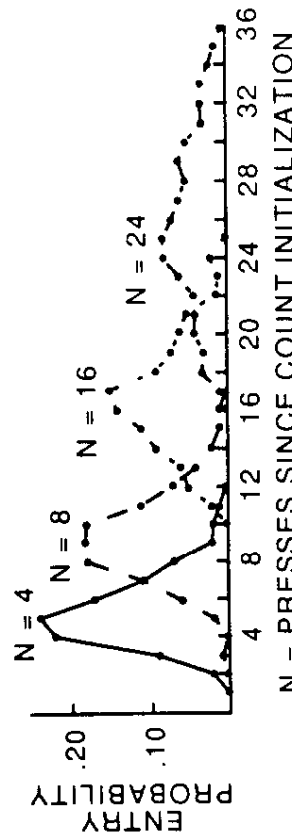


FIG 6.2. The probability of breaking off to enter the food delivery area as a function of n , the number of presses made since the initialization of the response counter, for various values of N , the required number of presses, under conditions where a premature entry triggers a 10-second time-out but does not reset the response counter. (Redrawn from Platt & Johnson 1971, p. 401, by permission of author and publisher.)

feeder. Under these penalty conditions, the median number of presses before trying the feeder was equal to the required number for required numbers ranging from 4 to 24.

As striking as the systematic increase in the median number of presses given as a function of the number required is the variability in the number given even when the number required was as small as 4. It appears that the process by which the rat represents a number has approximately a scalar variance property—the variance in the estimate of a numerosity is proportionate to the magnitude of the estimate—and that the Weber fraction, the constant of proportionality, is high. The Weber fraction is high enough that 4 is often confused with 3 or 5 and, frequently, even with 2 or 6. Increasing the penalty for premature entry does not substantially reduce this variance. Rather, it causes the animal to adopt a more conservative criterion, only trying the feeder when its estimate of the number of presses given is on average a fixed amount greater than the number required.

Meck and Church (1983) proposed and tested a serially incrementing counting mechanism as a model of the process by which an animal estimates numerosity (generates numerons). Their counting mechanism is a minor modification of Gibbon's very successful model of the mechanism by which animals estimate temporal durations (Gibbon, 1977). In the model for estimating the duration of a temporal interval, a switch closes at the start of a to-be-timed interval, gating a stream of pulses from a clock into an accumulator. The switch opens at the end of the interval, at which point the value in the accumulator (the sum of the pulses in the stream), is multiplied by a constant approximately (but not usually exactly) equal to 1 and stored in memory for comparison to subsequent intervals. To make this model into a counter, Meck and Church (1983) proposed that the switches that gate the pulse stream from the clock to the accumulator can also operate in what they call an "Event mode." In this mode, each sequential event or each item in an array of simultaneously presented stimuli closes the switch for a fixed duration, so that the value in the accumulator at the end of a series of such 'one-shot' closures is proportionate to the number of closures. Thus, in this model numbers are represented by magnitudes (the scaled contents of the accumulator), the very same magnitudes or at least the same kind of magnitudes that may be used to represent a continuous scalar variable like temporal duration.

The mechanism proposed by Meck and Church maps from to-be-counted events or stimuli to states of the integrator, thus states of the integrator constitute the numerons in this counting system. The mapping is one—one because each event gates one and only one (approximately equal) burst of pulses to the integrator, hence each successively counted entity is paired with a successively higher quantity in the integrator. The states of the integrator are run through in an order that is always the same because the ordering relation for quantity or magnitude (i.e., for the successive states of the integrator) is the same as the ordering relation for numerosity. The fact that adding successive increments to a quantity produces an ordered set of quantities just as adding successive ones

produces an ordered set of numbers is part of the reason that quantity or magnitude can be represented numerically, and vice versa. Finally, the state of the integrator at the end of the series of events is taken to represent a property of the series (its numerosity), not a property of the final event itself. Thus the process envisioned by Meck and Church conforms to the principles that define a counting process (Gelman and Gallistel, 1978).

Meck and Church (1983) and Meck, Church, and Gibbon (1985) gave several lines of experimental evidence in support of their hypothesis that numerons in animals are magnitudes like the magnitudes that represent durations. The most persuasive evidence comes from an experiment in which they correctly predicted a transfer of a discrimination between two durations to a discrimination between two numerosities. If in both tasks the behavioral decision (which level to choose) depends on a comparison between two magnitudes, then one might expect to see immediate transfer from a judgment based on duration to one based on numerosity when the two counts yielded the same magnitudes as the two durations.

Meck and Church (1983) first taught the rat to press one lever when it heard a steady noise of 2 seconds' duration and the other lever when it heard a steady noise of 4 seconds' duration. Next they determined the duration discrimination function. They interspersed the training trials with test trials, on which the rat heard steady noise bursts of intermediate duration ($2 < d < 4$) and was not rewarded for either choice (there being no "correct" choice when the stimulus was intermediate between the trained values of 2 and 4). By means of these interspersed trials, they could measure the discrimination function, that is, the tendency to choose the "4-sec lever" as a function of the duration of the stimulus.

Finally, they interspersed counting trials. On these trials, the rats heard 1-second noise bursts alternating with 1-second silent segments. On any one trial, the discriminative test stimulus consisted of somewhere between 10 and 20 cycles of this stimulus. Both the total duration of the test stimuli (20–40 seconds) and the summed durations of their sound-on or sound-off segments (10–20 seconds) greatly exceeded both training durations (2 versus 4 seconds). Pilot work indicated that in the counting or "Event" mode, the switch gated pulses to the accumulator for about 200 milliseconds for each counted event. Thus, the accumulation from the counting of a sequence of 10 second bursts would match the accumulation from the timing of a single burst of 2 seconds' duration ($0.2 \times 10 = 2$) and the accumulation from the counting of 20 bursts would match the accumulation from the timing of a single 4-second burst ($0.2 \times 20 = 4$). The accumulations from timing these cyclical test stimuli would be very much greater than the accumulations that were rewarded during duration training, but the accumulations from counting the number of bursts would fall in the same range as the accumulations rewarded during duration training. Thus the rats might be expected to apply the decision rule learned for durations to the quantities generated by their counts of the number of bursts, which in fact they did (Figure 6.3). The results in Figure 6.3, which were replicated and extended by Meck et al.

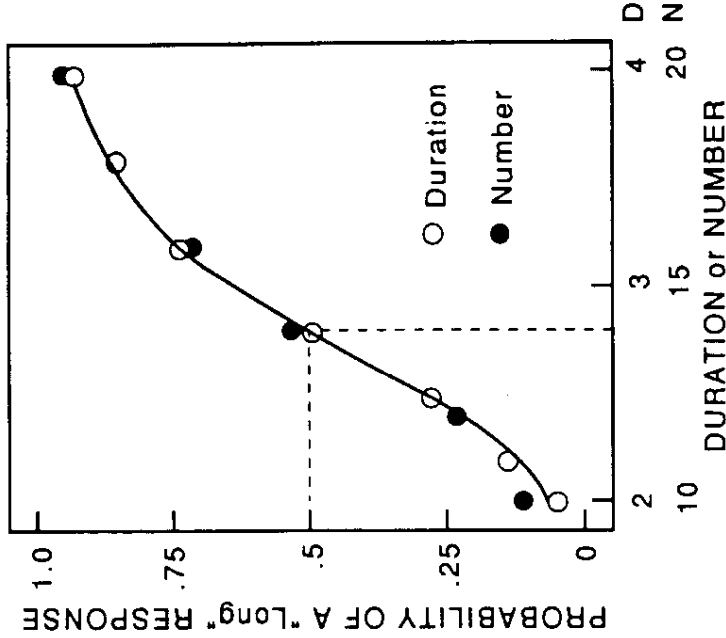


FIG 6.3. Probability of choosing the lever rewarded after the "long" (4-second) stimulus as a function of either the duration of a steady test stimulus or as a function of the number of cycles in a test stimulus consisting of alternating bursts of noise and silence. All the trials with durations other than the training durations (2 and 4 seconds) were interpolated unrewarded trials, as were all the trials with multiple bursts of noise (the number trials). The rats' generalization from the representation of a continuous variable (duration) to the numerosity of a discrete variable (noise bursts) suggests that both variables are represented by magnitudes. (Redrawn from Meck and Church 1983, p. 331, by permission of author and publisher.)

(1985), imply that the rat counts by making successive, roughly equal increments in the magnitude of some mental variable. In effect, the mind inverts the representational convention whereby numbers are used to represent linear magnitudes. Instead of using number to represent magnitude, the rat uses magnitude to represent number. Linear magnitude represents numerosity in the brain in the manner in which the lengths of the bars in a frequency histogram represent the underlying numbers of observations. The relevance of this model for the animal representation of number for our theory of subitizing is that it suggests that the primitive or

preverbal mental representation of numerosity is by magnitudes, the same mental magnitudes or the same kinds of mental magnitudes that are used to represent continuous variables like durations. In our theory of subitizing, we assume that this is the preverbal representation of numerosity in human beings as well.

The high scalar variance in the rat's representatives of numerosity may have at least two sources: miscountings (failing to close the switch for every item or closing it more times than there are items) and an inherent imprecision in the magnitude of the values retained in memory as representatives of the experienced numerosities. This imprecision could arise from the characteristics of memory itself or it might arise from variability in the magnitudes resulting from a fixed number of switch closures, hence in the magnitudes corresponding to a fixed numerosity. The latter kind of scalar variance would arise if, for example, the stream of pulses obeyed Poisson statistics, so that the number of pulses gated to the accumulator varied from one 200 msec switch closure to the next. The data on the discrimination of temporal intervals—where omitted or supernumerary switch closures (miscountings) are presumably not a factor—reveal a similar high scalar variance. This suggests that a significant fraction of the variance in number estimates derives from the second source, a scalar variance inherent in the process by which a numerosity is mapped to the mental magnitude that represents it, a variance that cannot be reduced by reducing miscountings.

VERBAL AND NONVERBAL REPRESENTATIONS OF NUMEROSITY IN HUMANS

The verbal representation of numerosity by arbitrarily derived morphemes is profoundly different from a representation based on magnitudes. Magnitude manifests the same ordering as numerosity, whereas there is no ordering inherent in the morphemes. The magnitude representing 2 differs more from the magnitude representing 4 than from the magnitude representing 3. No such ordering of differences inheres in the numerals 'two,' 'three,' and 'four' nor in the numerals '1,' '2,' '3,' and '4.' Thus, estimates of numerosity based on verbal counts cannot exhibit the second of the two kinds of variance just specified. Imprecision in generating or recording the magnitude that represents *four* yields a magnitude more appropriate to represent *three* or *five* much more often than it yields a magnitude appropriate to *two* or *six*, but imprecision in writing or '1' does not yield '2.' If anything, imprecision in writing '1' more often leads to confusion with '7' than with '2.' In short, the confusion matrices for numerals and numerographs are not determined by the relative magnitudes of the numerosities they represent. Assuming one has learned some conventional verbal counting list, the only source of error in the verbal representation of number would appear to be miscounting, which may be reduced by a careful execution of

the process that updates the partition between already-counted and not-yet-counted items in the course of a count.

The algorithms for addition and subtraction with verbal representatives of number also differ profoundly from the algorithms for performing these operations with magnitudes. The algorithm for combining two numerals to get a third numeral that represents their sum requires knowledge of the base for the particular verbal/orthographic code that is being used ($7 + 2 = 11$ in a base 8 system) and the rules that govern carrying in place-value notation. None of this is required when two magnitudes are combined to yield a third magnitude, as, for example, when two line segments are laid end to end to generate the segment that is their sum.

In view of the profound differences between the verbal and nonverbal representations of numerosity one might suppose that human number processing relied on one or the other but not both at the same time. However, work on the reaction time in number comparison tasks (Dehaene, 1989; Dehaene, Dupoux & Mehler, in press; Hinrichs, Yurko & Hu, 1981; Holyoak, 1978; Holyoak & Mah, 1982) suggests that when human beings work with some kinds of verbally or symbolically presented number problems, they commonly map from the verbal representatives of numerosity to the more primitive nonverbal representation, compute solutions using the nonverbal representation, then map back from it to verbal answers. Evidence of this is the long established fact that the reaction time for a "greater than" or "less than" judgment between a remembered reference number and a newly presented comparison number is quicker the greater the interval between them, even for numbers that straddle decade boundaries. (One can judge 59 to be greater than 49 more quickly than one can judge 51 to be greater than 49.)

The most startling evidence for the hypothesis that human beings routinely map back and forth between their verbal and preverbal representations of numerosity comes from a profoundly acalculic neuropsychiatric patient (Dehaene, personal communication). When asked whether it was true that $2 + 2 = 4$, the patient said "yes," but he also said "yes" when asked whether $2 + 2 = 3$ and when asked whether $2 + 2 = 5$, from which one is inclined to conclude that his answers were confabulations. However, when asked whether $2 + 2 = 9$, he answered "no." It appears that he has lost access to the verbally created addition table that we learn in our early schooling but that he could still map from the numerographs to their corresponding magnitudes (preverbal numerons), add the magnitudes to get an imprecisely specified third magnitude, and map from this third magnitude back to the numerographs that correspond to approximately that magnitude. Because the specification of this third magnitude is imprecise, he cannot readily distinguish between numerographic answers that correspond to closely adjacent magnitudes (nemrosities), but he can reject numerographs that correspond to grossly different magnitudes.

These findings strengthen the hypothesis that the rudiments of the verbal system of number are learned by reference to a preverbal magnitude-based representation of number that is part of our animal heritage and that mappings back and forth between the preverbal representatives and the verbal representatives of number are a routine part of the mental manipulation of number. This hypothesis leads to our theory of subitizing.

A MODEL OF SUBITIZING

We suggest that in the subitizing range, human subjects use the preverbal counting process revealed in research on animals. We suggest that for small numerosities, the preverbal counter can run much faster than the verbal counting routine, so that the magnitude (the preverbal numeron) that represents the numerosity of an array is specified long before the verbal counting process has arrived at the correct numerlog. It is specified so much sooner that one can get the correct numerlog more quickly by using the preverbal counting mechanism (subitizing) to get the preverbal numeron and use the preverbal numeron to retrieve the numerlog. It is likely that this strategy would work even for numerosities much larger than 4 if it were not for the imprecision with which the preverbal numeron is specified. This imprecision leads to too many inaccurate responses. Because the variance (inaccuracy) in the specification of the magnitude of the numeron is scalar, the probability of giving a response that is erroneous by ± 1 increases with the numerosity of the array. As the chance of making an error increases, the subject relies more frequently on the verbal count and less frequently on the answer generated by way of the subitizing process. When the size of the array exceeds four, the chance of an error in the answer generated by way of the subitizing process is so great that the subject always waits for the answer generated by verbal counting.

The verbal counting process and the preverbal counting process are both subject to error because of miscounting, that is, because of procedural failures that violate the one-one principle (Gelman & Gallistel, 1978). These miscounts are very common in children's counting, which is an important reason why the reaction-time function increases so much more steeply for kindergarten children than it does for adults (Figure 6.1 panel C). One of the skills that develops with extensive experience at counting is skill at choosing and applying the partitioning processes that segregate the already counted from the not yet counted in the course of enumeration (Gelman & Greeno, 1989). However, the preverbal counting process is subject to an additional source of variance. The values of the magnitudes yielded by this process on counts vary when there has been no miscounting. Hence the preverbal representative of a numerosity is a probability distribution defined over the equivalent of a mental number line. Thus, in the preverbal counting process, there is some likelihood of confounding the numeron

for a given numerosity with the numerons for adjacent numerosities even when there has been no miscounting. We call this the *inherent variance*. The variance due to miscounts and the inherent variance both increase as a function of numerosity. But, the Weber fraction for the second source of variance is much higher than the Weber fraction for the first source, at least in skilled verbal counters. This is why skilled counters increasingly shift to reliance on the slower but more accurate verbal counting process as the numerosity increases from 1 to 4 or 5.

Further evidence that the preverbal counting mechanism is fast but inaccurate comes from an experiment by Rilling and McDiarmid (1965), who applied a signal-detection analysis to the pigeon's ability to discriminate the numerosities of its pecks. The pigeon had to choose between two illuminated side keys on the basis of the number of pecks it had just made on a center key. The experimental program controlled this number, that is, the side keys were illuminated after a preprogrammed number of pecks on the center key. Which side key was the correct (rewarded) key varied as a function of the preprogrammed number for the center key. The pigeons discriminated numerosities in the range 45–50 even though they pecked the center key at a rate of 6 pecks per second, which is twice as fast as humans can count (Figure 6.4). However, they were so inaccurate in their counts that they confounded 45 with 50 on 10–20% of the trials, and when

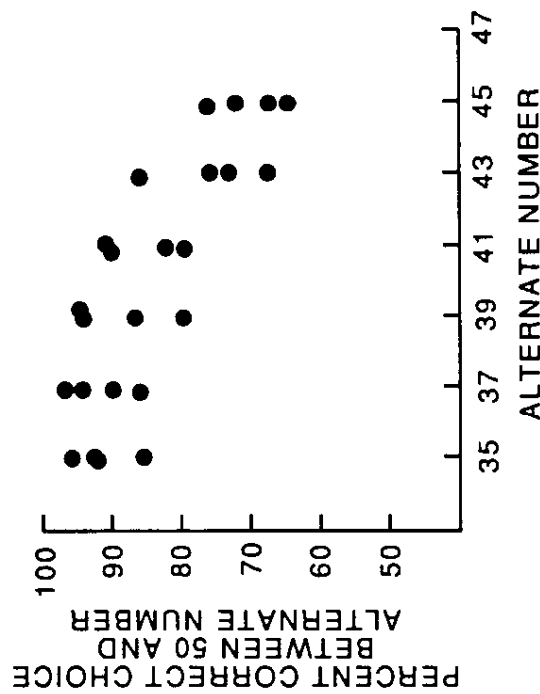


FIG 6.4. Scatter plot from four determinations of a pigeon's ability to discriminate between a run of fifty key pecks and an alternative run whose number varied systematically in the range from thirty-five to forty-seven. (Redrawn from Rilling and McDiarmid (1965), p. 527, by permission of author and publisher.)

the difference to be discriminated was .48 versus .50, discrimination failed altogether. For the evidence that the pigeons based their discrimination on the numerosity of their pecks and not on the duration of their pecking, which covaried with numerosity, see Rilling (1967).

RELEVANCE OF THE PRESENT MODEL FOR STUDIES OF HUMAN INFANTS

The numerical abilities of infants have been studied by several research groups using one or both of two methods. In one of these, infants are first habituated to a display or set of displays with arrays of N items (usually ranging from 1 to 4). Their tendency to start responding again when they are then shown displays of a different value is taken as an index of their ability to discriminate between the two numerosities. In the second of these methods, infants encounter two displays with arrays of different values, for example 2 and 3, and their preference for one display over the other serves as evidence of their ability to discriminate between the two numerosities. A variant of the preference technique has been used to assess infants' abilities to respond intermodally to numerosity. For example, Starkey, Spelke, and Gelman (1990) showed infants a pair of side by side, heterogeneous displays that varied in numerosity (2 vs. 3) while two or three drum beats were heard through a hidden speaker placed midway between the displays. An infant's tendency to look systematically across trials at the display that contains either the same of a different number of items than the number of drumbeats serves as an index of their ability to link stimuli presented via different modalities on the basis of common numerosity.

Different groups of investigators, using one or another variant of the habituation or preference methods with either homogeneous or heterogeneous items in their displays, have converged on a common set of findings. Infants between 1 and 10 months of age discriminate reliably between two and three items (Antell & Keating, 1983; Starkey & Cooper, 1980; Strauss & Curtis, 1981; Starkey, Spelke & Gelman, 1983); they sometimes discriminate between three- and four-item displays (Strauss & Curtis, 1981); and they fail to discriminate between four and five items or four and six items (Starkey & Cooper, 1980; Strauss & Curtis, 1981). In addition, 4- to 6-month-old infants respond intermodally to numerosity. Starkey et al. (1990) have shown that 6- to 8-month-old infants have a reliable tendency to look at the display which matches in numerosity the number of drumbeats they hear on a given trial. Indeed, in one of these studies, infants even matched numerosity when then the visual and auditory inputs were presented sequentially. In a related study by Moore, Benenson, Reznick, Peerson, and Kagan (1987), infants looked at the numerosity that matched the number of sounds they heard during their first phase of testing. During the last phase of testing their behavior followed a common path in the infancy literature and

responded to novelty, that is, they switched to pairing sounds and slides that differed in numerosity (See Starkey et al., 1990, for the pertinent analyses of the Phase 1 and 2 differences in Moore et al., 1987). In both phases, their responding was a systematic function of the relative numerosities of the auditory sequence and the array.

The above findings on infants' numerical abilities are typically taken as evidence that they do not count. For example, Strauss and Curtis (1984) write, "Finally, there is no evidence that infants can abstract numerosities larger than about three or four. Logically, there is no reason why a counting strategy should have such a limitation" (p. 144). We have here developed an account of preverbal counting for which there is a clear reason for such a limitation. Because the infant findings are much like those in studies of animals' abilities to count, both in terms of their different degrees of success as set sizes increase from 2 and 3 to 4 and 5, they are fully consistent with the hypothesis that infants do count and compare when given either sequential events or simultaneous arrays. More importantly, our preverbal counting model makes intelligible the basis of the numerical comparisons implied by the infants' patterns of responses across the different sizes used. If, as set size increases, the preverbal counting mechanism generates increasing variance in the resulting numerons, then discrimination between adjacent numerosities should start to break down at numerosities of 4 and 5. Similarly, discriminations between four and six items will be unreliable. For the variance in the estimate of a numerosity 4 is already high enough to produce confusion with 3 and 5 and the variance of the representations of 4 and 6 overlap enough to make discriminations difficult.

Given that we can readily incorporate the extant data base on human infants' abilities to represent numerosity into our preverbal counting model, it is premature to reject the hypothesis that infants count. Our account points to the kind of research that would provide suitable tests of this model. Studies are needed where one systematically varies the difference between pairs of set sizes so as to select stimuli that will generate varying degrees of discriminability in the numbers representing numerosity. Our prediction is that infants will fare better on those pairs where the resulting variances do not overlap appreciably than on those pairs where there is appreciable overlap in the probability density distributions. In absence of such data, it is premature to conclude that human infants do not use a nonverbal counting procedure, let alone that they use some undefined perceptual apprehension process (e.g., von Glasersfeld, 1982).

CONCLUSION

Verbal concepts and verbally mediated human behaviors must have a preverbal substrate, which evolved prior to the very recent evolution of the language ability, yet it is surprising how seldom cognitive psychology is able to link the

verbal capacities to their preverbal antecedents. We believe that there is hope for establishing such a linkage in the number domain.

ACKNOWLEDGMENTS

Preparation of this chapter was supported in part by NSF grant #BNS 8961220 to Rochel Gelman.

REFERENCES

- Allport, D. A. (1975). The state of cognitive psychology. *Quarterly Journal of Experimental Psychology*, 27, 141-152.
- Antell, S. E., & Keating, D. P. (1983). Perception of numerical invariance in neonates. *Child Development*, 54, 695-701.
- Boysen, S. T., & Berntson, G. G. (1989). Numerical competence in a chimpanzee (*Pan troglodytes*). *Journal of Comparative Psychology*, 103, 23-31.
- Capaldi, E. J., & Miller, D. J. (1988). Counting in rats: Its functional significance and the independent cognitive processes which comprise it. *Journal of Experimental Psychology: Animal Behavior Processes*, 14, 3-17.
- Chi, M. T. H., & Klahr, D. (1975). Span and rate of apprehension in children and adults. *Journal of Experimental Child Psychology*, 19, 434-439.
- Cooper, R. G., Jr. (1984). Early number development: Discovering number space with addition and subtraction. In C. Sophian (Ed.), *The origins of cognitive skills* (pp. 157-192). Hillsdale, NJ: Lawrence Erlbaum Associates.
- Davis, H., & Pérusse, R. (1988). Numerical competence in animals. *Behavioral and Brain Sciences*, 11, 561-615.
- Dehaene, S. (1989). The psychophysics of numerical comparison: A reexamination of apparently incompatible data. *Perception and Psychophysics*, 45(6), 557-566.
- Dehaene, S., Dupoux, E., & Mehler, J. (in press). Is numerical comparison digital: Analogical and symbolic effects in two-digit number comparisons. *Journal of Experimental Psychology: Human Perception and Performance*.
- Fernandes, D. M., & Church, R. M. (1982). Discrimination of the number of sequential events by rats. *Animal Learning and Behavior*, 10, 171-176.
- Gallistel, C. R. (1990). *The organization of learning*. Cambridge, MA: Bradford Books/MIT Press.
- Gelman, R., & Gallistel, C. R. (1978). *The child's understanding of number*. Cambridge, MA: Harvard University Press.
- Gelman, R., & Greeno, J. G. (1989). On the nature of competence: Principles for understanding in a domain. In L. B. Resnick (Ed.), *Knowing and learning: Issues for a cognitive science of instruction*. Hillsdale, NJ: Lawrence Erlbaum Associates, pp. 125-186.
- Gibson, J. (1977). Scalar expectancy theory and Weber's Law in animal training. *Psychological Review*, 84, 279-335.
- Hinrichs, J. V., Yurko, D. S., & Hu, J. M. (1981). Two-digit number comparison: Use of place information. *Journal of Experimental Psychology: Human Perception and Performance*, 7, 890-901.
- Holyoak, K. J. (1978). Comparative judgments with numerical reference points. *Cognitive Psychology*, 10, 203-243.

- Holyoak, K. J., & Mah, W. A. (1982). Cognitive reference points in judgments of symbolic magnitudes. *Cognitive Psychology*, 14, 328-352.
- Jensen, E. M., Reese, E. P., & Reese, T. W. (1950). The subitizing and counting of visually presented fields of dots. *Journal of Psychology*, 30, 363-392.
- Kaufman, E. L., Lord, M. W., Reese, T. W., & Volkman, J. (1949). The discrimination of visual number. *American Journal of Psychology*, 62, 498-525.
- Klahr, D., & Wallace, J. G. (1973). The role of quantification operators in the development of conservation of quantity. *Cognitive Psychology*, 4, 301-327.
- Mandler, G., & Shebo, B. J. (1982). Subitizing: An analysis of its component processes. *Journal of Experimental Psychology: General*, 11, 1-22.
- Meck, W. H., & Church, R. M. (1983). A mode control model of counting and timing processes. *Journal of Experimental Psychology: Animal Behavior Processes*, 9, 320-334.
- Meck, W. H., Church, R. M., & Gibbon, J. (1985). Temporal integration in duration and number discrimination. *Journal of Experimental Psychology: Animal Behavior Processes*, 11, 591-597.
- Moore, D., Benenson, J., Reznick, S., Pearson, M., & Kagan, J. (1987). Effect of auditory numerical information on infants' looking behavior: Contradictory evidence. *Developmental Psychology*, 23, 665-670.
- Platt, J. R., & Johnson, D. M. (1971). Localization of position within a homogeneous behavior chain: Effects of error contingencies. *Learning and Motivation*, 2, 386-414.
- Rilling, M. (1967). Number of responses as a stimulus in fixed interval and fixed ratio schedules. *Journal of Comparative and Physiological Psychology*, 63, 60-65.
- Rilling, M., & McDiamid, C. (1965). Signal detection in fixed ratio schedules. *Science*, 148, 526-527.
- Rumbaugh, D. M., Savage-Rumbaugh, S., & Hegel, M. T. (1987). Summation in the chimpanzee (*Pan troglodytes*). *Journal of Experimental Psychology: Animal Behavior Processes*, 13, 107-115.
- Saltzman, I. J., & Garner, W. R. (1948). Reaction time as a measure of span of attention. *The Journal of Psychology*, 25, 227-241.
- Shipley, E., & Shepperson, B. (1980). Countable entities: Developmental changes. *Cognition*, 34, 109-136.
- Starkey, P., & Cooper, R. G. (1980). Perception of numbers by human infants. *Science*, 210, 1033-1035.
- Starkey, P., Spelke, E. S., & Gelman, R. (1983). Detection of intermodal correspondences by human infants. *Science*, 222, 179-181.
- Starkey, P., Spelke, E. S., & Gelman, R. (1990). Numerical abstraction in human infants. *Cognition*, 97-128.
- Strauss, M. S., & Curtis, L. E. (1981). Infant perception of numerosity. *Child Development*, 52, 1146-1152.
- Strauss, M. S., & Curtis, L. E. (1984). Development of numerical concepts in infancy. In C. Sophian (Ed.), *Origins of cognitive skills* (pp. 131-155). Hillsdale, NJ: Lawrence Erlbaum Associates.
- von Glasersfeld, E. (1982). Subitizing: The role of figural patterns in the development of number concepts. *Archives de Psychologie*, 50, 191-218.
- Wynn, K. (in press). Children's understanding of counting. *Cognition*.
- Woodworth, R. S., & Schlosberg, H. (1954). *Experimental psychology: Revised edition*. New York: Holt, Rinehart & Winston.