Variability signatures distinguish verbal from nonverbal counting for both large and small numbers

SARA CORDES, ROCHEL GELMAN, and C. R. GALLISTEL
University of California, Los Angeles, California
and Rutgers University, New Brunswick, New Jersey

and

JOHN WHALEN
University of Delaware, Newark, Delaware

Humans appear to share with animals a nonverbal counting process. In a nonverbal counting condition, subjects pressed a key a numeral-specified number of times, while saying "the" at every press. The mean number of presses increased as a power function of the target number, with a constant coefficient of variation (c.v.), both within and beyond the proposed subitizing range (1–4 or 5), suggesting small numbers are represented on the same continuum as larger numbers and subject to the same noise process (scalar variability). By contrast, when subjects counted their presses out loud as fast as they could, the c.v. decreased as the inverse square root of the target value (binomial variability instead of scalar variability). The unexpected power-law relation between target value and mean number of presses in nonverbal counting suggests a new hypothesis about the development of the function relating number symbols to mental magnitudes.

Humans appear to share with animals a nonverbal counting process that generates mental magnitudes representing numerosities (both small, ≤ 5, and large; Gallistel & Gelman, 1992, 2000; Gelman & Cordes, 2001). In Meck and Church's (1983) model of nonverbal counting, each enumerated unit is represented by an additional fixed increment of magnitude in the contents of a mental accumulator. The cardinal value of the counted set or sequence is represented by the final magnitude, which is read into memory (Figure 1).

These magnitudes in memory are noisy, like the magnitudes that represent continuous quantities. Following Meck and Church (1983), we assume that the mean value and the variability of the mental magnitude distribution for a numerosity are proportional to the numerosity. We further make the standard signal-detection assumption that discrimination failures occur to the extent that the signal distributions for nearby numerosities overlap.

Other accumulator models have been proposed. Reynvoet and Brysbaert (1999), as well as Dehaene, Dupoux, and Mehler (1990), hold that the mental magnitudes are proportional to the logarithm of the objective magnitudes and have constant variability (signal noise independent of signal magnitude). However, this logarithmic explanation is inconsistent with results that implicate the ability to mentally subtract with both durations and numerosities (Brannon, Wusthoff, Gallistel, & Gibbon, 2001; Gibbon & Church, 1981). The Brannon et al. (2001) results imply that the mental magnitudes in memory are proportional to the objective magnitude of the numerosities they represent.

The magnitude model explains a variety of human data, including the Weber-law characteristic of numerical order judgments—namely, the accuracy and rapidity with which the numerical order of two numerals may be discriminated is determined by their ratio (e.g., Fetterman, 1993; Moyer & Landauer, 1967, 1973; see Dehaene, Dehaene-Lambertz, & Cohen, 1998; Gallistel & Gelman, 1992, for reviews). The effects of numerical size and distance on the speed and accuracy of human order judgments have also been observed in primates (Brannon & Terrace, 2000; Rumbaugh, Savage-Rumbaugh, & Hegel, 1987; Washburn, 1994). Results like these fit well with Moyer and Landauer's (1967) proposal that judgments of numerical order are represented by noisy mental magnitudes. On this hypothesis, the discriminations obey Weber's law because the overlap between two signal distributions with scalar variability is determined by the ratio of their means.

The importance of scalar variability and, more generally, the Weber-law characteristic in numerical estimation and discrimination is that scalar variability and Weber's law
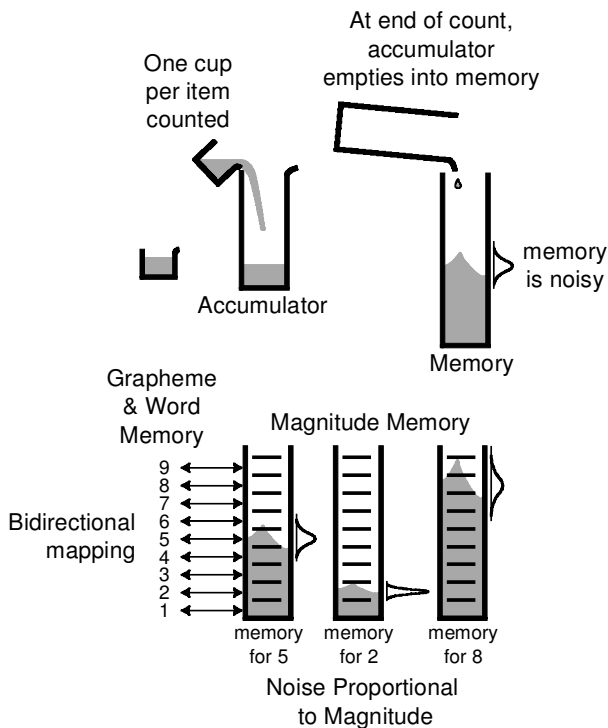
Figure 1. The accumulator model. A magnitude representing a numerosity is formed through accumulation of "cup fulls" of activation, one cup for each item or event enumerated. Accumulated magnitudes from an ongoing count may be compared to a magnitude stored in memory or may be mapped to symbols for quantities. However, magnitudes read from memory have inherent scalar variability that may result in errors. The greater the magnitude, the more likely an error.

are not predicted by another obvious source of variability in numerical estimation and discrimination—counting errors. Counting is a multistep process, with the number of steps proportional to the numerosity counted. If there is some probability of error (either skipping an item or counting it twice) at every step, then the more steps there are, the greater the expected accumulation of miscounts. The variability in counts from this source should, however, obey binomial statistics; it should grow in proportion to the square root of the numerosity. Thus, the coefficient of variation, the ratio of the standard deviation to the mean, should be negatively correlated with target number. Discriminations should get *relatively* better as numerosities increase. That is, the discrimination of 30 from 20 should be more accurate than the discrimination of 3 from 2, contrary to Weber's law.

Whalen, Gallistel, and Gelman (1999) provided some evidence that adults use a nonverbal counting mechanism that generates magnitude representations with scalar variability. In their key press task, subjects were presented with an odd Arabic numeral from 7 to 25 and were asked to rapidly press a key, without counting, until they felt they had arrived at that number. Results paralleled the animal counting data (e.g., Platt & Johnson, 1971); the number of presses

increased in proportion to target number, as did the trial-to-trial variability in the number of presses.

To assess whether subjects responded on the basis of duration rather than numerosity, Whalen et al. (1999) had subjects attempt to reproduce time intervals. The coefficients of variation for subjects' duration estimates were much higher than the coefficients for their key press responses. Whalen et al. also asked subjects to estimate the amount of time a given number of presses took. These estimates were grossly inaccurate.

To counter the hypothesis that their subjects counted subvocally, they adduced results from an analysis of interresponse intervals and other arguments. Still, this is a somewhat indirect line of evidence. Because this is a key point, we wanted to strengthen the case for the nonverbal nature of the counting process used under their conditions. To prevent our subjects' use of subvocal articulation of count words, we had them say out loud the word "the" coincident with each press, while they pressed a key as fast as they could to generate a target numerosity specified by a numeral. Data from this condition contrasted with those from two verbal counting conditions, where subjects explicitly counted their presses. In these latter conditions, the coefficient of variation declined as the inverse square root of the target number (binomial variability).

We also extended the range of target numbers into the so-called subitizing range to look for evidence of discontinuity in variability around the transition from subitizing to nonverbal counting. Many authors have argued that in this range, numerosity is directly given by low-level perceptual mechanisms that yield discrete (as opposed to magnitude) representatives of numerosity and that do not employ a serial counting process (e.g., Simon, 1999; Spelke, 2000). If the mental representatives of numerosities less than or equal to 5 are discrete-valued variables rather than real-valued magnitudes, then we should not observe scalar variability in this range.

METHOD

Subjects

Eight adult volunteers (2 males, 6 females) from the University of California, Los Angeles, participated in the "The" (nonverbal counting) condition. Two subjects participated twice.[1] Six of these volunteers also participated in the full count and tens count (verbal counting) conditions. One of these subjects participated twice in each of these conditions.[2] (The distribution of subjects across conditions and stimuli is summarized in Table 1.)

Apparatus

Stimulus presentation and response input were similar to those of Whalen et al. (1999). A Macintosh G3 computer and Psyscope stimulus presentation software were used to present the stimuli. A computer joystick connected to a Psyscope Buttonbox recorded both responses and latencies.

Procedure

Each trial began with a "Ready?" message in the center of the computer screen. After subjects pressed a button on the computer joystick to start the trial, an Arabic numeral appeared on the screen, spec-

Table 1
Distribution of Subjects Across Conditions and Presented Stimuli

| | Condition | | |
Stimuli	"The"	Full Count	Tens Count
(3, 5, 8, 13, 20, 32)	S100	S100	S100
	S101	S101 (excluded)	S101
(2, 3, 4, 5, 8, 13, 20, 32)	S102*	S102*	S102*
	S103*		
	S106		
	S109		
(7, 9, 11, 13, 15, 17, 19, 21, 23, 25)	S102*	S102*	S102*
	S103*	S103	S103
	S104	S104	S104
	S105	S105	S105

Denotes subjects participating in more than one session within a given condition.

ifying the target number of key presses for that trial. The range of target stimuli tested varied somewhat across subjects (Table 1).

Subjects were told to press the left key on the joystick with their preferred hand as fast as they could until they felt they had arrived at the target number of key presses. In the "The" (nonverbal counting) condition, subjects were instructed not to count their presses, but instead to repeat the word "the" at every press. By contrast, in the full count and tens count conditions, subjects were told to count their presses out loud. In the full count condition, they were told not to count by tens and not to shorten the count words (e.g., say "twe-four" for "twenty-four"). In the tens count condition, they were told to count aloud by tens (e.g., 1, 2, … 9, 10, 1, 2, …).[3] Subjects signaled completion of a sequence of presses by pressing the right key once.

From each subject, we obtained 20 data points per target number. The total number of trials per subject varied from 120 to 200, depending on the number of targets they had. An experimental session,
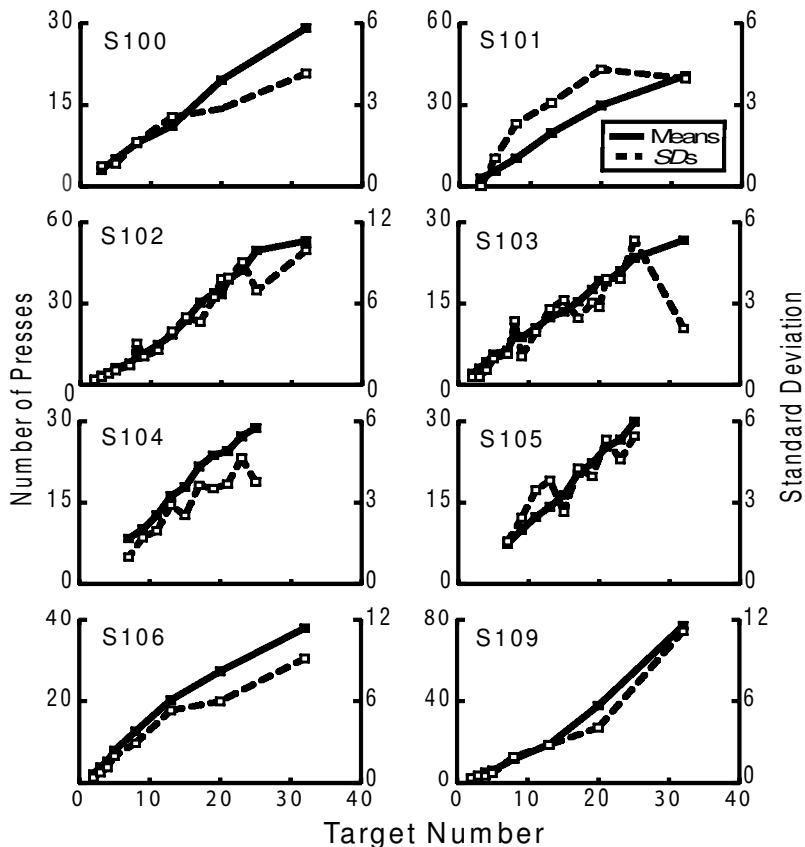

Figure 2. Mean number of key presses and the standard deviation in the distribution of presses as a function of target number for each subject in the "The" condition. Note that different *y*-axis scales are used to depict the data for Subjects 101, 102, 106, and 109.
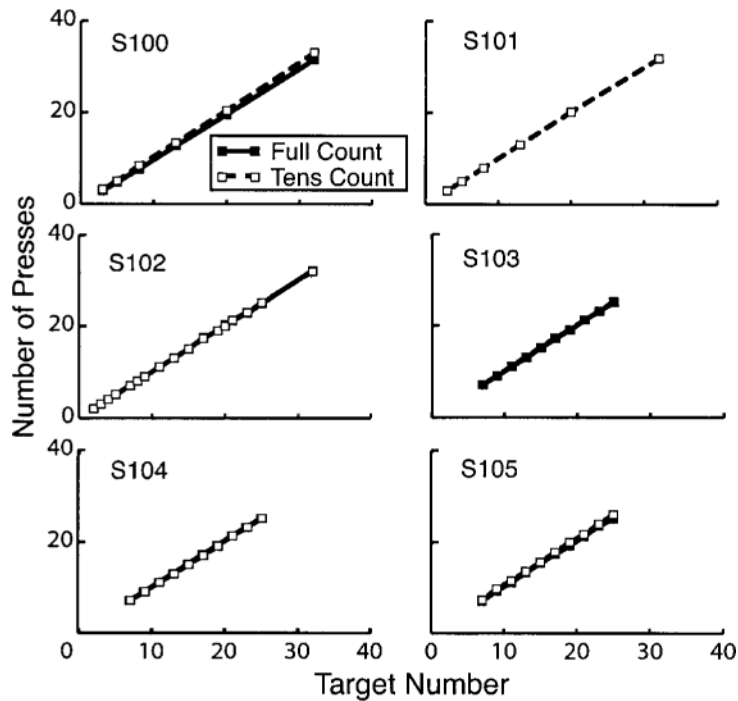
Figure 3. Mean number of presses as a function of target number for each subject in both verbal counting conditions (full count and tens count).

however, did not last more than 15 min. In order to ease performance anxiety, subjects were told that responses were not expected to be completely accurate. To ensure compliance with instructions, the experimenter sat nearby, monitored performance, and audio-recorded the sessions.

RESULTS

Audio recordings and experimenter observation verified compliance with instructions by all subjects in both the "The" and tens count conditions, and by 5 of the 6 subjects in the full count condition. One subject was unable to synchronize each count word with one and only one key press and so data from this session were excluded from analyses.

In all three conditions, the mean number of key presses increased in proportion to the target number, with a near one-to-one correspondence in both verbal counting conditions (Figures 2 and 3).[4] The standard deviations of the response distributions also increased with target number in all conditions. This pattern was most evident in the "The" (nonverbal counting) condition, where distributions were notably wider and more drastically dependent upon target number than in the other two conditions (Figures 2 and 4). (Although less apparent, regression analyses did reveal an increase in variability with target number in every subject in the full count and 5 of 6 subjects in the tens count experiment.)

Because the variability in the estimates of means and standard deviations increases with target number, so does the variability in their ratio. Thus, simple regression analyses cannot be used to test hypotheses about the relation between this ratio and $\overline{n}$. Logarithmic transformation equates variability in the estimated coefficient of variation, $\sigma_n/\overline{n}$, at different values of $\overline{n}$. Thus, to test hypotheses about the relation between the coefficient of variation and the mean, we plotted $\log(\sigma) - \log(\overline{n})$.[5] For the scalar variability hypothesis (the nonverbal counting hypothesis), the slope of this relation should be zero. For the binomial variability hypothesis (the verbal counting error hypothesis), the slope should be $-.5$. [The hypothesis is that $\sigma_n = k\overline{n}^{-.5}$. Dividing through by $\overline{n}$ gives $\sigma_n/\overline{n} = k\overline{n}^{-.5}\,\overline{n}^{-1} = k\overline{n}^{-.5}$, and so $\log(\sigma_n/\overline{n}) = \log(k) - 0.5\log(\overline{n})$.]

In the "The" condition, for 5 of the 8 subjects, the slope of the plot did not differ significantly from zero ($p > .05$), whereas for 7 of them, it differed significantly from $-.5$ ($p < .05$). The mean of these regression slopes (one slope per subject, $M = -0.06$) did not differ significantly from zero [$t(7) = 0.39, p > .05$] but did differ significantly from $-.5$ [$t(7) = 2.72, p < .05$]. These results are consistent with the scalar variability hypothesis and inconsistent with the hypothesis that verbal counting errors are the principal source of variability in the "The" condition.

By contrast, for 3 of the 5 subjects in the full count condition and 5 of the 6 subjects in the tens count condition, the slope of the plot differed significantly from zero ($p < .05$), although only 1 subject in each condition obtained a slope differing from $-.5$ ($p < .05$). In addition, in the tens count condition, the mean slope of the regression ($-.69$) differed significantly from zero [$t(5) = 3.04, p < .05$]. The
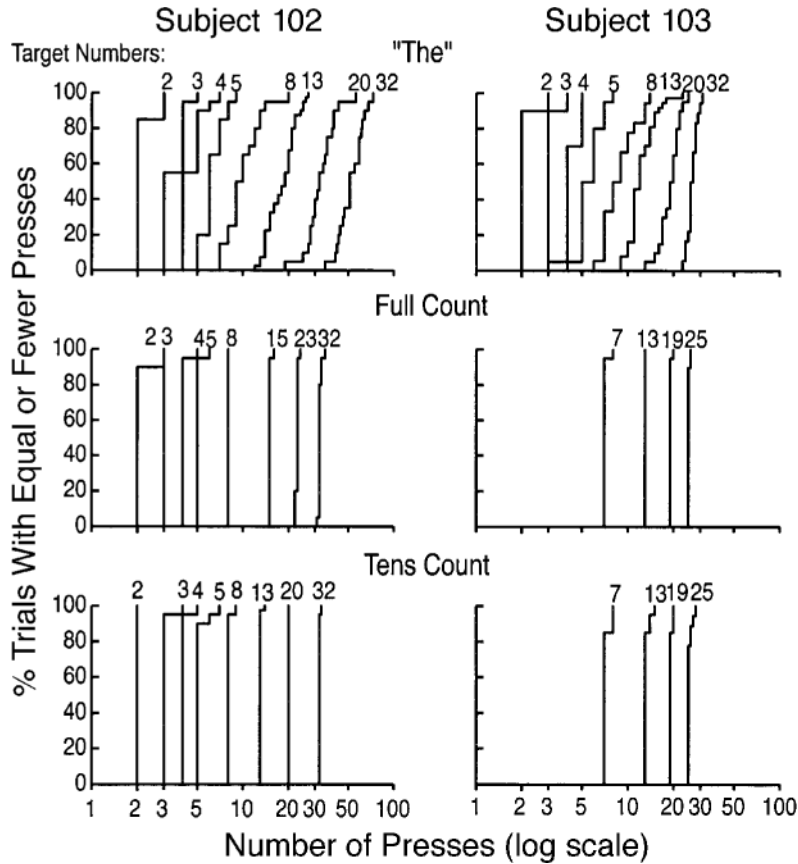
Figure 4. The cumulative normalized distributions of the number of presses made by 2 representative subjects (Subjects 102 and 103) for selected target numbers in the "The" (nonverbal counting), full count (verbal counting), and tens count (verbal counting) conditions.

difference between the mean slope (−.51) and a zero value also approached significance in the full count condition [$t(4) = 2.42, p = .07$], suggesting that verbal counting did not produce scalar variability. Neither mean slope, however, differed significantly from −.5 [$t(5) = 0.83$ and $t(4) = 0.05, p > .05$, respectively], the value expected by the binomial variability hypothesis. Thus verbal counting and nonverbal counting are distinguished by the manner in which variability increases with the mean count: It increases in accord with the binomial law in the first case and in accord with the scalar law in the second. (The coefficients of variation are plotted against the mean number of presses in Figure 5.)

Graphs of mean interresponse times versus press number indicated a constant linear relationship with a slope near zero in the "The" condition. For example, the amount of time between the 4th and 5th key presses was the same as that between the 16th and 17th key presses. There were no observable patterns in the data to suggest the use of chunking strategies by the subjects. These low, flat interresponse times were also observed in the tens count condition, but

they were not observed in the full count condition. As expected, given the requirement to count out loud as fast as possible for values into the thirties and not "cheat" by dropping the decade term, interresponse times in this condition increased at each decade boundary (Figure 6).

To address the subitizing questions, we examined data from the "The" condition for evidence of a discontinuity at the limits of the supposed subitizing process. A paired comparisons t test revealed no significant difference between the regression slopes of the data for target numbers in the subitizing range (2–5) and the slopes of the data for targets beyond that range [$t(5) = 0.486, p > .05$]. This result was also obtained using a more conservative subitizing range (2–4) as well as when the examination was confined to data from those subjects presented with the full range of targets, from 2 to 32. Thus our data are consistent with the hypothesis of quantitative continuity between the range in which subitizing is supposed to operate and the range beyond that: Scalar variability is seen both within and outside the putative subitizing range, and the coefficient of variation is the same both within and outside that range.
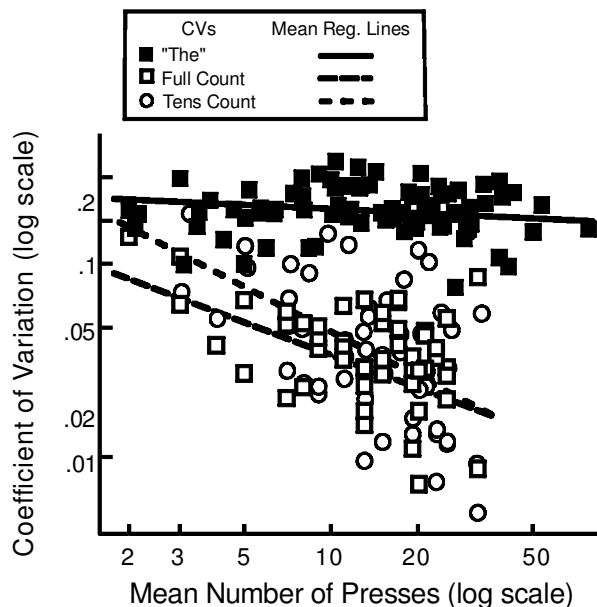
Figure 5. Coefficients of variation versus mean number of presses on double-log coordinates for all three conditions. The lines drawn through these group data have the mean slope and mean intercepts of the population of slopes and intercepts obtained from the subject-by-subject regressions.

DISCUSSION

Evidence for a Magnitude-Based Representation of Remembered Number and a Nonverbal Counting Mechanism

Our results give further evidence that adult humans have access to a system of noisy magnitudes as a form of nonverbal number representation, as first suggested by Moyer and Landauer (1967), and that they have a nonverbal counting mechanism that generates magnitudes, as first suggested by Meck and Church (1983). We believe that requiring subjects to talk out loud while pressing at faster than five presses per second a target number of times effectively precluded their using verbal counting, including subvocal counting. Nonetheless, subjects in the "The" condition were able to approximate the target number of key presses.

The nature of the variability in their approximations reinforces the hypothesis that they were not verbally counting. When subjects counted under conditions intended to suppress verbal counting, the variability in the achieved numbers of presses was proportional to the mean for a given target (scalar variability). By contrast, when the same subjects counted verbally, the variability was proportional to the square root of the mean number achieved for a given target (binomial variability).

Experimenter observation, subject reports, and response time data further reinforce the conclusion that the verbal rehearsal of the word "the" coincident with each press prevented subvocal counting. The results of an earlier sim-

ilar study by Logie and Baddeley (1987) are consistent with this interpretation. They had subjects count a sequence of fast aperiodic flashes while rapidly repeating "the" out loud. This greatly reduced correct counting; subjects got the correct count less than 50% of the time as the number of flashes increased beyond 6 or 7. However, Logie and Baddeley noted that the mean number that subjects reported after a count was proportional to the true number. Moreover, the average magnitude of the errors increased in proportion to the target number (see Logie & Baddeley, Figures 5 and 7). Logie and Baddeley did not consider the possibility of a nonverbal counting process, but they recognized that their data were difficult to explain in terms of subvocal counting using an articulatory loop. Our interpretation is that under their conditions, subjects used a nonverbal counting mechanism and that the variability in their responses arose from a noisy mapping within memory between the nonverbal magnitude and the verbal symbols for numerosity (the bidirectional mapping hypothesis, Whalen et al., 1999).

Magnitudes are formally equivalent to the real numbers. It is because discrete, like continuous, quantities (Meck & Church, 1983) are represented by magnitudes that we argue the nonverbal real numbers are the foundation of numerical representations (Gallistel & Gelman, 1992, 2000; Gallistel, Gelman, & Cordes, in press). This hypothesis counters a widely held assumption: Numerical representations are based on the integers and that the problem for development and learning is to construct the real numbers from them (Carey, 1998, 2001; Hauser & Carey, 1998; Leslie, Xu, Tremoulet, & Scholl, 1998; Spelke, 2000).

Scalar variability is inconsistent with models that posit a mental representation of numerosity lacking the properties of the system of magnitudes (real numbers). Proportions cannot be defined in symbolic systems with only ordinal or weaker properties. Thus, any model in which the mental representation of numerosity lacks numerical structure altogether or has only ordinal structure (more/less) cannot explain why numerosities separated from target numerosities by a fixed proportion are equally likely to be confounded with (or given in place of) the target numerosities. That is why scalar variability is such a theoretically important empirical property.

Our results extend the earlier results of Whalen et al. (1999) in another theoretically important direction by showing quantitative continuity between the so-called subitizing range and a range of numbers that, by general agreement, lie beyond the reach of any subitizing mechanism so far hypothesized. The subitizing range is the range within which many researchers have hypothesized a direct perception of numerosity. This perception is assumed to yield a discrete symbolization of numerosity rather than an analog symbolization by mental magnitudes (Mandler & Shebo, 1982; Trick & Pylyshyn, 1994). It is also generally assumed that these numerical percepts ("twoness," "threeness") are not generated by any form of counting, verbal or nonverbal (Butterworth, 1999; Starkey & Cooper, 1995). On this hypothesis, it is unclear why the represen-
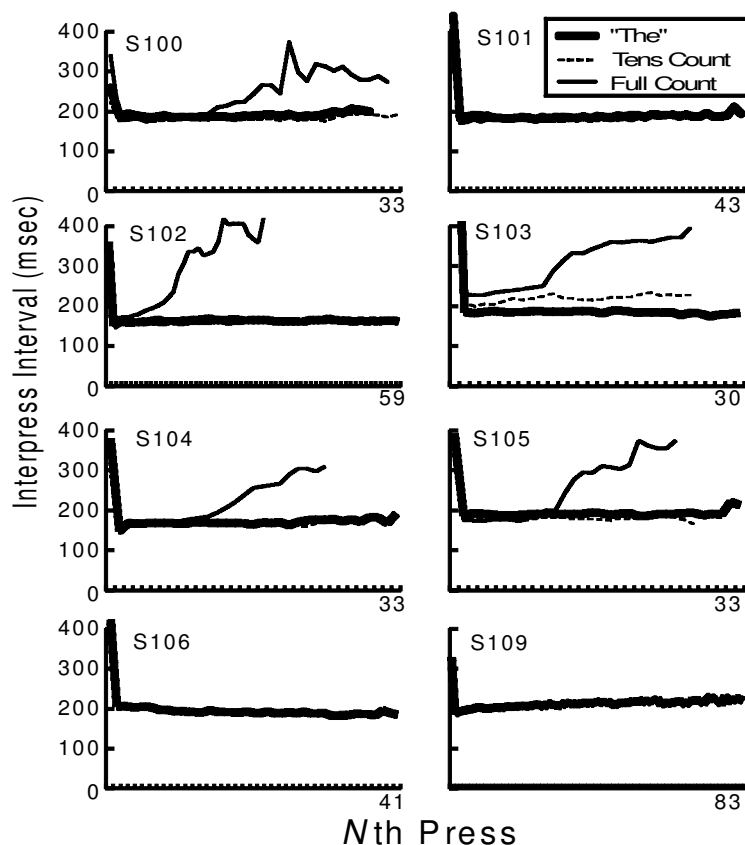
Figure 6. Mean interresponse interval (in milliseconds) versus *n*th key press for each subject in the "The," tens count, and full count conditions. Note that not every subject participated in every condition.

tation of the numbers between 2 and 5 should exhibit scalar variability. Indeed, a lack of variability is often taken to be a signature of this mechanism (Carey, 2001; Spelke, 2000). It is even less clear why the constant of proportionality relating the variability to the mean should be the same for numbers inside and outside the subitizing range. (For more on this topic, see Gelman, 1993; Gelman & Cordes, 2001; Gelman & Meck, 1992.)

In our task, a numeral indicated a target numerosity and subjects tried to produce a correspondingly numerous sequence of responses. The behavior-controlling representation of numerosity may have different properties when sets are small, presented simultaneously, and with different task demands, which is usually the case where subitizing has been posited (e.g., see Trick & Pylyshyn, 1994). We do not think, however, that data from numerosity-naming reaction time studies and symbolic distance and size studies support such a hypothesis. Rather, they suggest a common mode of representation throughout the single-digit range (see Balakrishnan & Ashby, 1992; Moyer & Landauer, 1967). In short, several lines of evidence favor the view that, at least in adults, numerosities of all sizes are represented by mental magnitudes, whether those magnitudes

are evoked by number words, numerals, sequentially occurring events, or simultaneously presented sets.

The Quantitative Relation Between Symbolic Numerosity and the Mental Magnitudes

Our task requires the comparison of a magnitude summoned from memory by a numeral (Process 1) and a magnitude generated, we suppose, by a nonverbal counting process (Process 2). For the number of presses made to approximate the number signified by the numeral, both processes must yield roughly similar magnitudes for similar numerosities— and they do, as Figure 2 shows. However, a close look at the plots shows interesting systematic departures from the expected relation. In particular, some subjects show a tendency to make an ever-increasing excess of presses over the number signified by the numerals (e.g., Subjects 102 and 109). This could be scalar memory error, which is well documented in the psychophysics of duration memory (Gibbon, Church, & Meck, 1984). However, it is also possible that there is a systematic distortion in Process 1 that is not present in Process 2. It is conceivable that the mapping between the objective dimension along which numerals are appropriately arrayed (the objective number line) and the
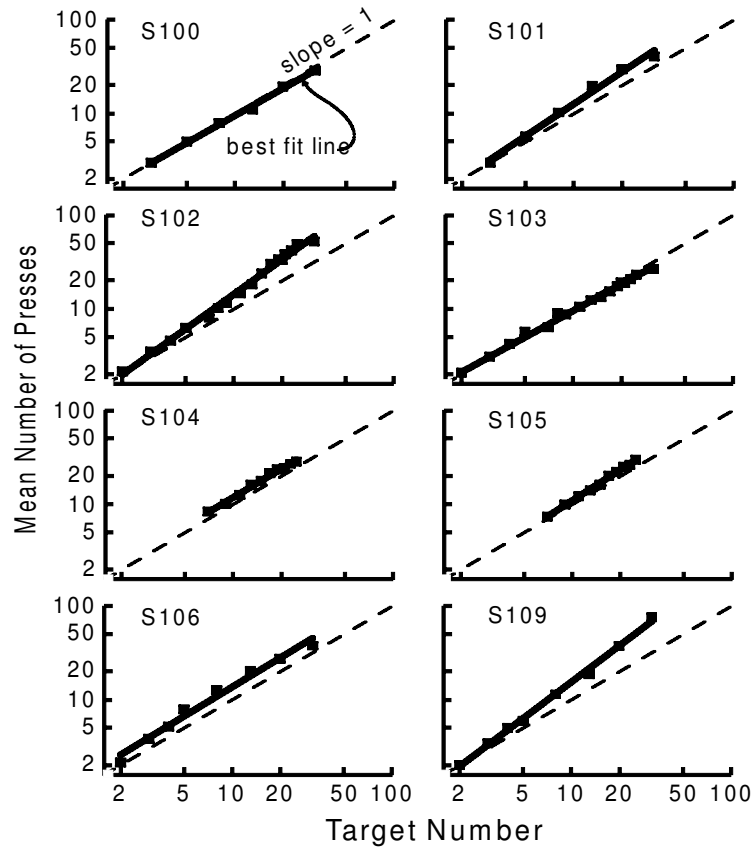
Figure 7. Mean number of key presses for each subject in the "The" (nonverbal counting) condition as a function of target number on double logarithmic coordinates. While most subjects produced data with slopes not significantly different from one, the data from 2 subjects (Subjects 102 and 109) are best fit by lines whose slopes are significantly greater than 1.

mental continuum that gives their numerical meaning (the subjective number line) obeys the power law that Stevens and his colleagues argued was characteristic of the mappings between objective stimulus dimensions and their mental counterparts (Stevens, 1956; Stevens & Harris, 1962).

The linear–linear plots in Figure 2 are not well suited to reveal subtle departures from linearity because the random (unsystematic) scatter in the means increases as the target number increases. Log–log plots are more likely to reveal subtle departures from linearity because they equate scatter about the regression line. Moreover, if the relation between the dimension of the numerals and the mental magnitudes generated by nonverbal counting is a power relation (the Stevens hypothesis), then log–log plots of the data will yield straight lines, with slopes equal to the power.

A scalar function is a power function with exponent 1. Thus, if the magnitudes that numerals summon from memory are proportional to the magnitudes generated by nonverbal counting, then the slopes of these log–log regression lines will not deviate significantly from 1. If the magnitudes summoned from memory are logarithmically compressed (Process 1 logarithmically compressed), but successive increments in nonverbal counting are equal (Process 2 linear), then this plot will not be straight, but will be downwardly accelerated.

The log–log plots of the mean number of key presses for each subject in the "The" condition are shown in Figure 7. In the majority of cases, they are well approximated by straight lines, and there are no trends discernible in the residuals.[6] In no case does one see the strong downward curvature to be expected on the logarithmic compression hypothesis. For 2 subjects (Subjects 102 and 109), however, the slopes are slightly but very significantly greater than 1, consistent with the Stevens hypothesis.

Like scalar variability, the power-law departures from linearity in our data extend into the subitizing range—another example of quantitative continuity. Like the continuity in measured variability, this continuity is not consonant with the hypothesis that small numbers are represented in a fundamentally different way than larger numbers—by discrete symbols rather than by mental magnitudes. It would also seem to be strong evidence that these subjects, at least, were not counting subvocally. If they were, their

mean count would not show a systematic nonlinear (accelerating) increase as the target value increased.

Our finding of power relations with non-unity exponents in some subjects suggests a generalization of the parameter-learning hypothesis of Joram, Subrahmanyam, and Gelman (1998). They argued that learning to think with conventional measures of, for example, length requires learning the scale factor relating the numbers generated by measurement to the mental magnitudes representing values on the measured dimension. On this parameter-learning hypothesis, the numbers generated by measuring constitute the *x*-axis of a mental graph (the arguments of a mental function); the mental magnitudes that represent the corresponding lengths constitute the *y*-axis (the range of the mental function); and the function is scalar (it has the form $y = kx$), with k being the parameter that must be learned. This hypothesis assumes that the conventional symbols are themselves represented mentally at equally spaced locations along a mental continuum. If, in accord with the generalized Stevens hypothesis, the template for mappings between mental continua is a power function, $y = kx^p$, then the construction of the bidirectional mapping between conventional symbols for numerosity and the mental magnitudes generated by nonverbal counting requires the learning of two parameters. We suggest that subjects learn to obtain from numerals and other conventional symbols for numerosity approximately appropriate mental magnitudes by learning the values of the two parameters (k and p) that specify the requisite power function. Our results suggest that there are small but measurable differences in the parameters that adult subjects learn.

REFERENCES

BALAKRISHNAN, J. D., & ASHBY, F. G. (1992). Subitizing: Magical numbers or mere superstition. *Psychological Research*, **54**, 80-90.

BRANNON, E. M., & TERRACE, H. S. (2000). Representation of the numerosities 1–9 by rhesus macaques (*Macaca mulatta*). *Journal of Experimental Psychology: Animal Behavior Processes*, **26**, 31-49.

BRANNON, E. M., WUSTHOFF, C. J., GALLISTEL, C. R., & GIBBON, J. (2001). Numerical subtraction in the pigeon: Evidence for a linear subjective number scale. *Psychological Science*, **12**, 238-243.

BUTTERWORTH, B. (1999). *The mathematical brain*. London: MacMillan.

CAREY, S. (1998). Knowledge of number: Its evolution and ontogeny. *Science*, **282**, 641-642.

CAREY, S. (2001). Cognitive foundations of arithmetic: Evolution and ontogenesis. *Mind & Language*, **16**, 37-55.

DEHAENE, S., DEHAENE-LAMBERTZ, G., & COHEN, L. (1998). Abstract representations of numbers in the animal and human brain. *Trends in Neuroscience*, **21**, 355-361.

DEHAENE, S., DUPOUX, E., & MEHLER, J. (1990). Is numerical comparison digital? Analogical and symbolic effects in two-digit number comparison. *Journal of Experimental Psychology: Human Perception & Performance*, **16**, 626-641.

FETTERMAN, J. G. (1993). Numerosity discrimination: Both time and number matter. *Journal of Experimental Psychology: Animal Behavior Processes*, **19**, 149-164.

GALLISTEL, C. R., & GELMAN, R. (1992). Preverbal and verbal counting and computation. *Cognition*, **44**, 43-74.

GALLISTEL, C. R., & GELMAN, R. (2000). Non-verbal numerical cognition: From reals to integers. *Trends in Cognitive Sciences*, **4**, 59-65.

GALLISTEL, C. R., GELMAN, R., & CORDES, S. (in press). The cultural

and evolutionary history of the real numbers. In S. Levinson & P. Jaisson (Eds.), *Culture and evolution*. Oxford: Oxford University Press.

GELMAN, R. (1993). A rational-constructivist account of early learning about numbers and objects. In D. Medin (Ed.), *Learning and motivation* (Vol. 30, pp. 61-96). Academic Press: New York.

GELMAN, R., & CORDES, S. (2001). Counting in animals and humans. In E. Dupoux (Ed.), *Language, brain and cognitive development: Essays in honor of Jacques Mehler* (pp. 279-303). Cambridge, MA: MIT Press.

GELMAN, R., & MECK, E. (1992). Early principles aid early but not later conceptions of number. In J. Bideaud, C. Meljac, & J. P. Fischer (Eds.), *Pathways to number* (pp. 171-189). Hillsdale, NJ.: Erlbaum.

GIBBON, J., & CHURCH, R. M. (1981). Time left: Linear versus logarithmic subjective time. *Journal of Experimental Psychology: Animal Behavior Processes*, **7**, 87-107.

GIBBON, J., CHURCH, R. M., & MECK, W. H. (1984). Scalar timing in memory. In J. Gibbon & L. Allan (Eds.), *Timing and time perception* (New York Academy of Sciences, Vol. 423, pp. 52-77). New York: New York Academy of Sciences.

HAUSER, M., & CAREY, S. (1998). Building a cognitive creature from a set of primitives: Evolutionary and developmental insights. In E. D. D. Cummins, E. C. Allen, et al. (Eds.), *The evolution of mind* (pp. 51-106). Oxford: Oxford University Press.

JORAM, E., SUBRAHMANYAM, K., & GELMAN, R. (1998). Measurement estimation: Learning to map the route from number to quantity and back. *Review of Educational Research*, **68**, 413-449.

LESLIE, A. M., XU, F., TREMOULET, P. D., & SCHOLL, B. (1998). Indexing and the object concept: Developing *What* and *Where* systems. *Trends in Cognitive Sciences*, **2**, 10-18.

LOGIE, R. H., & BADDELEY, A. D. (1987). Cognitive processes in counting. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **13**, 310-326.

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.

MOYER, R. S., & LANDAUER, T. K. (1967). Time required for judgments of numerical inequality. *Nature*, **215**, 1519-1520.

MOYER, R. S., & LANDAUER, T. K. (1973). Determinants of reaction time for digit inequality judgments. *Bulletin of the Psychonomic Society*, **1**, 167-168.

PLATT, J. R., & JOHNSON, D. M. (1971). Localization of position within a homogeneous behavior chain: Effects of error contingencies. *Learning & Motivation*, **2**, 386-414.

REYNVOET, B., & BRYSBAERT, M. (1999). Single-digit and two-digit Arabic numerals address the same semantic number line. *Cognition*, **72**, 191-201.

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.

SIMON, T. J. (1999). The foundations of numerical thinking in a brain without numbers. *Trends in Cognitive Sciences*, **3**, 363-364.

SPELKE, E. S. (2000). Core knowledge. *American Psychologist*, **55**, 1233-1243.

STARKEY, P., & COOPER, R. G. (1995). The development of subitizing in young children. *British Journal of Developmental Psychology*, **13**, 399-420.

STEVENS, S. (1956). The direct estimation of sensory magnitudes—Loudness. *American Journal of Psychology*, **69**, 1-25.

STEVENS, S., & HARRIS, J. R. (1962). The scaling of subjective roughness and smoothness. *Journal of Experimental Psychology*, **64**, 489-494.

TRICK, L. M., & PYLYSHYN, Z. W. (1994). Why are small and large numbers enumerated differently? A limited-capacity preattentive stage in vision. *Psychological Review*, **101**, 80-102.

WASHBURN, D. A. (1994). Stroop-like effects for monkeys and humans: Processing speed or strength of association? *Psychological Science*, **5**, 375-379.

WHALEN, J., GALLISTEL, C. R., & GELMAN, R. (1999). Nonverbal count-

ing in humans: The psychophysics of number representation. *Psychological Science*, **10**, 130-137.

NOTES

1. Since subject data obtained during the different nonverbal counting sessions did not differ significantly [$t(14) = 0.404$, $t(14) = 1.18$, $p > .05$], data from both sessions were combined for each of these subjects (Subjects 102 and 103, respectively).

2. Again, the data from both sessions within each condition for Subject 102 were combined since they did not differ significantly [$t(14) = 1.37$, $t(14) = 0.87$, $p > .05$ for the full count and tens count conditions, respectively].

3. This condition was included since most subjects in the full count condition found it difficult to count rapidly using "teens" and "twenties" words. It is an interesting question, how subjects retain the tens count when using this strategy. We speculate that they retain it as a nonverbal magnitude.

4. In cases where the joystick button incorrectly registered two presses for one (which resulted in presses much closer together than humanly possible), the computer program was unable to keep up and registered an illegal press, indicated by a negative interresponse time. These button presses, clearly segregated from real presses by the interresponse time, were excluded from response time analyses. Infrequently, this occurred on the initial key press of a trial. In that case, the target number would flash on the screen too quickly to be seen by the subject. When that happened, the trial was ended immediately and these data points were excluded from analyses. In addition, outlier data points, ones more than three standard deviations away from the mean, were excluded from analyses.

5. For purposes of analyses in the logarithmic domain, zero values were excluded because the logarithm of x tends to minus infinity as x tends to zero.

6. Two sessions resulted in data that were not well fit by a straight line on a log–log plot; a small downward curvature was noted in both graphs. These data are, however, better fit by a power function than by a log function.