In K Holyoak & R. Morrison (Eds) The Cambridge handbook of thinking and reasoning. Cambridge University Press (pp 559-588) 2005

Mathematical Cognition

C. R. Gallistel
Rochel Gelman
Rutgers University[1]

Mathematics is a system for representing and reasoning about quantities, with arithmetic as its foundation. Its deep interest for our understanding of the psychological foundations of scientific thought comes from what Eugene Wigner called the unreasonable efficacy of mathematics in the natural sciences. From a formalist perspective, arithmetic is a symbolic game, like tic-tac-toe. Its rules are more complicated, but not a great deal more complicated. Mathematics is the study of the properties of this game and of the systems that may be constructed on the foundation that it provides. Why should this symbolic game be so powerful and resourceful when it comes to building models of the physical world? And on what psychological foundations does the human mastery of this game rest?

The first question is metaphysical—why is the world the way it is? We do not treat it, because it lies beyond the realm of experimental behavioral science. We review the answers to the second question that experimental research on human and non-human animal cognition suggests.

The general nature of the answer is that the foundations of mathematical cognition appear does not lie in language and the language faculty. The ability to estimate quantities and to reason arithmetically with those estimates exists in the brains of animals that have no language. The same or very similar non-verbal mechanisms appear to operate in parallel with verbal estimation and reasoning in adult humans. And, they operate to some extent before children learn to speak and before they have had any tutoring in the elements of arithmetic. These findings suggest that the verbal expression of number and of arithmetic thinking is based on a non-verbal system for estimating and reasoning about discrete and continuous quantity, which we share with many non-verbal animals. A reasonable supposition is that the neural substrate for this system arose far back in the evolution of brains precisely because of the puzzle that Wigner called attention to: arithmetic reasoning captures deeply important properties of the world, which the animal brain must represent in order to act effectively in it.

1

Department of Psychology
Rutgers Center for Cognitive Science
Rutgers University
152 Frelinghuysen Rd
Piscataway, NJ 08854-8020
email: galliste@ruccs.rutgers.edu
rgelman@ruccs.rutgers.edu

The authors' research reported here was supported by the following grants: NSF DFS-920974 to R Gelman and SRB-97209741 to R Gelman et al., NIH Grant MH 63866 to C.R. Gallistel.

The recognition that there is a nonverbal system of arithmetic reasoning in human and many nonhuman animals is recent, but it influences most contemporary experimental work on mathematical cognition. This review is organized around the questions: 1) What are the properties of this non-verbal system? 2) How is it related to the verbal system and written numerical systems?

What is a Number?

Arithmetic is one of the few domains of human thought that has been extensively formalized. This formalization did not begin in earnest until the middle of the 19th century (Boyer & Merzback: , 1989). In the process of formalizing the arithmetic foundations of mathematics, mathematicians changed their minds about what a number was. Before formalization, an intuitive understanding of what a number was determined what could legitimately be done with it. Once the formal "games" about number were made explicit, anything that played by the rules was a number.

This formalist viewpoint is crucial to an understanding of issues in the current scientific literature on mathematical cognition. Many of them turn on questions of how we are to recognize and understand the properties of mental magnitudes. *Mental magnitude* refers to an inferred (but, one supposes, potentially observable and measurable) entity in the head that represents either numerosity (for example, the number of oranges in a case) or another magnitude (for examples, the length, width, height and weight of the case) and that has the formal properties of a real number.

For a mental magnitude to represent an objective magnitude, it must be causally related to that objective magnitude. It must also be shown that it is a player in mental game (a functionally cohesive collection of brain processes) that operates according to at least some of the rules of arithmetic. When putative mental numbers do not validly enter into, at a minimum, mental addition, mental subtraction and mental ordering, then they do not function as numbers.

Kinds of Numbers

The ancient Greeks had considerable success axiomatizing geometry, but mathematicians did not axiomatize the system of number until the 19th century, after it had undergone a large, historically documented expansion. Before this expansion, it was too messy and incomplete to be axiomatized, because it lacked *closure*. A system of number is closed under a combinatorial operation if, when you apply the operation to any pair of numbers, the result is a number. Adding or multiplying two positive integers always produces a positive integer, so the positive integers are closed under addition and multiplication. They are also closed under the operation of ordering. For any pair of numbers $a \geq b = 1$ if a is greater than b, and 0 if not. These three operations—addition, multiplication and ordering—are the core operations of arithmetic. They and their inverses make the system what it is.

The problem comes from the inverse operations of subtraction and division. When you subtract a bigger number from a smaller, the result is not a positive integer. Should one regard the result as a number? Until well into the 19th century, many professional mathematicians did not. Thus, subtracting a bigger number from a smaller number was not a legitimate mathematical operation. This was inconvenient, because it meant that in the course of algebraic reasoning (reasoning about unspecified numbers), one might unwittingly do something that was illegitimate. This purely practical consideration strongly motivated the admission of the negative numbers and 0 to the set of numbers acknowledged to be legitimate.

When one divides one integer by another, the result, called a rational number, or, more colloquially, a fraction, is rarely an integer. From the earliest times from which we have written records, people who worked with written numbers included at least some rational numbers among the numbers, but, like school children to this day, they had extraordinary difficulties in figuring out how to do arithmetic with rational numbers in general. What is the sum of 1/3 and 11/17? That was a hard question in ancient Egypt and remains so today, in school classrooms all over the world.

The common notation for a fraction specifies a number not by giving it a unique name like *two*, but rather by specifying a way of generating it ("divide the number *one* by the number *two*). The practice of specifying a number by giving an arithmetic procedure that will generate it to whatever level of precision is required has grown stronger over the millenia. It is the key to a rigorous handling of both irrational and complex numbers, and to the way in which digital computers operate with real numbers. But it is discomfiting, for several reasons. First, there are an infinity of different notations for the same number: 1/2, 2/4, 3/6, and so on, all specifying the same number. Moreover, for most rational numbers, there is no complete decimal representation. Carrying out the division gives a repeating decimal. In short, you cannot write down a symbol for most rational numbers that is both complete and unique.[2]

Finally, when fractions are allowed to be numbers, the *discrete* ordering of the numbers is lost. It is no longer possible to specify the next number in the sequence, because there is an infinite number of rational numbers between any two rational numbers. For all these reasons, admitting fractions to the system of number makes the system more difficult to work with in the concrete, albeit more powerful in the abstract, because the system of rational numbers is, with one exception, closed under division.

Allowing negative numbers and fractions to be numbers also creates problems with what otherwise seem to be sound principles for reasoning about numbers. For example, it seems to be sound to say that dividing the bigger of two numbers by the smaller gives a number that is bigger than the number one

[2] Technically, this is not really true, because Cantor discovered a way to assign a unique positive integer to every rational number. The integers his procedure assigns are, however, useless for computational purposes

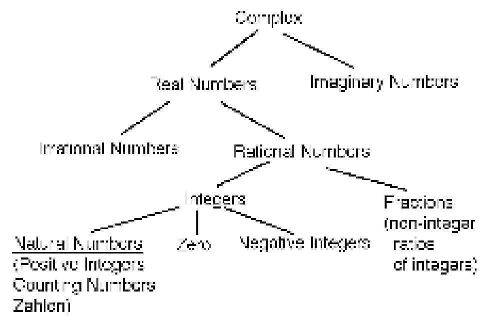
gets if one divides the smaller by the bigger. What then are we to make of the "fact" that $1/-1 = -1/1 = -1$?

Clearly, caution and clear thinking are going to be necessary if we want to treat as numbers entities that you do not get by counting. But, humans do want to do this, and they have wanted to since the beginning or recorded history. We measure quantities like lengths, weights, and volumes in order to represent them with numbers. What the measuring does—if it is done well—is give us "the right number" or at least one useable for our purposes. Measuring and the resulting representation of continuous quantities by numbers goes back to the earliest written records. Indeed, it is often argued that writing evolved from a system for recording the results of measurements made in the course of commerce (bartering, buying and selling), political economy (taxation), surveying and construction (Menninger, 1969).

The ancient Greeks believed that all measurable magnitudes could in principle be represented by rational numbers. Everything was a matter of proportion and any proportion could be expressed as the ratio of two integers. They were also the first to try to formalize mathematical thinking. In doing so, they discovered, to their horror, that fractions did not suffice to represent all possible proportions. They discovered that the proportion between the side of a square and its diagonal could not be represented by a fraction. The Pythagorean formula for calculating the diagonal of a square says that the diagonal is equal to the square root of the sum of the squares of the sides. In this case, the diagonal is equal to $\sqrt{(1^2 + 1^2)} = \sqrt{(1+1)} = \sqrt{2}$. The Greeks proved that there is no fraction that when multiplied by itself is equal to 2. If only the integers and the fractions are numbers, then the length of the diagonal of the unit square cannot be represented by a number. Put another way, you can measure the side of the square or you can measure its diagonal, but you cannot measure them both exactly within the same measuring system—unless you are willing to include among the numbers in that system numbers that are not integers (cannot be counted) and are not even the ratio of two integers. You must include what the Greeks called the irrational numbers. But if you do include the irrational numbers, how do you go about specifying them in the general case?

Many irrationals can be specified by the operation of extracting roots, which is the inverse of the operation of raising a number to a power. Raising any positive integer to the power of any other always produces a positive integer. Thus, the system of positive integers is closed under raising to a power. The problem, as usual, comes from the inverse operation, extracting roots. For most pairs of integers, a and b, the ath root of b is not a positive integer, nor even a rational number; it is an irrational number. The need within algebra to have an arithmetic that was closed under the extraction of roots was a powerful motivation for mathematicians to admit both the irrational numbers and complex numbers to the set of numbers. By admitting the irrational numbers, they created the system of so-called real numbers, which was essential to the calculus. To this day there are professional mathematicians who question the legitimacy of the irrational numbers. Nonetheless, the real numbers, which include the irrationals (see Figure 1), are taken for granted by all but a very few contemporary mathematicians.

Figure 1. *The number system on which modern mathematics is based. Not shown in this diagram are the algebraic numbers, which are the numbers that may be obtained through the extraction of roots (the solving of polynomial equations) nor the transcendental numbers, which may be obtained only by the solving of an equations with trigonometric, exponential or logarithmic terms. These are subcategories of the irrational numbers.*



The notion of a real number and the notion of a magnitude (for example, the length of a line) are formally identical. This means among other things that, for every line segment, there is a real number that uniquely represents the length of that segment (in a given system of measurement) and conversely, for every real number, there is a line segment that represents the magnitude of that number. Therefore, in what follows, when we mention a mental magnitude, we mean an entity in the mind (brain) that functions within a system with the formal properties of the real number system. Like the real number system, we assume that this system is a closed system: all of its combinatorial operations, when applied to any pair of mental magnitudes, generate another mental magnitude.

As this brief sketch indicates, the system of number recognized by almost all contemporary professional mathematicians as "*the* number system"—the ever more inclusive hierarchy of kinds of numbers shown in Figure 1—has grown up over historical time, with much of the growth culminating only in the preceding two centuries. The psychological question is, what is it in the minds of humans (and perhaps also non-human animals) that has been driving this process? And how and under what circumstances does this mental machinery enable educated modern humans to master the basics of formal mathematics, when, and to the extent that, they do so?

Numerical Estimation and Reasoning in Animals

The development of verbalized and written reasoning about number that culminated in a formalized system of real numbers isomorphic to continuous magnitudes was driven by the fact that humans apply numerical reasoning to continuous quantity just as much as they do to discrete quantity. In considering the literature on numerical estimation and reasoning in animals, we begin by reviewing the evidence that they estimate and reason arithmetically about the quintessentially continuous quantity, time.

The common laboratory animals such as the pigeon, the rat and the monkey, measure and remember continuous quantities, such as duration, as has been shown in a variety of experimental paradigms. One of these is the so-called peak procedure. In this procedure, a trial begins with the onset of a stimulus signaling the possible availability of food at the end of a fixed interval, called the feeding latency. Responses made at or after the interval has elapsed trigger the delivery of food. Response prior to that time have no consequences. On 20-50% of the trials, food is not delivered. On these trials, the key remains illuminated or the lever remains extended or the hopper remains illuminated for between 4 and 6 times longer than the feeding latency. On these trials, called probe trials, responding after the feeding latency has past is pointless.

Peak-procedure data come from these unrewarded trials. On such trials, the subject abruptly begins to respond some while before the interval ends (in anticipation of its ending) and continues to peck or press or poke for some while after, before abruptly stopping. The interval during which the subject responds brackets its subjective estimate of the fixed interval. Representative data are shown in Figure 2.

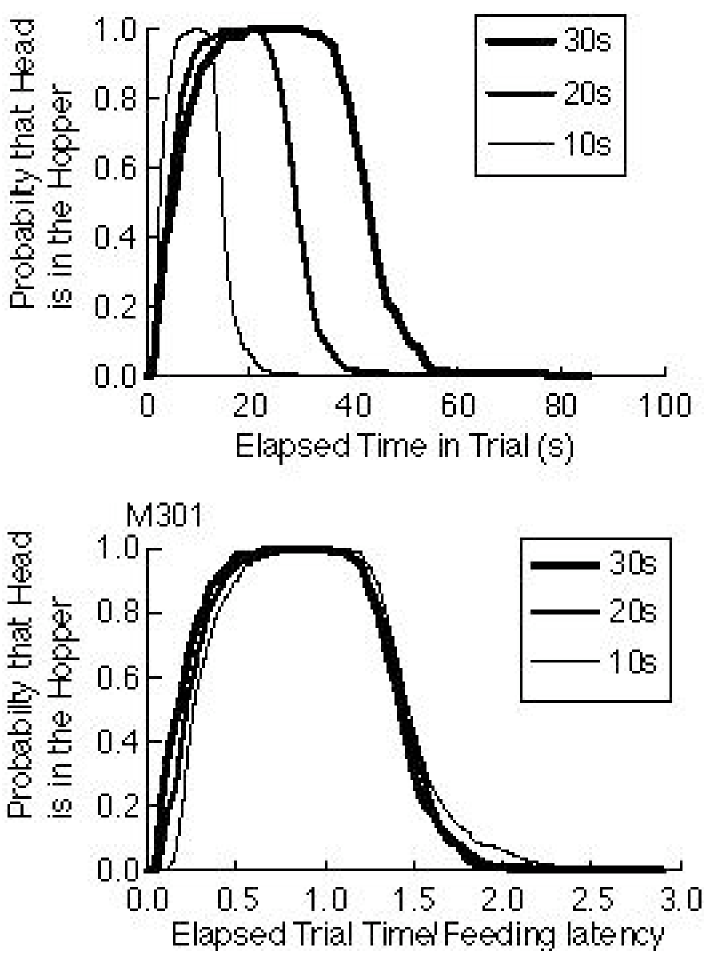


Figure 2. *Representative peak procedure data: Probability that the mouse's head was in the feeding hopper as a function of the time elapsed since the beginning of a trial and the feeding latency. (The feeding latency varied between blocks of trials.) A. The original data. These peak curves are the cumulative distribution of start times (rising phase) minus cumulative distribution of stop times (falling phase). These are the raw distributions (no curve has been fitted.) B. Data replotted as a proportion of the feeding latency. Because the variability in the onsets and offsets of responding is proportional to the feeding latency, as are the location of the means of the distributions relative to the target times, the peak curves superpose when plotted as a proportion of this latency. Data originally* published by (King, McDonald, & Gallistel, 2001).

Figure 2A, shows seemingly smooth increases and decreases in the probability that the mouse is making an anticipatory response (poking its head into the feeding hopper in anticipation of food delivery) on either side of the feeding latency. The smoothness is an averaging artifact. On any one trial the onset and offset of anticipatory responding is abrupt, but the temporal locus of these onsets and offsets varies from trial to trial (R.M. Church, Meck, & Gibbon, 1994). The peak curves in Figure 2, like peak curves in general, are the cumulative start distributions minus the cumulative stop distributions, where start and stop refer to the onset and offset of sustained food anticipatory behavior.

When the data in Figure 2A are replotted against the proportion of the feeding latency elapsed, rather than against the latency itself, the curves superpose (Figure 2B). Thus, both the location of the distributions relative to the target latency and the trial-to-trial variability in the onsets and offsets of responding is proportional to the remembered latency. Put another way, the probabilities that the subject will have begun to respond or will have stopped responding are determined by the proportion of the remembered arming latency that has elapsed. This property of remembered durations is called <u>scalar variability</u>.

Rats, pigeons and monkeys also count and remember numerosities (Elizabeth M. Brannon & Roitman, 2003; R. M. Church & Meck, 1984; S. Dehaene, 1997; S. Dehaene, Dehaene-Lambertz, & Cohen, 1998; Gallistel, 1990; Gallistel & Gelman, 2000). One of the early protocols for assessing counting and numerical memory was developed by Mechner (1958) and later used by Platt and Johnson (1971). The subject must press a lever some number of times (the target number) in order to arm the infrared beam at the entrance to a feeding alcove. When the beam is armed, interrupting it releases food. Pressing too many times before trying the alcove incurs no penalty beyond that of having made supernumerary presses. Trying the alcove prematurely incurs a 10-second time-out, which the subject must endure before returning to the lever to complete the requisite number of presses. Data from such an experiment are shown in Figure 3. They look strikingly like the temporal data. The number of presses at which subjects are maximally likely to break off pressing and try the alcove peaks at or slightly beyond the required number, for required numbers ranging from 4 to 24. As the remembered target number gets larger, the variability in the break-off number also gets proportionately greater. Thus, behavior based on number also exhibits scalar variability

The fact that behavior based on remembered numerosity exhibits scalar variability just like the scalar variability seen in behavior based on the remembered magnitude of continuous quantities like duration suggests that numerosity is represented in the brains of non-verbal vertebrates by mental magnitudes, that is by entities with the formal properties of the real numbers, rather than by discrete symbols like words or bit patterns. When a device such as an analog computer represents numerosities by different voltage levels, noise in the voltages leads to confusions between nearby numbers. If, by contrast, a device represents countable quantity by countable (that is, discrete) symbols, as digital computers and written number systems do, then one does not expect to see the kind of variability seen in Figures 2 and 3. For example, the bit-pattern symbol for fifteen is 01111 while for sixteen it is 10000. Although the numbers are adjacent in the ordering of the integers, the discrete binary symbols for them differ in all five bits. Jitter in the bits (uncertainty about whether a given bits was 0 or 1) would make fourteen (01110), thirteen (01101), eleven (01011) and seven (00111) all equally and maximally likely to be confused with fifteen, because the confusion arises in each case from the misreading of one bit. These dispersed numbers should be confused with fifteen much more often than is the adjacent sixteen. (For an analysis of the error patterns to be expected in cascade counters, see Killeen & Taylor, 2001). Similarly, a scribe copying a handwritten English text is presumably more likely to confuse "seven" and "eleven" than "seven" and "eight". Thus, the nature of the variability in a remembered target number suggests that what is being remembered is a magnitude, something that behaves like a continuous quantity, which is to say something with the formal properties of a real number.

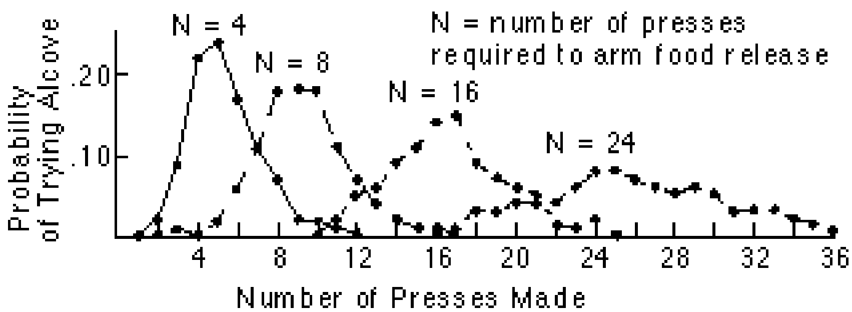


Figure 3. *The probability of breaking off to try the feeding alcove as a function of the number of presses made on the arming lever and the number required to arm the food-release beam at the entrance to the feeding alcove. Subjects were rats. (Redrawn from Platt & Johnson, 1971, by permission of the authors and publishers.)*

Numerosity and Duration are Represented by Comparable Mental Magnitudes

Meck and Church (1983) pointed out that the mental accumulator model that Gibbon (1977) had proposed to explain the generation of mental magnitudes representing durations could be modified to make it generate mental magnitudes representing numerosities. Gibbon had proposed that while a duration was being timed a stream of impulses fed an accumulator, so that the accumulation grew in proportion to the duration of the stream. When the stream ended (when timing ceased), the resulting accumulation was read into memory, where it represented the duration of the interval. Meck and Church postulated that to get magnitudes representing numerosity, the equivalent of a pulse former was inserted into the stream of impulses, so that for each count there was a discrete increment in the contents of the accumulator, as happens when a cup of liquid is poured into a graduate (see Figure 4). At the end of the count, the resulting accumulation is read into memory where it represents the numerosity.

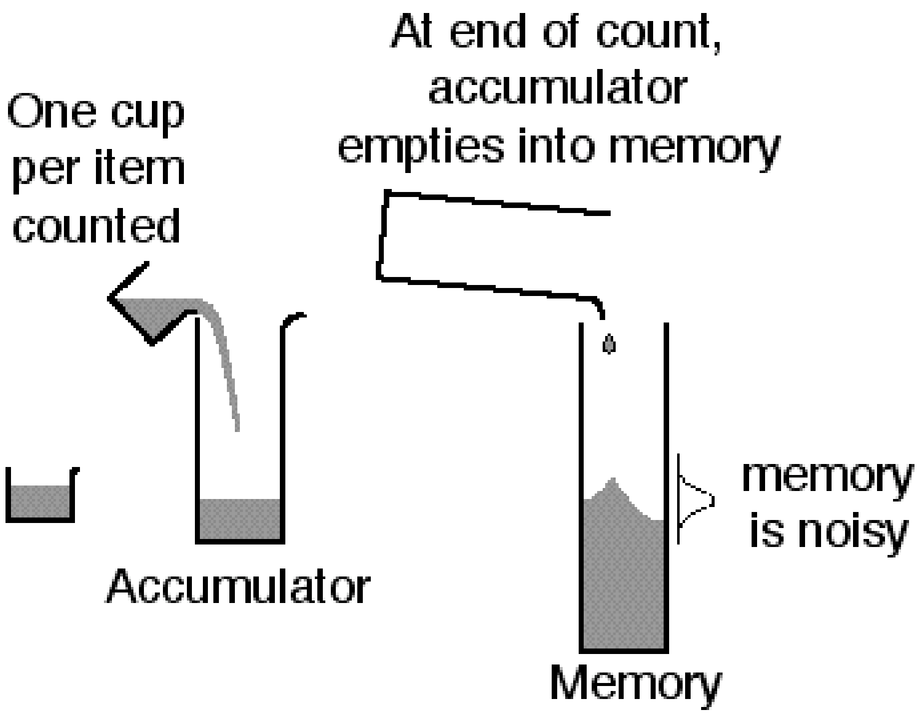


Figure 4. *The accumulator model for the non-verbal counting process. At each count, the brain increments a quantity, an operation formally equivalent to pouring a cup into a graduate. The final magnitude (the contents of the graduate at the conclusion of the count) is stored in memory, where it represents the numerosity of the counted set. Memory is noisy (represented by the wave in the graduate), which is to say that the values read from memory on different occasions vary. The variability in the values read from memory is proportional to the mean value of the distribution (scalar variability).*

The model in Figure 4 is the well known accumulator model for non-verbal counting by the successive incrementation of mental magnitudes. It is also the origin of the hypothesis that the mental magnitudes representing duration and the mental magnitudes representing numerosity are essentially the same, differing only in the mapping process that generates them and, hence, in what it is they refer to. Put another way, both numerosity and duration are represented mentally by real numbers. Meck and Church (1983) compared the psychophysics of number and time representation in the rat and concluded that the coefficient of variation, the ratio between the standard deviation and the mean, was the same, which is further evidence for the hypothesis that the same system of real numbers is used in both cases.

The model in Figure 4, was originally proposed to explain behavior based on the numerosity of a set of serial events (for example, the number of responses made), but it may be generalized to the case where the items to be counted are presented all at once, for example, as a to-be-enumerated visual array. In that case, each item in the array can be assigned a unit magnitude, and the unit magnitudes can then be summed (accumulated) across space, rather than over time. Dehaene and Changeux (1993) developed a neural net model based on this idea. In their model, the activity aroused by each item in the array is reduced to a unit amount of activity, so that it is no longer proportional to the size, contour, etc of the item. The units of activity corresponding to the entities in the array are summed across the visual field to yield a mental magnitude representing the numerosity of the array.

Non-Human Animals Reason Arithmetically

We have repeatedly referred to the real number system because numbers (or magnitudes) are truly that only if they are arithmetically manipulated. Being causally connected to something that can be represented numerically, does not make an entity in the brain or anywhere else a number. It must also be suitably processed. The defining features of a numerical representations are: 1) There is a causal mapping from discrete and continuous quantities in the world to the numbers. 2) The numbers are arithmetically processed. 3) The mapping is usefully (validly) invertible: the numbers obtained through arithmetic processing correctly refer through the inverse mapping back to the represented reality.

There is a considerable experimental literature demonstrating that laboratory animals reason arithmetically with mental magnitudes representing numerosity and duration. They add, subtract, divide and order subjective durations and subjective numerosities; they divide subjective numerosities by subjective durations to obtain subjective rates of reward; and they multiply subjective rates of reward by the subjective magnitudes of the rewards to obtain subjective incomes. Moreover, the mapping between real magnitudes and their subjective counterparts is such that their mental operations on subjective quantities enable these animals to behave effectively. Here we summarize a few of the relevant studies. (For reviews, see Sarah T. Boysen & Hallberg, 2000; Elizabeth M. Brannon & Roitman, 2003; S. Dehaene, 1997; Gallistel, 1990; Spelke & Dehaene, 1999)

Adding numerosities

Boysen and Berntson (1989) taught chimpanzees to pick the Arabic numeral corresponding to the number of items they observed. In the last of a series of tests of this ability, they had their subjects go around a room and observe either caches of actual oranges in two different locations or Arabic numerals that substituted for the caches themselves. When they returned from a trip, the chimps picked the Arabic numeral corresponding to the sum of the two numerosities they had seen, whether the numerosities had been directly observed (hence, possibly counted) or symbolically represented (hence not counted). In the latter case, the magnitudes corresponding to the numerals observed were presumably retrieved from a memory map relating the arbitrary symbols for number (the Arabic numerals) to the mental magnitudes that naturally represent those numbers. Once retrieved, they could be added very much like the magnitudes generated by the non-verbal counting of the caches. For further evidence that nonverbal vertebrates sum numerical magnitudes (see Beran, 2001; R. M. Church & Meck, 1984; Hauser, 2001, and citations therein;

Olthof, Iden, & Roberts, 1997; Olthof & Roberts, 2000; D. M. Rumbaugh, Savage-Rumbaugh, & Hegel, 1987).

Subtracting durations and numerosities

On each trial of the time-left procedure (Gibbon & Church, 1981), subjects are offered an ongoing choice between a steadily diminishing delay, on the one hand (the time-left option), and a fixed delay, on the other hand (the standard option). At an unpredictable point in the course of a trial, the opportunity to choose ends. Before it gets its reward, the subject must then endure the delay associated with the option it was exercising at that moment. If it was responding at the so-called standard station, it must endure the standard delay; if it was responding at the time-left station, it must endure the time left. At the beginning of a trial, the time left is much longer than the standard delay, but it grows shorter as the trial goes on, because the time so far elapsed in a trial is subtracted from the initial value to yield the time left. When the subjective time left is less than the subjective standard, subjects switch from the standard option to the time-left option. The subjective time left is the subjective duration of a remembered initial duration (subjective initial duration) minus the subjective duration of the interval elapsed since the beginning of the trial. Thus, in this experiment subjects' behavior depends on the subjective ordering of a subjective difference and a subjective standard (two of the basic arithmetic operations).

In the number-left procedure (E.M. Brannon, Wusthoff, Gallistel, & Gibbon, 2001), pigeons peck a center key in order both to generate flashes and to activate two choice keys. The flashes are generated on a variable ratio schedule, which means that the number of pecks required to generate each flash varies randomly between one and eight. When the choice keys are activated, the pigeons can get a reward by pecking either of them, but only after their pecking generates the requisite number of flashes. For one of the choice keys, the so-called standard key, the requisite number is fixed and independent of the number of flashes already generated. For the other choice, the number-left key, the requisite number is the difference between a fixed starting number and the tally of flashes already generated by pecking the center key. The flashes generated by pecking a choice key are also delivered on a variable ratio schedule.

The use of variable ratio schedules for flash generation partially dissociates time and number. The number of pecks required to generate any given number of flashes--and, hence, the amount of time spent pecking--varies greatly from trial to trial. This makes possible an analysis to determine whether subjects' choices are controlled by the time spent pecking the center key or by the number of flashes thus generated. The analysis shows that it was number not duration that controlled the pigeons' choices.

In this experiment, subjects chose the number-left key when the subjective number left was less than some fraction of the subjective number of flashes required on the standard key. Thus, their behavior was controlled by the subjective ordering of a subjective numerical difference and a subjective numerical standard. For an example of spontaneous subtraction in monkeys, see Sulkowski and Hauser (2001).

There is also evidence that the mental magnitudes representing duration and rates are signed, that is, there are both positive and negative mental magnitudes (Gallistel & Gibbon, 2000; Savastano & Miller, 1998). In other words, there is evidence not only for subtraction but also for the hypothesis that the system for arithmetic reasoning with mental magnitudes is closed under subtraction.

Dividing Numerosity by Duration

When vertebrates from fish to humans are free to forage in two different nearby locations, moving back and forth repeatedly between them, the ratio of the expected durations of the stays in the two locations matches the ratios of the numbers of rewards obtained per unit of time (R. J. Herrnstein, 1961). Until recently, it had been assumed that this matching behavior depended on the law of effect. When subjects do not match, they get more reward per unit of time invested in one patch than per unit of time invested in the other. Only when they match do they get equal returns on their investment. Thus, matching could be explained on the assumption that subjects try different ratios of investments (different ratios of expected stay durations) until they discover the ratio that equates the returns (R.J. Herrnstein & Vaughan, 1980).

Gallistel, Mark, King and Latham (2001) have shown that rats adjust to changes in the scheduled rates of reward as fast as it is in principle possible to do so; they are ideal detectors of such changes. They could not adjust so rapidly if they were discovering by trial and error the ratio of expected stay durations that equated their returns. The importance of this in the present context is that a rate is the number of events, that is, a discrete or countable quantity, which is the kind of thing naturally represented by the positive integers, divided by a continuous or (uncountable) quantity--the duration of the given interval—which is the kind of thing that can only be represented by a real number.

Gallistel and Gibbon (2000) review the evidence that both Pavlovian and instrumental conditioning depend on subjects' estimating rates of reward. They argue that rate of reward is the fundamental variable in conditioned behavior. The importance of this in the present context is twofold. First, it is evidence that subjects divide mental magnitudes. Second, it shows why it is essential that countable and uncountable quantity be represented by commensurable mental symbols, symbols that are part of the same system and can be arithmetically combined without regard to whether they represent countable or uncountable quantity. If countable quantity were represented by one system (say, a system of discretely ordered symbols, formally analogous to the list of counting words) and uncountable (continuous) quantity by a different system (a system of mental magnitudes), it would not be possible to estimate rates. The brain would have to divide mental apples by mental oranges[3].

Multiplying Rate by Magnitude

When the magnitudes of the rewards obtained in two different locations differ, then the ratio of the expected stay durations is determined by the ratio of the incomes obtained from the two locations (Catania, 1963; Harper, 1982; Keller & Gollub, 1977; Leon & Gallistel, 1998). The income from a location is the product of the rate and the magnitude. Thus, this result implies that subjects multiply subjective rate by subjective magnitudes to obtain subjective incomes. The signature of multiplicative combination is that changing one variable by a given factor--for example, doubling the rate, changes the product by the same factor (doubles the income) regardless of the value of the other factor (the magnitude of the rewards). Leon and Gallistel (1998) showed that changing the ratio of the rates of reward by a given factor changed the ratio of the expected stay durations by that factor, regardless of the ratio of the reward magnitudes, thereby proving that subjective magnitudes combine multiplicatively with subjective rates to determine the ratio of expected stay durations.

Ordering Numerosities

Most of the paradigms that demonstrate mental addition, subtraction, multiplication and division also demonstrate the ordering of mental magnitudes, because the subject's choice depends on this ordering. Brannon and Terrace (2000) demonstrated more directly that monkeys order numerosities by presenting simultaneously several arrays differing in the numerosity of the items constituting each array and requiring their macaque subjects to touch the arrays in the order of their numerosity. When subjects had learned to do this for numerosities between one and four, they generalized immediately to numerosities between five and nine.

The most interesting feature of Brannon and Terrace's results was that they found it impossible to teach subjects to touch the arrays in an order that did not conform to the order of the numerosities (either ascending or descending). This implies that the ordering of numerosities is highly salient for a monkey. It cannot ignore their natural ordering in order to learn an unnatural one. It also suggests that the natural ordering is not itself learned; it is inherent in the monkey's representation of numerosity. What is learned is to respond on the basis of numerical order, not the ordering itself.

For further evidence that non-verbal vertebrates order numerosities and durations, see (Biro & Matsuzawa, 2001; Elizabeth M. Brannon & Roitman, 2003; Elizabeth M. Brannon & Terrace, 2002; Carr & Wilkie, 1997; Olthof et al., 1997; D.M. Rumbaugh & Washburn, 1993; Washburn & Rumbaugh, 1991).

---

[3] Fortran and C programmers, who have made the mistake of dividing an integer variable by a floating point variable will know whereof we speak

In summary, research with vertebrates, some of which have not shared a common ancestor with man since before the rise of the dinosaurs, implies that they represent both countable and uncountable quantity by means of mental magnitudes. The system of arithmetic reasoning with these mental magnitudes is closed under the basic operations of arithmetic, that is, mental magnitudes may be mentally added, subtracted, multiplied, divided and ordered without restriction.

Humans Also Represent Numerosity with Mental Magnitudes

The Symbolic Size and Distance Effects

It would be odd if humans did not share with their remote vertebrate cousins (pigeons) and near vertebrate cousins (chimpanzees) the mental machinery for representing countable and uncountable quantity by means of a system of real numbers. That humans do represent integers with mental magnitudes was first suggested by Moyer and Landauer (1967; 1973) when they discovered what has come to be called the symbolic distance effect (Figure 5). When subjects are asked to judge the numerical order of Arabic numerals as rapidly as possible, their reaction time is determined by the relative numerical distance: the greater the distance between the two numbers, the more quickly their order may be judged. Subsequently, Parkman (1971) showed further that the greater the numerical value of the smaller digit, the longer it takes to judge their order (the size effect). The two effects together may be summarized under a single law, namely that the time to judge the numerical order of two numerals is a function of the ratio of the numerical magnitudes that they represent. Thus, Weber's law applies to symbolically represented numerical magnitude. Weber's law is that the discriminability of two magnitudes is a function of their ratio.

The size and distance effects in human judgments of the ordering of discrete and continuous quantities are robust. They are observed when the numerosities being compared are actually instantiated (by visual arrays of dots) and when they are represented symbolically by Arab numerals (Buckley & Gillman, 1974, see Figure 15). The *symbolic* distance and size effects are observed in the single digit range and in the double digit range (S Dehaene, Dupoux, & Mehler, 1990; Hinrichs, Yurko, & Hu, 1981). That this effect of numerical magnitude on the time to make an order judgment should appear for symbolically represented numerosities between 1 and 100 is decidedly counterintuitive. If introspection were any guide to what one's brain was doing, one would think that the facts about which numbers were greater than which were stored in a table of some kind and simply looked up. In that case, why would it take longer to look up the ordering of 2 and 3 (or 65 and 62) than 2 and 5 (or 65 and 47)? It does, however, and this suggests that the comparison that underlies these judgments operates with noisy mental magnitudes. On this hypothesis, the brain maps the numerals to the noisy mental magnitudes that would be generated by the non-verbal numerical estimation system if it enumerated the corresponding numerosity. It then compares those two noisy mental magnitudes to decide which numeral represents the bigger numerosity.

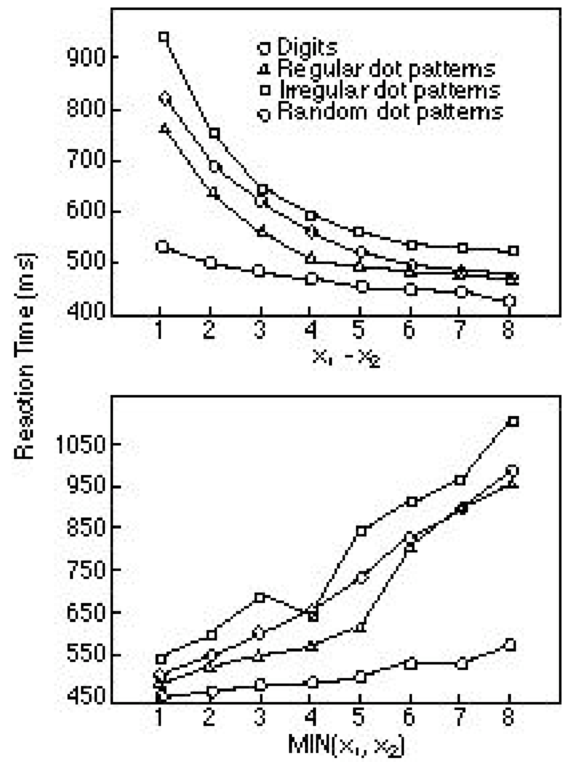


Figure 5. *The symbolic and non-symbolic size and distance effects on the human reaction time while judging numerical order in the range from 1 to 9. In three of the conditions, the numerosities to be judged were instantiated by two dot arrays (non-symbolic numerical ordering). The dots within each array were in either a regular configuration, an irregular configurations that did not vary upon repeated presentation, or in randomly varying configurations. In the fourth condition, the numerosities were represented symbolically by Arabic numerals. The top panel plots mean reaction times as a function of the numerical difference. The bottom plots it as a function of the size of the smaller comparand. (Replotted from Figures 1 and 2 in Buckley & Gillman, 1974.)*

On this hypothesis, the comparison that mediates the verbal judgment of the numerical ordering of two Arabic numerals uses the same mental magnitudes and the same comparison mechanism as is used by the non-verbal numerical reasoning system that we are assumed to share with many non-verbal animals. Consistent with this hypothesis is Brannon and Terrace's (2002) finding that reaction time functions from humans and monkeys for judgments of the numerical ordering of pairs of visually presented dot arrays are almost exactly the same (Figure 6).

Buckley and Gillman modeled the underlying comparison process (1974). In their model, numbers are represented in the brain by noisy signals (mental magnitudes), with overlapping distributions. The closer two numerosities are in the ordering of numerosities, the more their corresponding signal distributions in the brain overlap. When the subject judges the ordering of two numerosities, the brain subtracts the signal representing the one numerosity from the signal representing the other, and puts the signed difference in an accumulator, a mechanism that adds up inputs over, in this case, time. The accumulator for the ordering operation has fixed positive and negative thresholds. When its positive threshold is exceeded, it reports the one number to be greater than the other, and vice versa when its negative threshold is exceeded. If neither accumulator threshold is exceeded, the comparator resamples the two signals, computes a second difference, based on the two new samples, and adds it to the accumulator. The resampling explains why it takes longer (on average) to make the comparison when the numerosities being compared are closer. The closer they are, the more their corresponding signal distributions overlap. The more these distributions overlap, the more samples will have to be made and added together (accumulated) before (on average) a decision threshold is reached.

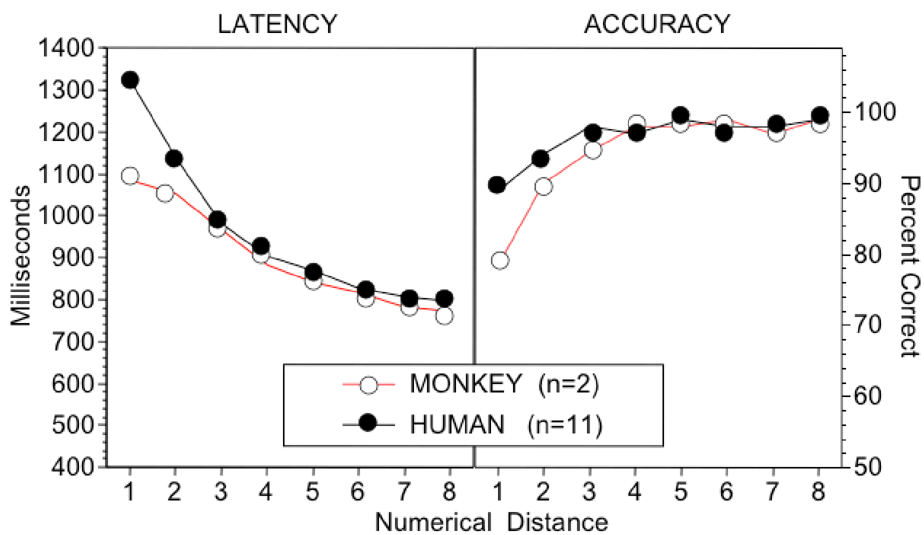


Figure 6. *The reaction time and accuracy functions for monkey (Rhesus macaque) and human subjects in touching the more numerous of two random dot visual arrays presented side by side on a touch-screen video monitor. Reproduced from (Elizabeth M. Brannon & Terrace, 2002 by permission of the authors and publisher.)*

Non-Verbal Counting in Humans

Given the evidence from the symbolic size and distance effects that humans represent number with mental magnitudes, it seems likely that they share with the non-verbal animals in the vertebrate clade a non-verbal counting mechanism that maps from numerosities to the mental magnitudes that represent them. If so, then it should be possible to demonstrate non-verbal counting in humans when verbal counting is suppressed. Whalen, Gallistel and Gelman (1999) presented subjects with Arabic numerals on a computer screen and asked them to press a key as fast as they could without counting until it felt like they had pressed the number signified by the numeral. The results from humans looked very much like the results from pigeons and rats: the mean number of presses increased in proportion to the target number and the standard deviations of the distributions of presses increased in proportion to their mean, so that the coefficient of variation was constant.

This result suggests, firstly, that subjects could count non-verbally, and, secondly, that they could compare the mental magnitude thus generated to a magnitude obtained by way of a learned mapping from numerals to mental magnitudes. Finally, it implies that the mapping from numerals to mental magnitudes is such that the mental magnitude given by this mapping approximates the mental magnitude generated by counting the numerosity signified by a given numeral.

In a second task, subjects observed a dot flashing very rapidly but at irregular intervals. The rate of flashing (8 per second) was twice as fast as estimates of the maximum speed of verbal counting (Mandler & Shebo, 1982). Subjects were asked not to count but to say about how many times they thought the dot had flashed. As in the first experiment, the mean number estimated increased in proportion to the number of flashes and the standard deviation of the estimates increased in proportion to the mean estimate. This implies that the mapping between the mental magnitudes generated by non-verbal counting and the verbal symbols for numerosities is bi-directional; it can go from a symbol to a mental magnitude that is comparable to the one that would be generated by non-verbal counting, and it can go from the mental magnitude generated by a non-verbal count to a roughly corresponding verbal symbol. In both cases, the variability in the mapping is scalar.

Whalen, et al (1999) gave several reasons for believing that their subjects did not count subvocally. We will not review them here, because a subsequent experiment speaks more directly to this issue (Sara Cordes, Gelman, Gallistel, & Whalen, 2001).

Cordes, et al (2001) suppressed articulation by having their subjects repeat a common phrase ("Mary had a little lamb") while they attempted to press a target number of times, or by having subjects say "the" coincident with each press.

In control experiments, subjects were asked to count their presses out loud In all conditions, subjects were asked to press as fast as possible.

The variability data from the condition where subjects were required to say "the" coincident with each press are shown in Figure 7 (filled squares). As in Whalen et al. (1999), the coefficient of variation was constant (scalar variability). The best fitting line has a slope that does not differ significantly from zero. The contrasting results from the control conditions, where subjects counted out loud are the open squares. Here, the slope--on this log-log plot---does deviate very significantly from zero. In verbal counting, one would expect counting errors--double counts and skips--to be the most common source of variability. On the assumption that the probability of a counting error is approximately the same at successive steps in a count, the resulting variability in final counts should be binomial rather than scalar. It should increase in proportion to the square root of the target value, rather than in proportion to the target value. If the variability is binomial rather than scalar, then when the coefficient of variation is plotted against the target number on a log-log plot, it should form a straight line with a slope of -0.5. This is what was in fact observed in the out-loud counting conditions: the variability was much less than in the non-verbal counting conditions and, more importantly, it was binomial rather than scalar. The mean slope of the subject-by-subject regression lines in the two control conditions was significantly less than zero and not significantly different from -0.5. The contrasting patterns of variability in the counting-out-loud and non-verbal counting conditions strengthen the evidence against the hypothesis that subjects in the non-verbal counting conditions were subvocally counting.

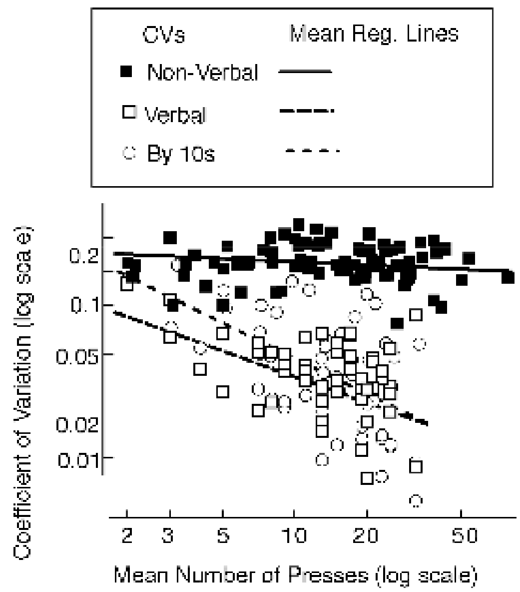


Figure 7. *The coefficients of variation (σ/μ) are plotted against the numbers of presses for the conditions in which subjects counted non-verbally and for the condition in which they fully pronounced each count word (double logarithmic coordinates). In the former condition, there is scalar variability, that is, a constant coefficient of variation. The slope of the regression line relating the log of the coefficient of variation to the log of mean number of presses does not differ from zero. In the latter, the variability is much less and it is binomial; the coefficient of variation decreases in proportion to the square root of the target number. In the latter case, the slope of the regression line relating the log of the coefficient of variation to the log of the mean number of presses differs significantly from zero but does not differ significantly from -0.5, which is the slope predicted by the binomial variability hypothesis. (Reproduced from Sara Cordes et al., 2001, by permission of the authors and the publisher.)*

In sum, non-verbal counting may be demonstrated in humans, and it looks just like non-verbal counting in non-humans. Moreover, mental magnitudes (real numbers) comparable to those generated by non-verbal counting appear to mediate judgments of the numerical ordering of symbolically presented integers. This suggests that the non-verbal counting system is what underlies and gives meaning to the linguistic representation of numerosity.

Non-verbal Arithmetic Reasoning in Humans

In humans as in other animals, nonverbal counting would be pointless if they did not reason arithmetically with the resulting mental magnitudes. Recent experiments give evidence that they can.

Barth (, 2001 #6452, see also H. Barth et al., under review (2003)) tested adults' performance on tasks that required the addition, subtraction, multiplication and division of non-verbally estimated numerosities, under conditions where verbally mediated arithmetic was unlikely. In her experiments, subjects were given instances of two numerosities in rapid sequence, each instance presented too quickly to be verbally countable. Then, they were given an instance of a third numerosity, and they indicated by pressing one of two buttons whether the sum, or difference, or product, or quotient of the first two numerosities was greater or less than the third.

The numerosities were presented either as dot arrays (with dot density and area covered controlled) or as tone sequences. In some conditions, presentation modalities were mixed, so, for example, subjects compared the sum of a tone sequence and a dot array to either another tone sequence or another dot array.

In Barth's results, there was no effect of comparand magnitude on reaction time or accuracy, only an effect of their ratio. That is, it did not matter how big the two numerosities were; only the proportion of the smaller to the larger affected reaction time and accuracy. The same proved to be true in her experiments involving mental magnitudes derived by arithmetic composition. This enables a comparison between the case in which the comparands are both given directly and the case in which one comparand is the estimated sum or difference of two estimated numerosities. As Figure 8 shows, the accuracy of comparisons involving a sum was only slightly less at each ratio of the comparands than the accuracy of a comparison between directly given comparands.

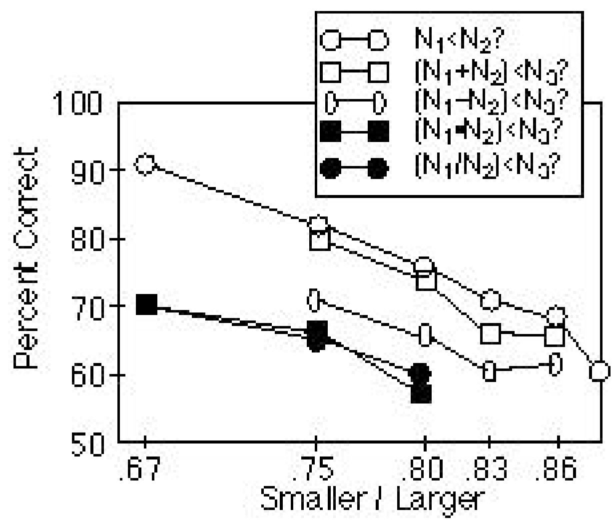


Figure 8. *The accuracy of order judgments for two non-verbally estimated numerosities. The estimates of numerosity were based on direct instantiations in the first condition ($N1<N2$). In the other conditions, one of them was derived from the composition of two other estimates. (Data replotted from H. C. Barth, 2001, p. 109).*

At a given comparand ratio, the accuracy of comparisons involving differences was less than the accuracy of a comparison between directly given comparands (Figure 8). This could hardly be otherwise. For addition, the sum increases as the magnitude of the pair of operands increases, but for subtraction, it does not; the difference between a billion and a billion and one is only one. The uncertainty (estimation noise) in the operands must propagate to the result of the operation, so the uncertainty about the true value of a difference must depend in no small measure on the magnitude of the operands from which it derived. If one looks only at the ratio of the difference to the other comparand, one fails to take account of the presumably inescapable impact of operand magnitude on the noise in the difference.

Bath's experiments establish by direct test the human ability to combine noisy non-verbal estimates of numerosity in accord with the combinatorial operations that define the system of arithmetic. In her data (Figure 8), as the proportion between the smaller and larger comparand increases toward unity, the accuracy of the comparisons degrades in a roughly parallel fashion regardless of the derivation of the first comparand. This suggests that the scalar variability in the nonverbal estimates of numerosity propagates to the mental magnitudes produced by the composition of those estimates.

Barth's data do not, however, directly demonstrate the variablity in the results of composition nor allow one to estimate the quantitative relation between the noise in the operands and the noise in the resultant. Cordes, Gallistel, Gelman and Latham (in preparation (2003)) used the above-described key-tapping paradigm to demonstrate the non-verbal addition and subtraction of non-verbal numerical estimates and the quantitative relation between the variability in the estimates of the sums and differences and the variability in the estimates of the operands.

In the baseline condition of the Cordes et al. (in preparation (2003)) experiment, subjects saw a sequence of rapid arhythmic variable duration dot flashes on a computer screen, at the conclusion of which they attempted to make an equivalent number of taps on one button of a 2-button response box, tapping as rapidly as they could while saying *the* out loud coincident with each tap. In the compositional conditions, subjects saw one sequence on the left side of the screen, a second sequence on the right side, and were asked to tap out either the sum or the difference. In the subtraction condition, they pressed the button on the side they believed to have had the fewer flashes as many times as they felt was required to make up the difference.

Sample results are shown in Figure 9. The numbers of responses subjects made were in all cases approximately linear functions of the numbers they were estimating, demonstrating the subjects' ability to add and subtract the mental magnitudes representing numerosities. In the baseline condition, the variability in the numbers tapped out was an approximately scalar function of the target number, although there was some additive and binomial variability.

The variability in the addition data was also, to a first approximation, a scalar function of the objective sum. Not surprisingly, however, the variability in the subtraction data was not. In addition, answer magnitude covaries with operand magnitude: the greater the magnitude of the operands, the greater the magnitude of their sum[4]. In subtraction, answer magnitude is poorly correlated with operand magnitude, because large magnitude operands often produce small differences. Insofar as the scalar variability in the estimates of operand magnitudes propagates to the variability in the results of the operations, there will be large variability in these small differences.

[4] The magnitude of a pair of numbers is the square root of the sum of their squares.

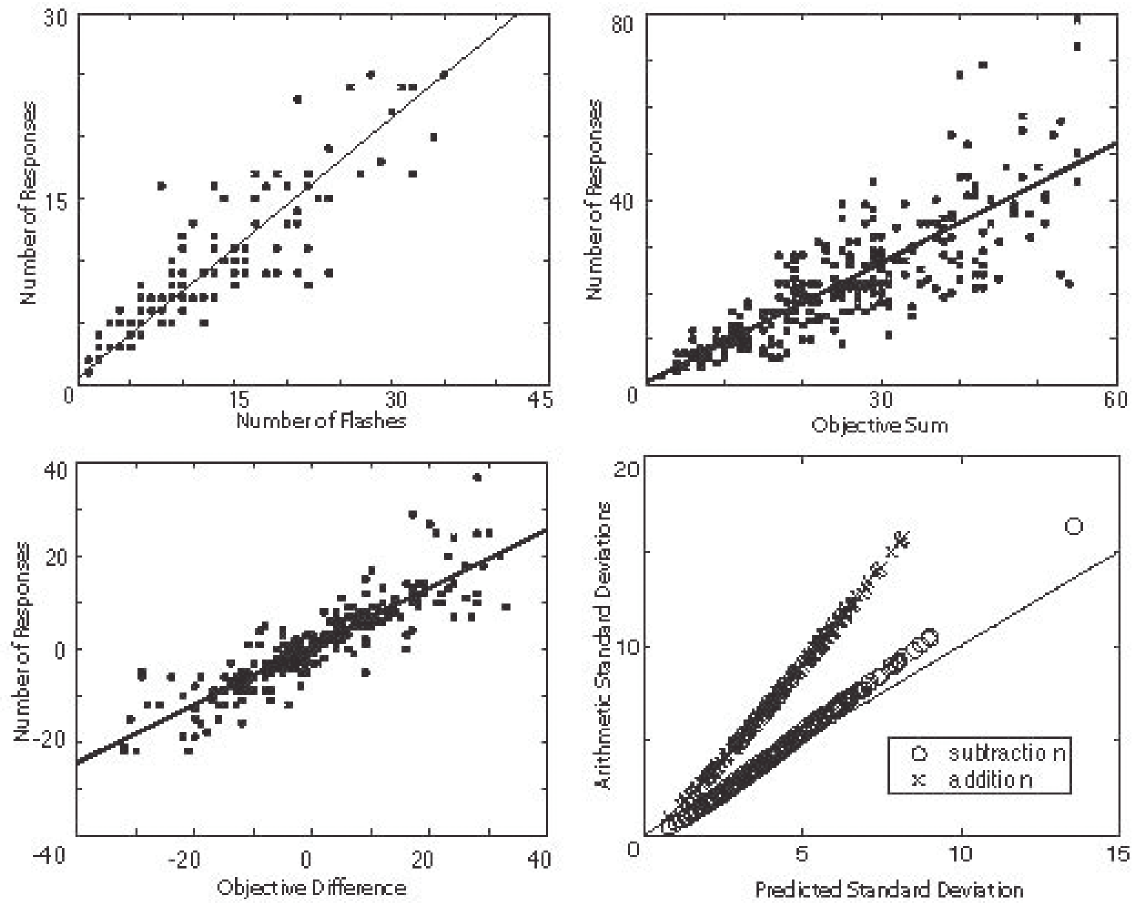


Figure 9.A *Number of responses (key taps) as a function of the number of flashes for one subject.* B. *Number of responses as a function of the sum of the numbers of flashes in two flash sequences*. C. *Number and sign (side) of the responses as a function of the difference between the numbers of flashes in two sequences of flashes*. D. *Predicting the variability in the sums and differences from the variability in the operands. Adapted from Cordes et al. (in preparation 2003) by permission of the authors and publisher.*

Cordes et al. (in preparation 2003) fit regression models with additive, binomial and scalar variance parameters to the baseline data, and to the addition and subtraction data. These fits enabled them to assess the extent to which the magnitude of the pair of operands predicted the variability in their sum and difference. On the assumption that there is no covariance in the operands, the variance in the results of both subtraction and addition should be equal to the sum of the variances for the two operands. When Cordes et al. plotted predicted variabilty against directly estimated variability (Figure 9D), they found that the subtraction data did conform approximately to expectations, but that the addition data clearly fell above the line. In other words, the variability in results of subtraction were approximately what was expected from the sum of the estimated variances in the operands, but the variability in the addition results was greater than expected.

Retrieving Number Facts

There is an extensive literature on reaction times and error rates in adults doing single-digit arithmetic (Ashcraft, 1992; J. I. Campbell, 1999; J. I. Campbell & J. Fugelsang, 2001; J. I. Campbell & Gunter, 2002; J. I. D Campbell, 2003, in preparation); J. I. D. Campbell & J. Fugelsang, 2001; Noel, 2001). It resists easy summary. However, magnitude effects analogous to those found for order judgments are a salient and robust finding: The bigger the numerosities represented by a pair of digits, the longer it takes to recall their sum or product and the greater the likelihood of an erroneous recall. The same is true in children (J.I.D. Campbell & Graham, 1985). For both sets of number facts, there is a notable exception to this generalization. The sums and products of ties (for example, 4 + 4 or 9 x 9) are recalled much faster than is predicted by the regressions for non-ties, although ties, too, show a magnitude effect (Miller, Perlmutter, & Keating, 1984).

There is a striking similarity in the effect of operand magnitude on the reactions times for both addition and multiplication. The slopes of the regression lines (reaction time versus the sum or product of the numbers involved) are not statistically different (Geary, Widman, & Little, 1986). More importantly, Miller et al. (1984) found that the best predictor of reaction times for digit multiplication problems was the reaction times for digit addition problems, and vice versa. In other words, the reaction-time data for these two different sets of facts, which are mastered at different ages, show very similar microstructure.

These findings suggest a critical role for mental magnitudes in the retrieval of the basic number facts (the addition and multiplication tables), upon which verbally mediated computation strategies depend. Whalen's (1997) diamond arithmetic experiment showed that these effects depend primarily on the magnitude of the operands not on the magnitude of the answers, nor on the frequency with which different facts are retrieved (although these may also contribute). Whalen (1997) taught subjects a new arithmetic operation of his own devising, the diamond operation. It was such that there was no correlation between operand magnitude and answer magnitude. Subjects received equal practice on each fact, so explanations in terms of differential practice did not apply. When subjects had achieved a high level of proficiency at retrieving the diamond facts, Whalen measured their reaction times. He obtained the same pattern of results seen in the retrieval of the facts of addition and multiplication.

Two Issues

What is the form of the mapping from magnitudes to mental magnitudes?

Weber's law, that the discriminability of two magnitudes (two sound intensities or two light intensities) is a function of their ratio, is the oldest and best established quantitative law in experimental psychology. Its implications for the question of the quantitative relation between directly measurable magnitudes (hereafter called objective magnitudes) and the mental magnitudes by which they are represented (hereafter called subjective magnitudes) have been the subject of analysis and debate for more than a century. This line of investigation led to work on the mathematical foundations of measurement, work concerning the question what it means to measure something (D. Krantz, Luce, Suppes, & Tversky, 1971; D. H. Krantz, 1972; Luce, 1990; Stevens, 1951, 1970). The key insight from work on the foundations of measurement is that the quantitative form of the mapping from things to their numerical representives cannot be separated from the question of the arithmetic operations that are validly performed on the results of that mapping. The question of the form of the mapping is only meaningful at the point where the numbers (magnitudes) produced by the mapping enter into arithmetic operations.

The discussion began when Fechner used Weber's results to argue that subjective magnitudes (for example, loudness and brightness) are logarithmically related to the corresponding objective magnitudes (sound and light intensity). Fechner's reasoning is echoed down to the present day by authors who assume that Weber's law implies logarithmic compression in the mapping from objective numerosity to subjective numerosity. These conjectures are uninformed by the literature on the measurement of subjective quantities spawned by Fechner's assumption. In deriving logarithmic compression from Weber's law, Fechner assumed that equally discriminable differences in objective magnitude correspond to equal differences in subjective magnitude. However, when you directly ask subjects whether they think just discriminable differences in, for example, loudness represent equal differences, they do not; they think a just discriminable difference between two loud sounds is greater than the just discriminable difference between two soft sounds (Stevens, 1951).

The reader will recognize that Barth performed both experiments—the discrimination experiment (Weber's experiment) and the difference judging experiment—but with numerosities instead of noises. In the discrimination experiment, she found that Weber's law applied: Two pairs of non-verbally estimated numerosities can be correctly ordered 75% of the time when $N_1/N_2 = N_3/N_4 = .83$, where N now refers to the (objective) numerosity of a set (Figure 8). From Moyer and Landauer (1967) to the present (S Dehaene, 2002), this has been taken to imply that subjective numerosity is a logarithmic function of objective numerosity. If that were so, and if subjects estimated the arithmetic differences between objective magnitudes from the arithmetic differences in the corresponding subjective magnitudes, then the Barth (2001) and the Cordes et al. (in preparation 2003) subtraction experiments would have failed, and so would the experiments demonstrating subtraction of time and number in nonverbal animals, because the arithmetic difference between the logarithms of two magnitudes represents their quotient, not their arithmetic difference.

In short, when subjects respond appropriately to the arithmetic difference between two numerical magnitudes, their behavior is not based on the arithmetic difference between mental (subjective) magnitudes that are proportional to the logarithms of the objective magnitudes. That much is clear. Either: (Model 1) The behavior is based on the arithmetic difference in mental magnitudes that are proportional to the objective magnitudes (a proportional rather than logarithmic mapping). Or: (Model 2) Dehaene (2001) has suggested that mental magnitudes are proportional to the logarithms of objective magnitudes and that, to obtain from them the mental magnitude corresponding to the objective difference, the brain uses a look-up table, a procedure analogous to the procedure that Whalen's (1997) subjects used to retrieve the facts of diamond arithmetic. In this model, the arithmetic difference between two mental magnitudes is irrelevant; the two magnitudes serve only to specify where to enter the look-up table, where in memory the answer is to be found.

In summary, there are two intimately interrelated unknowns concerning the mapping from objective to subjective magnitudes—the form of the mapping and the formal character of the operations on the results of the mapping. Given the experimental evidence showing valid arithmetic processing, knowing either would fix the other. In the absence of firm knowledge about either, can behavioral experimental evidence decide between the alternative models? Perhaps not definitively, but there are relevant considerations. The Cordes et al. (in preparation 2003) experiment estimates the noise in the results of the mental subtraction operation at and around 0 difference (Figure 9C). There is nothing unusual about the noise around answers of approximately 0. It is unclear what assumptions about noise would enable a logarithmic mapping model to explain this. The logarithm of a quantity goes to minus infinity as the quantity approaches 0, and there are no logarithms for negative quantities. On the assumption that realizable mental magnitudes like realizable non-mental magnitudes cannot be infinite, the model has to treat 0 as a special case. How the treatment of that special case could exhibit noise characteristics of a piece with the noise well away from 0 is unclear.

It is also unclear how the logarithmic-mapping-plus-table-lookup model can deal with the fact that the sign of a difference is not predictable a priori. In this model a bigger magnitude (number) cannot be subtracted from a smaller, because the resulting negative number does not have a logarithm, that is, there is no way to represent a negative magnitude in a scheme where magnitudes are represented by their logarithms. Thus, this model is not closed under subtraction.

Is there a distinct representation for small numbers?

When instantiated as arrays of randomly arranged small dots, presented for a fraction of a second, small numerosities can be estimated more quickly than large ones, but only up to about 6. Thereafter, the estimates increase more or less linearly with the number of dots, but the reaction time is flat (Figure 10).

Subjects confidence in their estimates also falls off precipitously after 6 (Kaufman, Lord, Reese, & Volkman, 1949; Taves, 1941). This led Taves to argue that the processes by which subjects arrive at estimates for numerosities of 5 or fewer are distinct from the processes by which they arrive at estimates for numerosities of 7 or more. Kaufman et al. (1949) coined the term subitizing to describe the process that operates in the range below 6.

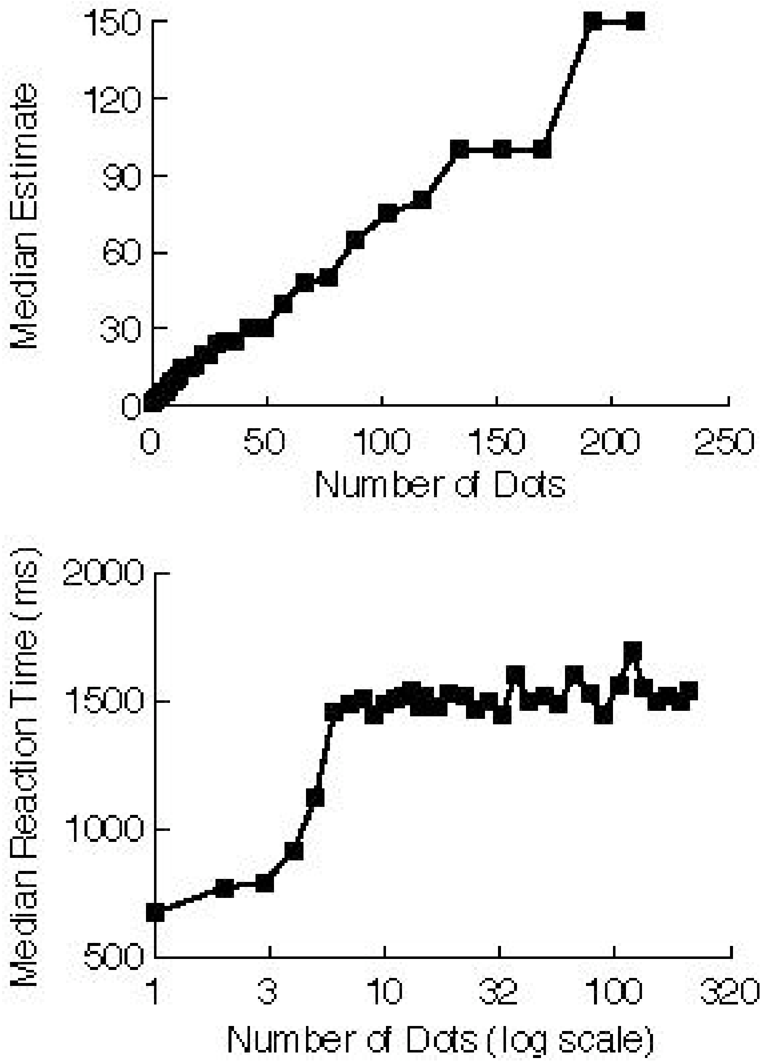


Figure 10. *Estimates of dot numerosity (top) and time to make an estimate (bottom) as functions of the number of dots in tachistoscopically presented arrays of randomly positioned dots. (Plotted from the data for the speeded instruction group in Table 1 of Kaufman et al., 1949, p. 510.)*

When the dot array to be enumerated is displayed until the subject responds, rather than very briefly by a tachistoscope, the reaction time function is superimposable on the one shown in Figure 10, up to and including numerosity 6. It does not level off at 6, however; rather, it continues with the same slope (about 325 ms/dot) indefinitely (Jensen, Reese, & Reese, 1950). This slope represents the time it takes to count subvocally. Thus, the discontinuity at 6 represents the point at which a non-verbal numerosity-estimating mechanism or process takes over from the process of verbal counting, because, presumably, it is not possible to count verbally more than 6 items under tachistoscopic conditions.

The non-verbal numerosity-estimating process is probably this process that is the basis for the demonstrated capacity of humans to compare (order) large numerosities instantiated either visually or auditorily (Hilary Barth, Kanwisher, & Spelke, 2003). The reaction times and accuracies for these comparisons show the Weber law characteristic, which is a signature of the process that represents numerosities by mental magnitudes rather than by discrete word-like symbols (Cordes et al 2001). The assumption that the representation is by mental magnitudes regardless of the mode of presentation is consistent with the finding that there is no cost to cross-modal comparisons of large numerosities; these comparisons take no longer and are no more inaccurate than comparisons within presentation modes (Hilary Barth et al., 2003).

There is controversy about the implications of the reaction time function within the subitizing range (the range below 6). In this range, there is approximately a 30 ms increment in going from one to two dots, an 80 msec increment in going from two to three, and a 200 ms increment in going from three to four. These are large increments. The net increment from one to four is about 300 ms, which is half the total latency to respond to a one-item array (Jensen et al., 1950; Kaufman et al., 1949; Mandler & Shebo, 1982). Moreover, the increments get bigger at each step. In particular, the step from 2 to 3 is significantly greater than the step from 1 to 2 in almost every data set.

It is often claimed that there is a discontinuity in the reaction time function within the subitizing range (Davis & Pérusse, 1988; Klahr & Wallace, 1973; Simon, 1999; Strauss & Curtis, 1984; Woodworth & Schlosberg, 1954) (Piazza, Giacomini, Le Bihan, & Dehaene, 2003), but it has also often been pointed out that there is no empirical support for this claim (Balakrishnan and Ashby, (1992). Because the reaction time function is neither flat nor linear in the range from 1 to 3, it offers no support for the common theory that very small numbers are directly perceived, as was first pointed out by the authors who coined the term subitizing (Kaufman, et al 1949).

Gallistel and Gelman (1992) and Dehaene and Cohen (1994) suggested that in the subitizing range there is a transition from a strategy based on mapping from nonverbally estimated mental magnitudes to a strategy based on verbal counting. This hypothesis has recently received important support from a paper by Whalen, West and Cook (J. Whalen, West, & Cook, under review, 2003). By strongly encouraging rapid approximate estimates and taking measures to make verbal counting more difficult, Whalen et al. (under review 2003) obtained a reaction time function with a slope of 47 ms per item, from 1 to 16 items.

The coefficient of variation in the estimated numbers was constant from 1 to 16 at about 14.5%, which is close to the value of 16% in the animal timing literature (Gallistel, King, & McDonald, in press). Thus, the Whalen et al. data show scalar variability in rapid number estimates all the way down to estimates of one and two, as do the data of Cordes et al. (2001). Whalen et al. (under review 2003) show that with this level of noise in the mental magnitudes being mapped to number words, the expected percent errors in the resulting verbal estimates of numerosity are close to zero in the range 1-3 and increase rapidly thereafter--in close accord with the experimentally observed percent errors in their speeded condition (Figure 11). This explains why subjects in experiments where it is not strongly discouraged switch to subvocal verbal counting somwhere between 4 and 6, and why their confidence in their speeded estimates falls off rapidly after 6 (Kaufman et al., 1949; Taves, 1941). Whalen et al. (under review) attribute the constant slope of 47 ms/item in the speeded reaction time function to a serial non-verbal counting process. In short, the reaction time function does not support the hypothesis that there are percepts of twoness and threeness, constituting a representation of small numerosities incommensurable with the mental magnitudes that represent other numerosities.

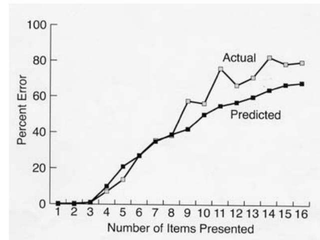


Figure 11. *The observed percent errors as a function of number of dots in Whalen's speeded condition compared to the percent expected on the hypothesis that the estimates were obtained by way of a mapping from non-verbal mental magnitudes to the corresponding number words and that the mental magnitudes had scalar variability with a coefficient of variation of 0.145. Reproduced from Whalen et al. (under review) by permission of the authors and the publisher.*

The Development of Verbal Numerical Competence

It appears that the system of non-verbal mental magnitudes plays fundamental role in verbal numerical behavior: When verbal counting is too slow to satisfy time constraints, it mediates the finding of a number word that specifies approximately the numerosity of a set. It mediates the ordering of the symbolic numbers and the numerosities they represent. And, it mediates the retrieval of the verbal number facts (the addition and multiplication tables) upon which verbal computational procedures rest. All of these roles require a mapping between the mental magnitudes that represent numerosity and number words and written numerals. Thus, in the course of ordinary development, humans learn a bidirectional mapping between the mental magnitudes that represent numerosity and the words and numerals that represent numerosity (Gallistel & Gelman, 1992; Gelman & Cordes, 2001). They make use of this bidirectional mapping in talking about number and the effects of combinatorial operations with numbers. There is broad agreement on this conclusion within the literature on numerical cognition, because of the abundant evidence for Weber-law characteristics in symbolic numerical behavior. The literature on the deficits in numerical reasoning seen in brain injured patients is broadly consistent with this same conclusion (S. Dehaene, 1997; Noel, 2001).

It also seems plausible that the nonverbal system of numerical reasoning mediates verbally expressed numerical reasoning. It seems plausible, for example, that adults believe that $(2 + 1) > 2$ and *four minus two is less than four*, because that is the way the mental magnitudes behave to which they (unconsciously) refer those symbols in order to endow them with meaning and with reference to the world.

Empiricists will offer as an alternative the hypothesis that adults believe these symbolic propositions because they have repeatedly observed that the properties of the world to which the words or symbols somehow refer behave in this way. Adults know, for example, that the word *two* refers to every set that can be placed in one-one correspondence with some uhr-two set, and likewise, *mutatis mutandis*, for the word *one,* and that the word *plus* refers to the uniting of sets, and that the word *greater than* refers to the relation between a set and its proper subsets, and so on. From an empiricist perspective, the words have these real world references only by virtue of the experiences adults have had, which are ubiquitous and universal.

Nativist/rationalists will respond that reference to the world by verbal expressions is mediated by preverbal world-referring symbolic systems in the mind of the hearer and that the ubiquity and universality of the experiences that are supposed to have created world-reference for these expressions are grounds for supposing that symbolic systems with these properties are part of the innate furniture of the mind. We will not pursue this old debate further, except to note the possible relevance of the experiments reviewed above demonstrating that non-verbal animals reason arithmetically about both numerosities (integer quantities) and magnitudes (continuous quantities).

We turn instead to the experimental literature on numerical competence in very young children. It is difficult to demonstrate conclusively behavior based on numerosity in infants, because it is hard not to confound variation in one or more continuous quantites with variation in numerosity, and infants often respond on the basis of continuous dimensions of the stimulus (Clearfield & Mix, 1999; Lisa Feigenson, Carey, & Spelke, 2002; see Mix, Huttenlocher, & Levine, 2002, for review). Nonetheless, there are studies that appear to demonstrate sensitivity to numerical order in infants (Elizabeth M. Brannon, 2002). Moreover, the ability of infants to discriminate sets on the basis of numerosity extends to pairs as large as 8 vs 16 (Lipton & Spelke, 2003; Xu & Spelke, 2000). Thus, there is reason to suppose that preverbal children share with non-verbal animals a non-verbal representation of numerosity.

The assumption that preverbal children represent numerosities by a system of mental magnitudes homologous to the system found in non-verbal animals is the foundation of the account of the development of verbal numerical competence suggested by Gelman and her collaborators (Gelman & Brenneman, 1994; Gelman & Cordes, 2001; Gelman & Williams, 1998). They argue that the development of verbal numerical competence begins with learning to count, which is guided from the outset by the child's recognition that verbal counting is homomorphic to non-verbal counting. In non-verbal counting, the pouring of successive cups into the accumulator (the addition of successive unit magnitudes to a running sum) creates a one-to-one correspondence between the items in the enumerated set and a sequence of mental magnitudes. Although the mental magnitudes thus created have the formal properties of real numbers, the process that creates generates a discretely ordered sequence of mental magnitudes, an ordering in which each magnitude has a next magnitude. The final magnitude represents the numerosity of the set. Verbal counting does the same thing; it assigns successive words from an ordered list to successive items in the set being enumerated, with the final word representing the cardinality of the set.

Gelman and her collaborators argue that the principles that govern non-verbal counting inform the child's counting behavior from its inception (Gelman & Gallistel, 1978). Children recognize that number words reference numerosities because they implicitly recognize that they are generated by a process homomorphic to the non-verbal counting of serially considered sets. Number words have meaning for the child, as for the adult, because it recognizes at an early age that they map to the mental magnitudes by which the non-verbal mind represents numerosities. On this account, the child's mind tries to apply from the outset the Gelman and Gallistel counting principles (Gelman & Gallistel, 1978): that counting must involve a one-one assignment of words to items in the set, that the words must be taken from a stably ordered list, and that the last word represents the cardinality of the set. However, it takes a long time to learn the list, and a long time to implement the verbal counting procedure flawlessly, because list learning is hard, because the implementation of the procedure is challenging (Gelman & Greeno, 1989), and because the child is often confused about what the experimenter wants.

Critical to Gelman's account is evidence that during the period when they are learning to count children already understand that the last count word represents a property of the set about which it is appropriate to reason arithmetically. Without such evidence, there is no grounds for believing that the child has a truly numerical representation. Evidence on this crucial point comes from the so-called magic experiments (Bullock & Gelman, 1977; Gelman, 1972, 1977, 1993). These experiments drew children into a game in which a winner and loser plate could be distinguished on the basis of the number of toy mice they contained. The task engaged children's attention and caused them to justify their judgments as to whether an uncovered plate was or was not the winner. Children as young as 2 1/2 indicated that the numerosity was the decisive dimension, and they spontaneously counted to justify their judgment that the plate with the correct numerosity was the winner. On magic trials, a mouse was surreptitiously added or subtracted from the winner plate during the shuffling, so that it had the same numerosity as the loser plate. Now, both plates when uncovered were revealed to be loser plates. In talking about what surprised them, children indicated that something must have been added or subtracted, and they counted to justify themselves. This is strong evidence that children as young as two and one half years of age understand that counting gives a representation of numerosity about which it is appropriate to reason arithmetically. This is well before they become good counters (Fuson, 1988; Gelman & Gallistel, 1978; Hartnett & Gelman, 1998). Surprised 2 1/2 year olds made frequent use of number words. They used them in idiosyncratic ways, but ways that nonetheless conformed to the counting principles (Gelman, 1993), including the cardinality principle.

An second account of the development of counting and numerical understanding grows, firstly, out of the conviction of many researchers that while two year olds count, albeit badly, they do not understand what they are attempting to do (Carey, 2001a, 2001b; Fuson, 1988; Mix et al., 2002; Karen Wynn, 1990; K. Wynn, 1992b). It rests, secondly, on evidence suggesting that in the spontaneous processing of numerosities by infants and by monkeys, there is a discontinuity between numbers of 4 or less and bigger numbers. In some experiments, the infant and monkey subjects discriminate all numerosity pairs in the range 1 to 4, but fail to discriminate pairs that include a numerosity outside that range (e.g., <3,6>), even when, as in the example, their ratio is greater than the ratio between discriminable pairs of 4 or less (L. Feigenson, Carey, & Hauser, 2002; Uller, Carey, Huntley-Fenner, & Klatt, 1999; Uller, Hauser, & Carey, 2001).

How to reconcile these latter findings with the finding that infants do discriminate the pair <8,16> (Lipton & Spelke, 2003; Xu & Spelke, 2000) is unclear. Similarly, it is unclear how to reconcile the monkey findings with the literature showing the discrimination of numerosities small and large in non-verbal animals. Particularly to be born in mind in this connection is the finding that monkeys cannot be taught to order numerosities in other than a numerical order (Elizabeth M. Brannon & Terrace, 2000), even though they can be taught to order things other than numerosities in a arbitrary, experimenter-imposed order (Terrace, Son, & Brannon, 2003). This implies that numerical order is spontaneously salient to a monkey.

The account offered by Carey (Carey, 2001a, 2001b, in press) begins with the assumption that convincing cases of infant number discrimination involving numbers less than 4 may depend on the object tracking system. For example, in Wynn's (1992a) experiment, the infants saw an object appear to join or leave one or two objects behind an occluding screen. They were surprised when the screen was removed to reveal a number of objects different from the number that ought to have been there. This surprise may have arisen only the infant's belief in object permanence.

When an infant sees an object move behind an occluding screen, whose subsequent removal fails to reveal an object, the infant is surprised (R. Baillargeon, 1995; R. Baillargeon, E. S., Spelke, & Wasserman, 1985). His/her surprise is presumably mediated by a system for tracking objects, such as the object file system suggested by (Kahneman, Treisman, & Gibbs, 1992) or the FINST system suggested by (Pylyshyn & Storm, 1988). This system maintains a marker (object file or FINST) for each object it is tracking, but it can only track about 4 objects (Scholl & Pylyshyn, 1999). On this account, infants in experiments like Wynn's are surprised for the same reason as in original object-permanence experiments: there is a missing object. The infant has an active mental marker or pointer that no longer points to an object. Or, there is an object for which it has no marker.

Carey argues that sets of object files are the foundations on which the understanding of the integers rests. The initial meaning of the words *one, two, three and four* does not come from the corresponding mental magnitudes; rather, it comes from sets of object files. The child comes to recognize the ordering of the referents of *one, two, three and four* because a set of two active object files has as a proper subset a set of one object file, and so on. The child comes to recognize that addition applies to the things referred to by these words because the union of two sets of object files yields another set of object files (provided the union does not create a set greater than 4). This is the foundation of the child's belief in the successor principle: every integer has a unique successor.

This account seems to ignore the basic function of a set of, for example, two object files (FINSTs, pointers), which is to point to two particular objects. If *two* referred to a particular set of two object files, it would presumably be useable only in connection with the two objects they pointed to. It would be a name for that pair of objects, not for all sets that share with that set the property of twoness.

A particular set of pointers cannot substitute for (is not equal to) another such set without loss of function, because its function is to point to one pair of objects, while the function of another such set is to point to a different pair. There is no reason to believe that there is any such thing as a general set of two pointers, a set that does not point to any particular set of two objects, but represents all the sets that do so point. Any set of two object files is an instance of a set with the twoness property (a token of twoness), but it can no more represent twoness than a name that picks out one particular dog, e.g., *Rover*, can represent the concept of a dog. A precondition of *Rover*'s serving the latter function is that it not serve the former. By contrast, any instance of the numeral *2* can be substituted for any other without loss of function, and so can a pair of hash marks.

A second problem with this account is that it is unclear how a system so lacking in closure be the basis for inferring a system whose function depends so strongly on closure. The Carey suggestion is motivated by findings that the maximum numerosity of a set of active object files is at most 4. There are only 9 numerically distinct unordered pairs of sets of 4 or less (<1,1>, <1,2>, <1,3>, <1,4>, <2,2>, <2,3>, <2,4>, <3,3>, and <3,4>). Five of the nine pairs, when composed (united) yield a set to numerous to be a set of object files. From this foundation, the mind of the child is said to infer that the numbers may be extended indefinitely by addition. One wants to know what the inference rule is that ignores the many negative instances in the base data set.

Conclusions and Future Directions

There is a widespread consensus, backed by a large and diverse experimental literature, that adult humans share with non-verbal animals a non-verbal system for representing discrete and continuous quantity that has the formal properties of continuous magnitudes. Mental magnitudes represent quantities in the same sense that, given a proper measurement scheme, real numbers represent line lengths. That is, the brains of non-verbal animals perform arithmetic operations with mental magnitudes; they add, subtract, multiply, divide and order them. The processes or mechanisms that map numerosities (discrete quantities) and magnitudes (continuous quantities) into mental magnitudes,and the operations that the brain performs on them, are together such that the results of the operations are approximately valid, albeit imprecise; the results of computations on mental magnitudes map appropriately back onto the world of discrete and continuous quantity.

Scalar variability is a signature of the mental magnitude system. Scalar variability and Weber's law are different sides of the same coin: models that generate scalar variability also yield Weber's law. There are two such models. One assumes that the mapping from objective quantity to subjective quantity (mental magnitude) is logarithmic; the other assumes that it is scalar. Both assume noise. That is, they assume that the signal corresponding to a given objective quantity varies from occasion to occasion, in a manner described by a Gaussian probability density function. The variation is on the order of 15% in both animal timing and human speeded number estimation.

The first model (logarithmic mapping) assumes that scalar behavioral variability reflects a constant level of noise in the signal distributions. This yields proportional (scalar) variability, because constant logarithmic intervals correspond to constant proportions in the corresponding non-logarithmic magnitudes. The second model (scalar mapping) assumes scalar variability in the underlying signal distributions. The overlap in the two signals distributions is a function only of the ratio between the represented numerosities in both models, which is why they both predict Weber's law.

Both models assume that there is only one mapping from objective quantities to subjective quantities (mental magnitudes), but there is no compelling reason to accept this assumption. The question of the quantitative form of the mapping only makes sense at the point at which the mental magnitudes enter into combinatorial operations. The form may differ for different combinatorial operations. In the future, the analysis of variability in the answers from nonverbal arithmetic may decide between the models. Thus, an important component of future models must be the specification of how variability propagates from the operands to the answers.

The system of mental magnitudes plays many important roles in verbalized adult number behavior. For example, it mediates judgments of numerical order and the retrieval of the verbal number facts (addition and multiplication tables) upon which verbalized and written calculation procedures depend. It also mediates the finding of number words to represent large numerosities, presented too briefly to be verbally counted, and, more controversially, the rapid retrieval of number words to represent numerosities in the subitizing range (1 - 6).

Any account of the development of verbal numerical competence must explain how subjects learn the bidirectional mapping between number words and mental magnitudes, without which mental magnitudes could not play the just described roles. One account of the development of verbal numerical competence assumes that it is directed from the outset by the mental magnitude system. The homomorphism between serial non-verbal counting and verbal counting is what causes the child to appreciate the enumerative function of the count words. The child attends to these words because of the homomorphism. Learning their meaning is the process of learning their mapping to the mental magnitudes. Another account assumes that the count words from *one* to *four* are initially understood to refer to sets of object files, mental pointers that pick out particular objects. On this account, the learning of the mapping to mental magnitudes comes later, after the child has extensive counting experience.

References

Ashcraft, M. H. (1992). Cognitive arithmetic: A review of data and theory. *Cognition, 44*(1-2), 75-106.

Baillargeon, R. (1995). Physical reasoning in infancy. In M. S. Gazzaniga (Ed.), *The cognitive neurosciences* (pp. 181-204). Cambridge, MA: MIT Press.

Baillargeon, R., E. S., Spelke, E., & Wasserman, E. (1985). Object permanence in 5-month-old infants. *Cognition, 20*, 191-208.

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

Barth, H., Kanwisher, N., & Spelke, E. (2003). The construction of large number representations in adults. *Cognition, 86*(3), 201-221.

Barth, H., La Mont, K., Lipton, J., Dehaene, S., Kanwisher, N., & Spelke, E. (under review (2003)). Nonsymbolic arithmetic in adults and young children.

Barth, H. C. (2001). *Numerical cognition in adults: Representation and manipulation of nonsymbolic quantities.* Unpublished PH.D., MIT, Cambridge, MA.

Beran, M. J. (2001). Summation and numerousness judgments of sequentially presented sets of items by chimpanzees (*Pan troglodytes*). *Journal of Comparative Psychology, 115*, 181-191.

Biro, D., & Matsuzawa, T. (2001). Chimpanzee numerical competence: Cardinal and ordinal skills. In *Matsuzawa, Tetsuro (Ed) Primate Research INstitute; Kyoto University (2001) Primate origins of human cognition and behavior* (pp. 199-225). New York, NY: Springer-Verlag Publishing.

Boyer, C. B., & Merzback: , U. C. (1989). *A history of mathematics. 2nd edition*. New York: Wiley.

Boysen, S. T., & Berntson, G. G. (1989). Numerical comptetence in a chimpanzee (*Pan troglodytes*). *J. Comp. Psychol., 103*, 23-31.

Boysen, S. T., & Hallberg, K. I. (2000). Primate numerical competence: Contributions toward understanding nonhuman cognition. *Cognitive Science, 24*(3), 423-443.

Brannon, E. M. (2002). The development of ordinal numerical knowledge in infancy. *Cognition, 83*(3), 223-240.

Brannon, E. M., & Roitman, J. D. (2003). Nonverbal representations of time and number in animals and human infants. In W. H. Meck (Ed.), *Functional and neural mechanisms of interval timing* (pp. 143-182). Boca Raton, FL: CRC Press.

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*(1), 31-49.

Brannon, E. M., & Terrace, H. S. (2002). The evolution and ontogeny of ordinal numerical ability. In *Bekoff, Marc (Ed); Allen, Colin (Ed); et al (2002) The cognitive animal: Empirical and theoretical perspectives on animal cognition* (pp. 197-204). Cambridge, MA: MIT Press.

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*(3), 238-243.

Buckley, P. B., & Gillman, C. B. (1974). Comparisons of digits and dot patterns. *Journal of Experimental Psychology, 103*, 1131-1136.

Bullock, M., & Gelman, R. (1977). Numerical reasoning in young children: The ordering principle. *Child Development, 48*, 427-434.
Campbell, J. I. (1999). Division by multiplication. *Memory & Cognition, 27*(5), 791-802.
Campbell, J. I., & Fugelsang, J. (2001). Strategy choice for arithmetic verification: effects of numerical surface form. *Cognition, 80*(3), B21-B30.
Campbell, J. I., & Gunter, R. (2002). Calculation, culture, and the repeated operand effect. *Cognition, 86*(1), 71-96.
Campbell, J. I. D. (2003, in preparation)). Cognitive architectures for numerical cognition. In J. I. Campbell (Ed.), *Handbook of mathematical cognition*: Psychology Press.
Campbell, J. I. D., & Fugelsang, J. (2001). Strategy choice for arithmetic verifcation: effects of numerical surface form. *Cognition, 80*(3), B21-B30.
Campbell, J. I. D., & Graham, D. J. (1985). Mental multiplication skill: Structure, process, and acquisition. *39*, 338-366.
Carey, S. (2001a). Cognitive foundations of arithmetic: Evolution and ontogenesis. *Mind and Language, 16*, 37-55.
Carey, S. (2001b). On the very possibility of discontinuities in conceptual development. In E. Dupoux (Ed.), *Language, brain, and cognitive development* (pp. 304-324). Cambridge, MA: MIT Press.
Carey, S. (in press). Evolutionary and Ontogenetic Foundations of Arithmetic. *Mind and Language.*
Carr, J. A. R., & Wilkie, D. M. (1997). Rats use an ordinal timer in a daily time-place learning task. *Journal of Experimental Psychology: Animal Behavior Processes, 23*(2), 232-247.
Catania, A. C. (1963). Concurrent performances: a baseline for the study of reinforcement magnitude. *Journal of the Experimental Analysis of Behavior, 6*, 299-300.
Church, R. M., & Meck, W. H. (1984). The numerical attribute of stimuli. In H. L. Roitblatt, T. G. Bever & H. S. Terrace (Eds.), *Animal cognition* (pp. 445-464). Hillsdale, NJ: Lawrence Erlbaum Assocs.
Church, R. M., Meck, W. H., & Gibbon, J. (1994). Application of scalar timing theory to individual trials. *Journal of Experimental Psychology: Animal Behavior Processes, 20*(2), 135-155.
Clearfield, M., & Mix, K. (1999). Number versus contour length in infants' discrimination of small visual sets. *Psychological Science, 10*, 408-411.
Cordes, S., Gallistel, C. R., Gelman, R., & Latham, P. E. (in preparation (2003)). Nonverbal arithmetic in humans.
Cordes, S., Gelman, R., Gallistel, C. R., & Whalen, J. (2001). Variability signatures distinguish verbal from nonverbal counting for both large and small numbers. *Psychonomic Bulletin & Review, 8*(4), 698-707.
Davis, H., & Pérusse, R. (1988). Numerical competence in animals: Definitional issues, current evidence, and a new research agenda. *Behav. Brain Scis., 11*, 561-615.
Dehaene, S. (1997). *The number sense*. Oxford: Oxford University Press.
Dehaene, S. (2001). Subtracting pigeons: Logarithmic or linear? *Psychological Science, 12*(3), 244-246.
Dehaene, S. (2002). Single neuron arithmetic. *Science, 297*, 1652-1653.

Dehaene, S., & Changeux, J. P. (1993). Development of elementary numerical abilities: A neuronal model. *Journal of Cognitive Neuroscience, 5*, 390-407.
Dehaene, S., Dehaene-Lambertz, G., & Cohen, L. (1998). Abstract representations of numbers in the animal and human brain. *Trends in Neurosciences, 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 and Performance, 16*, 626-641.
Feigenson, L., Carey, S., & Hauser, M. (2002). The Representations Underlying Infants' Choice of More: Object Files Versus Analog Magnitudes. *Psychological Science*, *13*, 150-156.
Feigenson, L., Carey, S., & Spelke, E. (2002). Infants' discrimination of number vs. continuous extent. *Cognitive Psychology, 44*(1), 33-66.
Fuson, K. (1988). *Children's counting and concepts of number*. New York: Springer Verlag.
Gallistel, C. R. (1990). *The organization of learning*. Cambridge, MA: Bradford Books/MIT Press.
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., & Gibbon, J. (2000). Time, rate and conditioning. *Psychological Review*, *107*, 289-344.
Gallistel, C. R., King, A., & McDonald, R. (in press). Sources of Variability and Systematic Error in Mouse Timing Behavior. *Journal of Experimental Psychology: Animal Behavior Processes*.
Gallistel, C. R., Mark, T. A., King, A., & Latham, P. E. (2001). The Rat Approximates an Ideal Detector of Changes in Rates of Reward: Implications for the Law of Effect. *Journal of Experimental Psychology: Animal Behavior Processes, 27*, 354-372.
Geary, D. C., Widman, K. F., & Little, T. D. (1986). Cognitive addition and mulitplication: Evidence for a single memory network. *Mem. Cog., 14*, 478-487.
Gelman, R. (1972). Logical capacity of very young children: number invariance rules. *Child Development, 43*, 75-90.
Gelman, R. (1977). How young children reason about small numbers. In N. J. Castellan, D. B. Pisoni & G. R. Potts (Eds.), *Cognitive theory* (Vol. 2, pp. 27-61). Hillsdale, NJ: Erlbaum.
Gelman, R. (1993). A rational-constructivist account of early learning about numbers and objects. In D. Medin (Ed.), *Learning and motivation. Vol 30* (Vol. 30, pp. 61-96). New York: Academic Press.
Gelman, R., & Brenneman, K. (1994). First principles can support both universal and culture-specific learning about number and music. In L. Hirschfeld & S. Gelman (Eds.), *Mapping the mind: domains, culture and cognition* (pp. 369-390). Cambridge: Cambridge University Press.
Gelman, R., & Cordes, S. A. (2001). Counting in animals and humans. In E. Dupoux (Ed.), *Language, brain, and cognitive development: Essays in honor of Jacques Mehler* (pp. 279-301). Cambridge, MA: MIT Press.

Gelman, R., & Gallistel, C. R. (1978). *The child's understanding of number*. Cambridge, Mass.: 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: Erlbaum.
Gelman, R., & Williams, E. (1998). Enabling constraints on cognitive development. In D. Kuhn & R. Siegler (Eds.), *Cognition, perception and language. Vol 2. Handbook of child psychology. 5th ed.* (pp. 575-630). New York: Wiley & Sons.
Gibbon, J., & Church, R. M. (1981). Time left: linear versus logarithmic subjective time. *Journal of Experimental Psychology: Animal Behavior Processes, 7*(2), 87-107.
Harper, D. G. C. (1982). Competitive foraging in mallards: ideal free ducks. *Anim. Behav., 30*, 575-584.
Hartnett, P., & Gelman, R. (1998). Early understandings of numbers: Paths or barriers to the construction of new understandings? *Learning & Instruction, 8*(4), 341-374.
Hauser, M. D. (2001). What do animals think about numbers? In *Sherman, Paul W (Ed); Alcock, John (Ed) (2001) Exploring animal behavior: Readings from American Scientist (3rd ed) (pp 76-83) iv, 316pp*.
Herrnstein, R. J. (1961). Relative and absolute strength of response as a function of frequency of reinforcement. *Journal of the Experimental Analysis of Behavior, 4*, 267-272.
Herrnstein, R. J., & Vaughan, W. J. (1980). Melioration and behavioral allocation. In J. E. R. Staddon (Ed.), *Limits to action: The allocation of individual behavior* (pp. 143-176). New York: Academic.
Hinrichs, J. V., Yurko, D. S., & Hu, J. M. (1981). Two-digit number comparison: Use of place information. *J. Exper. Psychol.: Hum. Perc. Perf., 7*, 890-901.
Jensen, E. M., Reese, E. P., & Reese, T. W. (1950). The subitizing and counting of visually presented fields of dots. *J. Psychol., 30*, 363-392.
Kahneman, D., Treisman, A., & Gibbs, B. J. (1992). The reviewing of object files: Object-specific integration of information. *Cog. Psychol., 24*(175-219).
Kaufman, E. L., Lord, M. W., Reese, T. W., & Volkman, J. (1949). The discrimination of visual number. *American Journal of Psychology, 62*, 498-525.
Keller, J. V., & Gollub, L. R. (1977). Duration and rate of reinforcement as determinants of concurrent responding. *Journal of the Experimental Analysis of Behavior, 28*, 145-153.
Killeen, P. R., & Taylor, T. (2001). How the propagation of error through stochastic counters affects time discrimination and other psychophysical judgments. *Psychological Review, 107*, 430-459.
King, A. S., McDonald, R., & Gallistel, C. R. (2001). Screening for mice that remember incorrectly. *International Journal of Comparative Psychology, 14*, 232-257.
Klahr, D., & Wallace, J. G. (1973). The role of quantification operators in the development of conservation. *Cog. Pschol., 4*, 301-327.
Krantz, D., Luce, R. D., Suppes, P., & Tversky, A. (1971). *The foundations of measurement*. New York: Academic.

Krantz, D. H. (1972). Measurement structures and psychological laws. *175*, 1427-1435.

Leon, M. I., & Gallistel, C. R. (1998). Self-Stimulating Rats Combine Subjective Reward Magnitude and Subjective Reward Rate Multiplicatively. *Journal of Experimental Psychology: Animal Behavior Processes*, *24*(3), 265-277.

Lipton, J. S., & Spelke, E. S. (2003). Origins of number sense: Large-number discrimination in human infants. *Psychological Science*, *14*(5), 396-401.

Luce, R. D. (1990). What is a ratio in ratio scaling. In S. J. Bolanowski & G. A. Gescheider (Eds.), *Ratio scaling of psychological magnitude: In honor of the memory of S. S. Stevens*. Hillsdale, NJ: Lawrence Erlbaum.

Mandler, G., & Shebo, B. J. (1982). Subitizing: an analysis of its component processes. *J. Exp. Psychol. Gen., 11*, 1-22.

Mechner, F. (1958). Probability relations within response sequences under ratio reinforcement. *J. Exp. Anal. Behav., 1*, 109-122.

Menninger, K. (1969). *Number words and number symbols*. Cambridge, MA: MIT Press.

Miller, K. F., Perlmutter, M., & Keating, D. (1984). Cognitive arithmetic: Comparison of operations. *J. exper. Psychol.: Learn. Mem. Cog., 10*, 46-60.

Mix, K. S., Huttenlocher, J., & Levine, S. C. (2002). *Quantitative development in infancy and early childhood*. Oxford: Oxford University Press.

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.

Noel, M.-P. (2001). Numerical cognition. In B. Rapp (Ed.), *The handbook of cognitive neuropsychology: What deficits reveal about the human mind* (pp. pp 495-518).

Olthof, A., Iden, C. M., & Roberts, W. A. (1997). Judgments of ordinality and summation of number symbols by Squirrel monkeys. *Journal of Experimental Psychology: Animal Behavior Processes, 23*(3), 325-339.

Olthof, A., & Roberts, W. A. (2000). Summation of symbols by pigeons (*Columbia livia*): the importance of number and mass of reward items. *Journal of Comparative Psychology, 114*(2), 158-166.

Parkman, J. M. (1971). Temporal aspects of digit and letter inequality judgments. *J. exper. Psychol., 91*, 191-205.

Piazza, M., Giacomini, E., Le Bihan, D., & Dehaene, S. (2003). Single-trial classification of parallel pre-attentive and serial attentive processes using functional magnetic resonance imaging. *Proceedings of the Royal Society (London) Series B*, *270*, 1237–1245.

Platt, J. R., & Johnson, D. M. (1971). Localization of position within a homogeneous behavior chain: effects of error contingencies. *Learn. Motiv., 2*, 386-414.

Pylyshyn, Z. W., & Storm, R. W. (1988). Tracking multiple independent targets: Evidence for a parallel tracking system. *Spatial Vision, 3*, 179-197.

Rumbaugh, D. M., Savage-Rumbaugh, S., & Hegel, M. T. (1987). Summation in the chimpanzee (*Pan troglodytes*). *J. Exp. Psychol.: Anim. Behav. Proc., 13*, 107-115.

Rumbaugh, D. M., & Washburn, D. A. (1993). Counting by chimpanzees and ordinality judgments by macaques in video-formatted tasks. In S. T. Boyese & E. J. Capaldi (Eds.), *The development of numerical competence: Animal and human models* (pp. 87-106). Hillsdale, NJ: Erlbaum.
Savastano, H. I., & Miller, R. R. (1998). Time as content in Pavlovian conditioning. *Behavioural Processes, 44*(2), 147-162.
Scholl, B. J., & Pylyshyn, Z. (1999). Tracking multiple items through occlusion: Clues to visual objecthood. *Cognitive Psychology, 38*, 259-290.
Simon, T. J. (1999). The foundations of numerical thinking in a brain without numbers. *Trends in Cognitive Sciences, 3*(10), 363-364.
Spelke, E., & Dehaene, S. (1999). Biological foundations of numerical thinking. *Trends in Cognitive Sciences, 3*(10), 365-366.
Stevens, S. S. (1951). Mathematics, measurement and psychophysics. In S. S. Stevens (Ed.), *Handbook of experimental psychology* (pp. 1-49). Wiley: New York.
Stevens, S. S. (1970). Neural Events and the Psychophysical Law. *Science, 170*(3962), 1043-1050.
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: Erlbaum Associates.
Sulkowski, G. M., & Hauser, M. D. (2001). Can rhesus monkeys spontaneously subtract? *Cognition, 79*(3), 239-262.
Taves, E. H. (1941). Two mechanismsfor the perception of visual numerousness. *Archives of Psychology, 37*(No. 265).
Terrace, H. S., Son, L. K., & Brannon, E. M. (2003). Serial Expertise of Rhesus Macaques. *Psychological Science, 14*, 66-73.
Uller, C., Carey, S., Huntley-Fenner, G., & Klatt, L. (1999). What representations might underlie infant numerical knowledge? *Cognitive Development, 14*(1), 1-36.
Uller, C., Hauser, M., & Carey, S. (2001). Spontaneous representation of number in cotton-top tamarins (Saguinus oedipus). *Journal of Comparative Psychology, 115*(3), 248-257.
Washburn, D. A., & Rumbaugh, D. M. (1991). Ordinal judgments of numerical symbols by macaques (*Macaca mulatta*). *Psychological Science, 2*, 190-193.
Whalen, J. (1997). *The influence of semantic representations of number on arithmetic fact retrieval.* Unpublished Ph.D., Johns Hopkins, Baltimore.
Whalen, J., Gallistel, C., & Gelman, R. (1999). Nonverbal counting in humans: The psychophysics of number representation. *Psychological Science, 10*(2), 130-137.
Whalen, J., West, V., & Cook, B. (under review, 2003). Don't count on subitizing: Evidence of non-verbal counting of small visual arrays.
Woodworth, R. S., & Schlosberg, H. (1954). *Experimental psychology: Revised edition.* New York: Holt, Rinehart & Winston.
Wynn, K. (1990). Children's understanding of counting. *Cognition, 36*, 155-193.
Wynn, K. (1992a). Addition and subtraction by human infants.
Wynn, K. (1992b). Children's acquisition of the number words and the counting system. *Cognitive Psychology, 24*, 220-251.
Xu, F., & Spelke, E. S. (2000). Large number discrimination in 6-month-old infants. *Cognition, 74*(1), B1-B11.