To appear in S. Levinson, & P. Jaisson (Eds.) *Culture and evolution*.

Cambridge, MA: MIT Press

The Cultural and Evolutionary History of the Real Numbers

C. R. Gallistel

Rochel Gelman

Sara Cordes

Department of Psychology
University of California
Los Angeles, CA 90095-1563

The cultural history of the real numbers began with the positive integers. Kronecker is often quoted as saying, "God made the integers; all else is the work of man," by which he meant that the system of real numbers had been erected by mathematicians on the intuitively obvious foundation provided by the integers. Taken as a statement about the cultural history of mathematics, this is beyond dispute. But if this is taken as a claim about the psychological foundations of arithmetic reasoning, then we suggest that here, as in many other areas of psychology, introspection and intuition are poor guides to the inner workings of the mind.

We suggest that it is the system of real numbers that is the psychologically primitive system, both in the phylogenetic and the ontogenetic sense. We review evidence that a system for arithmetic reasoning with real numbers evolved before language evolved. When language evolved, it picked out from the real numbers only the integers, thereby making the integers the foundation of the cultural history of the number. Secondly, we suggest that this ancestral non-verbal real number system becomes operative in the prelinguistic child and makes possible the acquisition of language-mediated counting and language-mediated arithmetic reasoning. It is the foundation on which an individual's language-mediated understanding of what numbers are and what may be done with them rests.

The Formal Relation Between the Integers and the Reals

The number system that can be used to represent continuous (uncountable) quantities is the system of real numbers. It includes the irrational numbers, like $\sqrt{2}$, and the transcendental numbers, like π. It is used by modern humans to represent many distinct systems of continuous quantity--duration, length, area, volume, density, rate, intensity, and so on. Because the system of real numbers is isomorphic to a system of magnitudes, the terms real number and magnitude are used interchangeably. Thus, when we refer to "mental magnitudes" we are referring to a real number system in the brain. Like the culturally specified real number system, the real number system in the brain is used to represent both continuous quantity and numerosity.

Magnitudes and real numbers have the property that there is no way to pick out a successor, the next number in the sequence. Given a line of some length, there is no procedure whereby one could pick out the next longer line. Similarly, given a real number, like, say, 2, there is no procedure that picks out the next real number, although there is, of course, a procedure that picks out the next integer. The real numbers are not discretely ordered but the integers are, so 2 qua real number has no successor, whereas 2 qua integer does.

The discrete ordering of the natural numbers (the positive integers) makes them uniquely suited to represent numerosity, that is countable quantity. The positive integers, however, taken by themselves, rather than as a component of the system of real numbers, have two serious failings, an algebraic failing and a geometric failing. The algebraic failing is that they are not closed under the inverse combinatorial operations of subtraction and division. Subtracting one positive integer from another often fails to yield a positive integer. If only the positive integers are regarded as legitimate numbers, then subtraction can be legitimately performed only when it is known to yield a positive integer. But in the course of algebraic reasoning, it is often desirable to subtract one unknown number from another. If only positive integers are allowed as numbers, then this maneuver will be of doubtful legitimacy, because one will not know whether the subtrahend is larger or smaller than the minuend. The division of one unknown number by another is likewise suspect, because only rarely does dividing one positive integer by another yield a positive integer.

The lack of closure in the system of natural numbers provided much of the motivation that drove the cultural expansion of that system to include zero, the negative integers, and the rational numbers. The rational numbers, include all the numbers that may be expressed as the proportion between two integers, that is, the fractions, including the improper fractions like 71/53.

The geometric failing of the integers and their offspring the rational numbers arises when we attempt to use proportions between integers to represent proportions between continuous quantities, as, for example when we say that one person is half again as tall as another, or one farmer has only a tenth as much land as another. These locutions show the seemingly natural expansion of the integers to the rational numbers, numbers that represent proportions. This expansion seemed so natural and unproblematic to the Pythagoreans that they believed that the natural numbers and the proportions between them (the rational numbers) were the inner essence of reality, the carriers of all the deep truths about the world. They were, therefore, greatly unsettled when they discovered that there were geometric proportions that could not be represented by a rational number, for example, the proportion between the diagonal and the side of a square. The Greeks proved that no matter how fine you made your unit of length, it would never go an integer number of times into both the side and the diagonal. Put another way, they proved that the square root of two is an irrational number, an entity that cannot be constructed from the natural numbers by a finite procedure.

As the name they gave to these would-be numbers implies, the Greeks found the existence of irrational numbers contrary to reason. They sensed that number ought to be able to represent geometric proportions like this, and they were right. They were right because if we follow the impulse to create a closed algebraic system to its natural end, then we are led to the system of real numbers, indeed, eventually to the system of complex numbers. The system of real numbers has numbers to represent every geometric proportion. Thus, in the process of creating the numbers needed to guarantee algebraic closure, mathematicians created the numbers needed to represent every possible proportion. Our thesis is that this cultural creation of the real numbers was a Platonic rediscovery of the underlying non-verbal system of arithmetic reasoning. The cultural history of the number concept is the history of our learning to talk coherently about a system of reasoning with real numbers that predates our ability to talk, both phylogenetically and ontogenetically.

Evidence for the Ancestral Status of the Real Numbers

Other Vertebrates Measure and Remember Uncountable Magnitudes

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 when a stimulus signaling the possible availability of food comes on. When pigeons are the subjects, the stimulus is the illumination of a key on the wall of the experimental chamber. When the subjects are rats, the stimulus is the extension into the cage of a lever. On 25-50% of the trials, the key or lever is armed at a fixed latency after the onset of the stimulus. Pecking or pressing before the key or lever is armed is pointless, but the first peck or press after the arming delivers food. On the remaining 50-75% of the trials, however, the key or lever is not armed. On these trials, the key remains illuminated or the lever remains extended for between 4 and 6 times longer than the arming latency. Pecking or pressing after the arming latency has elapsed is pointless, because if there has been no reward at that latency, then there will be none on that trial.

Peak-procedure data come from the unrewarded trials. On such trials, the subject abruptly begins to peck the key or press the lever some while before the arming latency (when it judges arming to be nigh) and continues to peck or press for some while after before abruptly stopping (when it judges that the arming latency has past). The interval during which the subject pecks or presses brackets its subjective estimate of the arming latency. Representative data are shown in Figure 1.

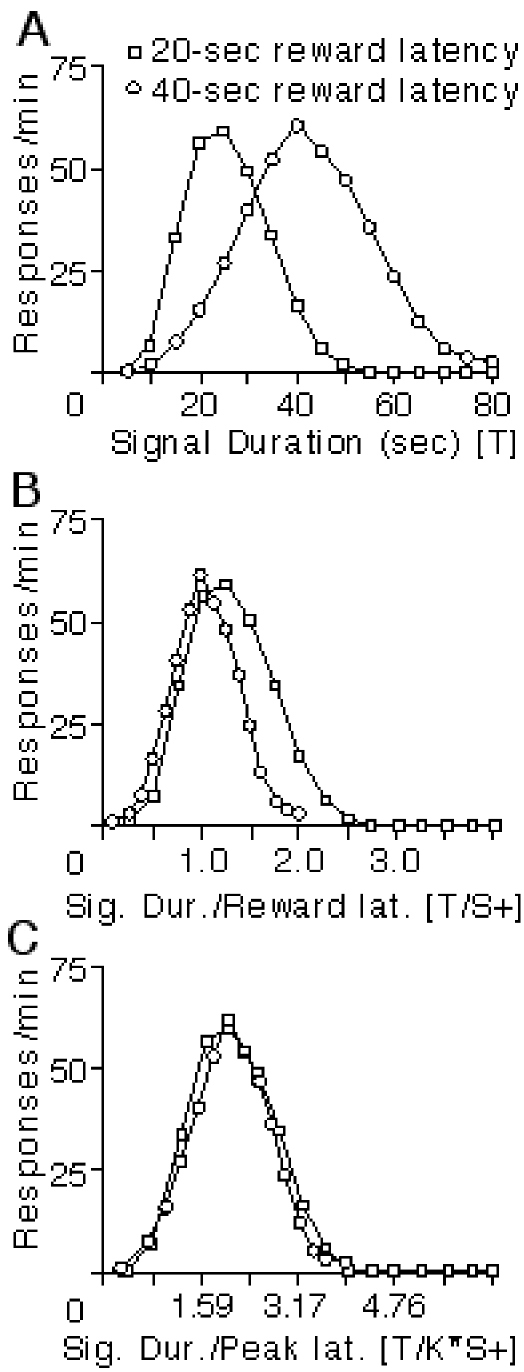

Figure 1. *Representative peak procedure data. Rat was the subject. In one block of many trials, the arming latency was 20 sec; in another, it was 40 sec. A. The original data. B. Data plotted as a proportion of the arming latency. C. Data plotted as a proportion of the latency at the mode of the distributions in A. Because the variability in the onsets and offsets of responding is proportional to the remembered arming latency, the distributions superimpose when plotted as a proportion of the modal latency.*

Figure 1A, shows smooth seemingly increases and decreases in response rates on either side of the arming latency. The smoothness is an averaging artifacts. On any one trial the onset and offset of responding is abrupt. The temporal locus of these onsets and offsets varies from trial to trial. Averaging across trials gives these approximately normal distributions. The curves in Figure 1A are best read as showing the probability that the subject will be responding as a function of the time elapsed since the warning signal came on. The mode of the distribution (the latency at which the distribution peaks) is the latency at which the subject is maximally likely to be responding. This latency does not necessarily coincide with the actual arming latency, because individual subjects often show small proportional errors in the mode; they misremember experienced durations by some multiplicative factor slightly greater or smaller than 1.

The distribution obtained with the 40 second arming latency is broader than the distribution obtained with the 20 second latency. As shown in Figure 1C, the broadening of the distribution at longer arming latencies is proportional to the remembered arming durations (the mode of the distribution), not to the actual arming durations (Figure 1B). When mean response rates are plotted against the elapsed proportion of the modal latency, the distributions obtained at different arming latencies superpose (Figure 1C). Thus, 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 the memory for durations is called scalar variability.

Scalar variability is a ubiquitous property of remembered mental magnitudes. It seems to be best explained by the assumption that the neural signals that come from the reading of a memory show trial-to-trial (reading-to-reading) variability, just as do the neural signals that come from the action of a stimulus (Gallistel, 1999; Gallistel & Gibbon, 2000). In other words, the reading of a mental magnitude in memory is a noisy process, and the noise is proportional to magnitude being read.

Other Vertebrates Count and Remember Numerosity

Rats, pigeons and monkeys also count and remember numerosities (see Dehaene, 1997; Gallistel, 1990; Gallistel & Gelman, 2000, for reviews; e.g. Roberts, Coughlin, & Roberts, 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 2. 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, the memory for number also exhibits scalar variability

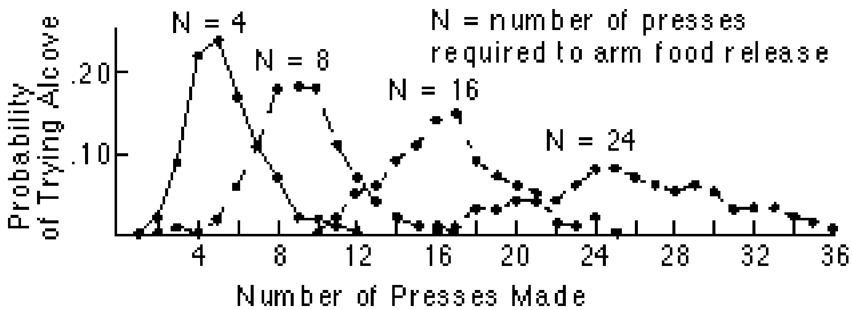

Figure 2. *The probability of breaking off to try the feeding alcove as a function of the number of presses made on the arming lever and the number required to arm the food-release beam at the entrance to the feeding alcove. Subjects were rats. Redrawn from* (Platt & Johnson, 1971) *by permission of the authors and publishers.*

The fact that the memory for numerosity exhibits scalar variability suggests that numerosity is represented in the brains of non-verbal vertebrates like rats, pigeons and monkeys by mental magnitudes, that is by 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 Figure 2. For example, the bit-pattern symbol for fifteen is 01111 while for sixteen it is 10000. Although the numbers are adjacent, 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. Similarly, a scribe copying a handwritten English text is presumably more likely to confuse "seven" and "eleven" than to confuse "seven" and "eight". Thus, the nature of the variability in a remembered target number implies that what is being remembered is a magnitude--a real number.

Numerosity and Duration are Represented by Comparable Mental Magnitudes

Meck and Church 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 (1983) 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 3). At the end of the count, the resulting accumulation is read into memory where it represents the numerosity.

The model in Figure 3 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 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.

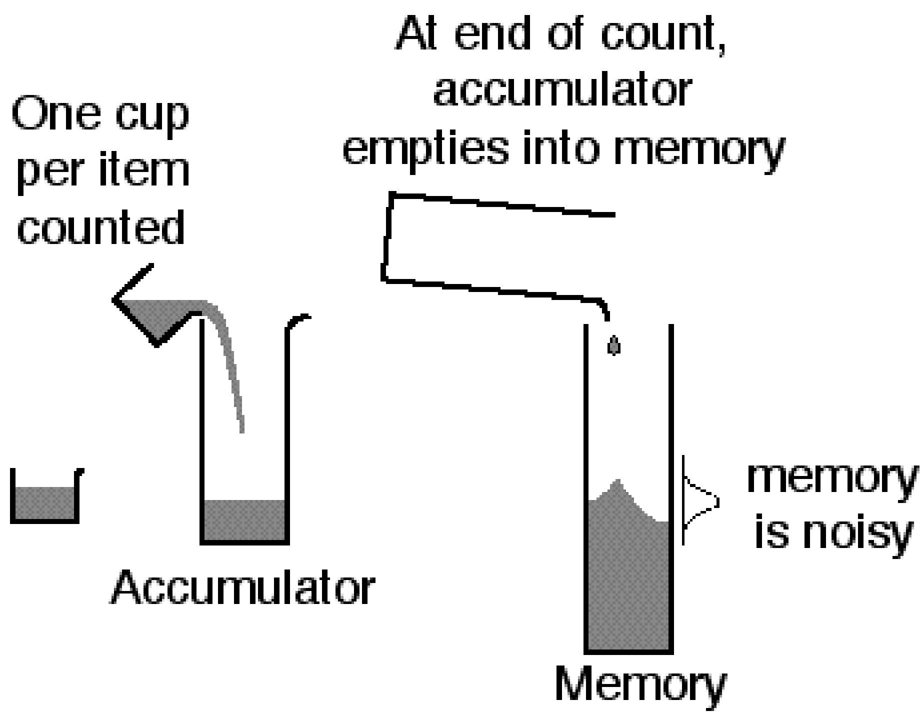

Figure 3. *The accumulator model for the non-verbal counting process. At each count, the brain increments a quantity, an operation formally equivalent to pouring a cup into a graduate. The final magnitude (the contents of the graduate at the conclusion of the count) is stored in memory, where it represents the numerosity of the counted set. Memory is noisy, which is to say that the values read from memory on different occasions vary. The variability in the values read is proportional to the mean value of the distribution (scalar variability).*

In the course of their work, Meck and Church (1983) were able to estimate the scale factor relating the mental magnitude scales for numerosity and duration. If the mental magnitude, $\hat{n}$, representing numerosity, n, is proportional to the numerosity represented, that is, if $\hat{n} = k_1 n$, and if the mental magnitude, $\hat{d}$, representing duration, d, is proportional to the duration represented, that is, if $\hat{d} = k_2 d$, then, letting the mental magnitudes be equal (letting $k_1 n = k_2 d$), gives $n = \frac{k_2}{k_1} d$. The scale factor $\frac{k_2}{k_1}$ tells us which numerosities are "mentally equivalent" to which durations, where by, "mentally equivalent" we mean "represented by real numbers of the same magnitude." Meck and Church got a value of about 0.2 s^{-1} for $\frac{k_2}{k_1}$, meaning that each unit increase in the mental magnitude representing numerosity corresponds to the increase generated by prolonging a duration by 0.2 seconds. Thus, an interval 2 seconds in duration generates a mental magnitude that, if it were used to represent a numerosity, would represent a numerosity of 10.

Knowing this scale factor, enabled Meck and Church (1983--see also Meck, Gibbon and Church, 1985)--to do an experiment directly demonstrating that the mental magnitudes representing numerosity and the mental magnitudes representing duration were interchangeable. They first taught rats to choose one lever after hearing a noise of two seconds duration and the other lever after hearing a noise of four seconds duration. The rats learned to make this discrimination under partial reinforcement conditions, that is, on 50% of the trials even a correct choice did not produce reward. Because the animal is thereby accustomed to not receiving a reward on many of the trials when it makes a correct choice, this procedure allows the experimenter to give the trained animal unrewarded "probe" trials. On a probe trials, the animal gets a stimulus different from the training stimuli (the reference stimuli) and the question is, how will it judge the probe stimulus? Which reference stimulus will it judge to be "more like" the probe stimulus? Its judgment is indicated by which lever it chooses. In this case, the reference stimuli are represented by the mental magnitudes corresponding to durations of 2 and 4 seconds.

The probe stimuli in the experiment were long sequences of noise bursts. The sequences were much longer than 4 seconds, so the mental magnitudes representing duration would be much greater than either of the reference magnitudes. If the magnitudes representing the duration of a probe sequence were compared to the reference magnitudes, they would be more similar to the greater of those two reference magnitudes, albeit basically outside the range of the reference magnitudes. However, the number of bursts in the sequence ranged from 10 to 20, which is to say that the mental magnitudes produced by the nonverbal counting of these sequences covered the same range of mental magnitudes produced by durations ranging from 2 seconds to 4 seconds. Meck and Church hoped that their subjects would make their choice between the levers by comparing the mental magnitudes generated by counting the bursts to the reference mental magnitudes, ignoring the fact that the reference magnitudes represented durations while the magnitudes being compared to them on these trials represented numerosity. This was the result they obtained: when there were there were 10 or close to 10 noise bursts, the rats chose the lever corresponding to the shorter duration; when there were 20 or close to 20 bursts, they chose the lever corresponding to the longer duration. This is strong evidence that a common system of mental magnitudes is used to represent both uncountable (continuous) and countable (discrete) quantity.

Other Vertebrates Reason Arithmetically with the Mental Magnitudes Representing both Countable and Uncountable Quantity

We have repeatedly referred to the real number system because numbers acquire their representational utility by virtue of the fact that they can be arithmetically manipulated--added, subtracted, multiplied, divided and ordered. From a formal point of view, if mental magnitudes could not be arithmetically manipulated, there would be no justification for calling them numbers. From a formalist perspective, numbers just are entities that are arithmetically manipulable. Thus, when we refer to the real numbers in the brain we mean magnitudes that can be arithmetically processed in the brain and that refer to countable and uncountable quantities.

There is a considerable experimental literature demonstrating that laboratory animals reason arithmetically with real numbers. 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. Here we summarize a few of the relevant studies.

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 simply 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 countable). 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 just like the magnitudes generated by the non-verbal counting of the caches.

Subtracting durations, numerosities and rates. 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 some unpredictable point in the course of a trial, the opportunity to choose suddenly terminates, and the subject must then endure the delay associated with the option it was exercising at that moment. If it was pecking the standard key, it is stuck with the standard delay; if it was pecking the time-left key, it is stuck with the time left. The initial value of the time left--the value at the beginning of a trial-- is much longer than the standard delay, but it grows ever shorter as the trial goes on, because the time left is the initial value minus the time so far elapsed in a trial. Therefore, the longer a trial persists before the loss of choice, the better the time-left option becomes relative to the standard. 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.

In the number-left procedure (Brannon, et al., in press), 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 varied 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 pecks generate 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 key, 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 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.

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.

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 for the hypothesis that the system for arithmetic reasoning with mental magnitudes is closed under subtraction.

Dividing Number 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 (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 (Herrnstein & Vaughan, 1980).

Despite the plausibility of this explanation, it has never been possible to construct a model based on this assumption that predicted the details of the behavior at all well (Lea & Dow, 1984). For one thing, because of the way rewards are scheduled in the customary experimental paradigm, the return to be expected from a given location increases the longer that location has gone unvisited. If the subject's decisions to leave one location to sample the other were sensitive to the returns on its behavioral investments, then the probability of leaving ought to get higher as the duration of a stay increases, but it does not. The probability of leaving is Markovian (Heyman, 1979); that is, it looks statistically as if the subjects repeatedly flipped a coin in order to decide whether to leave (Gibbon, 1995; Gibbon, Church, Fairhurst, & Kacelnik, 1988), with the outcome of later flips being independent of the outcome of earlier flips. This discovery led to the suggestion that matching behavior is an unconditioned response to the experience of a given ratio of rates of reward (Heyman, 1982).

Recently, Gallistel, et al. (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 anywhere near so rapidly as they in fact do adjust if they were discovering by trial and error the ratio of expected stay durations that equated their returns. This means that Heyman (1982) was right: matching behavior is an unconditioned or preprogrammed response to the experience of different rates of reward. The importance of this in the present context is that a rate is <u>countable</u> quantity--the number of rewards received in a given interval--<u>divided by</u> an <u>uncountable</u> quantity--the duration of the duration of the given interval.

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 all 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) and uncountable quantity by a different system (a system of continuously ordered magnitudes), it would not be possible to estimate rates.

<u>Multiplying Rate by Magnitude.</u> 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 mental ordering of the mental magnitudes, because the subject's choice depends on the ordering of the resulting magnitudes. 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. Perhaps most importantly, it was impossible to teach subjects to touch the arrays in an order that did not conform to the order of the numerosities. This implies that the ordering of the numerosities is highly salient for a monkey.

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 (real numbers). 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.

Evidence that Humans Represent Numerosity with Mental Magnitudes

The Symbolic Size and Distance Effects

From an evolutionary standpoint, it would be odd if humans did not share with their remote vertebrate cousins (pigeons) and near vertebrate cousins (macaques and chimpanzees) the mental machinery for representing countable and uncountable quantity by means of a system of real numbers. That humans do represent numbers with mental magnitudes was suggested by Moyer and Landauer (1967; 1973) when they discovered what has come to be called the symbolic size and distance effects. 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 (the distance effect), and, for a fixed difference, the greater the magnitude of the two numbers, the longer it takes to judge their order (the size effect).

Moyer and Landauer suggested that the effects of numerical magnitude on reaction time implied that Weber's law applied to symbolically represented numerical magnitude.

Weber's law is that the discriminability of two quantities is a function of their ratio. Moyer and Landauer (1973) suggested that symbolically represented numbers were translated into mental magnitudes in order to judge the numerical ordering of the represented numbers, and that noise in the mental magnitudes made it more difficult to determine which magnitude was greater. This, together with the assumption that more difficult discriminations take longer to make, explains the symbolic distance effect.

Weber's law is often taken to imply logarithmic compression in the mapping between objective magnitude and subjective magnitude. This assumption, together with the usually implicit (and physically implausible) assumption that the noise in mental magnitudes is magnitude independent, gives Weber's law, because magnitudes with a given ratio when mapped onto a logarithmic scale are separated by a given distance.

The assumption of logarithmic compression is, however, inconsistent with the results of the time-left and number-left experiments discussed above. If mental magnitudes were proportional to the logarithms of objective magnitudes, then equal differences in mental magnitudes would correspond to equal ratios between the corresponding objective magnitudes. Thus, when subjects are asked to compare the subjective difference between two numbers against some standard, the results should depend not on the objective difference between the two magnitudes but rather on their ratio. If the subjective difference between six and three is equal to the subjective magnitude of three, then the subjective difference between sixty and thirty should likewise be equal to the subjective magnitude of three, because $\log(60) - \log(30) = \log(6) - \log(3) = \log(2)$. This is implausible a priori, and it is contrary to experimental fact. In both the time-left and number-left experiments, the point of subjective equality increased linearly with the (objective) difference between the initial and standard magnitudes when their ratio was held constant.

The just mentioned experimental results imply that the mental magnitudes representing both numerosity and duration are approximately scalar mappings of the objective magnitudes. If these mental magnitudes have scalar noise, then this, too, gives Weber's law: the discriminability of two such magnitudes will depend on their ratio, not their difference. Thus, the symbolic size and distance effects are consistent with the assumption that when humans judge numerical order, they represent number by mental magnitudes with scalar variability, just as do other vertebrates. What is unique in humans (and a few human-trained laboratory subjects) is that these mental magnitudes can be evoked by way of a learned mapping from the culturally defined linguistic and graphemic symbols for the integers to the mental magnitudes that represent numerosity in the non-verbal or preverbal brain.

Non-verbal Counting in Humans

Given the evidence from the symbolic distance effect 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 (Figure 4): 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.

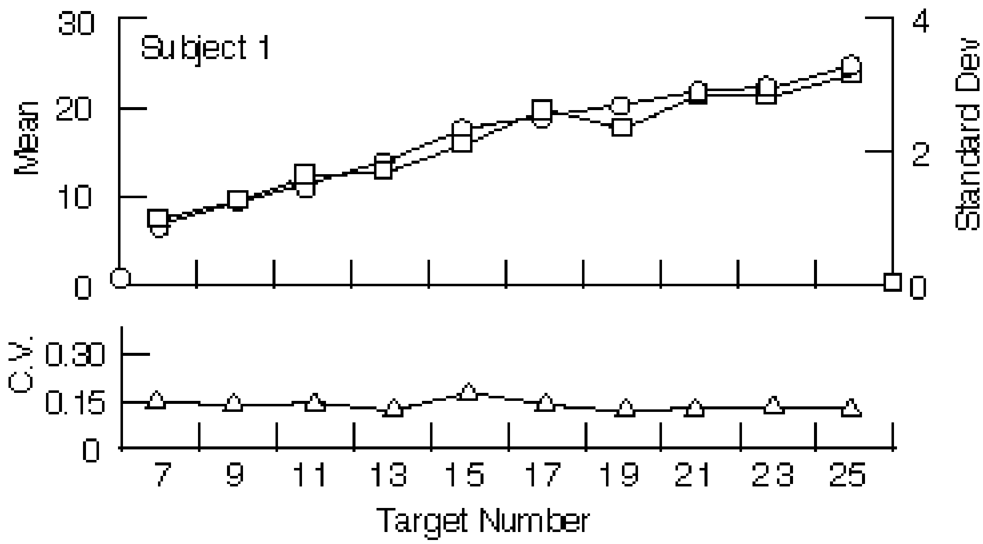

Figure 4. *Representative data from the human non-verbal counting experiment by Whalen, et al.* (1999). *The mean number of presses made increased in proportion to the target number (top panel, left ordinate) and so did the variability (top panel, right ordinate), so the coefficient of variation was constant (bottom panel). Human non-verbal counting exhibits the same scalar variability as non-human counting and timing (compare with Figures 1 and 2).*

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 the 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 about 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 recent further experiments by Cordes, Whalen, Gallistel, and Gelman (in preparation) speak more directly to this issue. Cordes, et al 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. These manipulations certainly suppress audible articulation. Cordes, et al. recorded the subjects while they pressed, and there was no audible counting. These manipulations presumably suppress subvocal articulation as well, because it does not seem likely that someone can subvocally articulate one word (a count word) at the same moment they audibly articulate a different, non-count word. The second of the two manipulations for suppressing articulatory coding--saying "the" with every press-- was particularly effective in that subjects found it easy to do and it required them to articulate a non-count word at the very moment when they would articulate a count word if they were verbally counting.

In control experiments, subjects were asked to count their presses out loud in one of two ways: the conventional way, fully pronouncing each count word; and a way that subjects found much easier, which was to use only the single digit count words, silently keeping track of the tens count. 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 5 (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 fully pronouncing each count word 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 considerably strengthen the evidence against the hypothesis that subjects in the non-verbal counting conditions were subvocally counting.

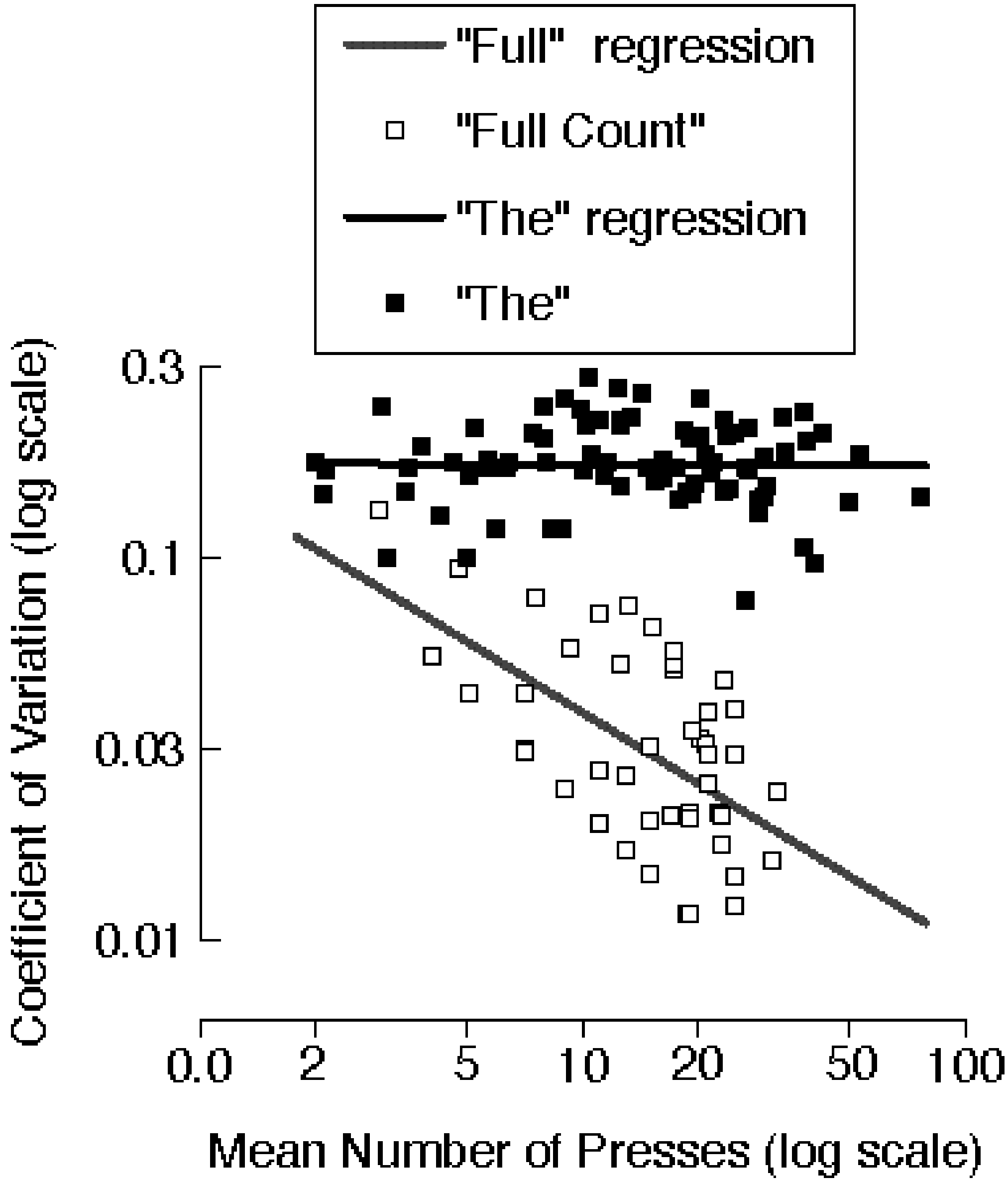

Figure 5. *The coefficients of variation (σ/μ) are plotted against the numbers of presses for the conditions in which subjects counted non-verbally and for the condition in which they fully pronounced each count word (double logarithmic coordinates). In the former condition, there is scalar variability, 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. Data from Cordes, et al. (in preparation).

A second feature of the data from some subjects in the Cordes, et al experiment further strengthens the evidence against the sub-vocal counting hypothesis: In two of the subjects, the relation between the target number and the mean number of presses was a power function with an exponent significantly greater than 1, that is, with a slight upward curvature (Figure 6). Power functions describe the relation between objective magnitudes and subjective magnitudes (Stevens, 1970). Their most interesting property is that they preserve equal proportions. Any two pairs of objective magnitudes that have the same objective ratio (for example, 3:2 and 30:20) have the same subjective ratio when the mapping from objective magnitudes to subjective magnitudes is a power function.

This finding is consistent with the scalar version of the magnitude-mapping hypothesis by which the symbolic distance effect is generally explained, the hypothesis that number words and numerals map to the same kind of mental magnitudes that represent continuous quantities like stimulus intensities, and that these remembered mental magnitudes have scalar noise. It suggests that the mapping is constructed by arranging the linguistic symbols along one continuum and the numerical magnitudes along an orthogonal continuum, as in a conventional scatter graph. The locus of points relating the two continua (the points that define the mapping, that is, the function relating the two continua) fall on a power curve whose slope is close to but not necessarily equal to 1. Learning the meaning of the count words, on this hypothesis, involves learning appropriate parameters for the power function that relates the continuum on which the symbols are arranged to the continuum that represents numerosity.

On the other hand, it is hard to understand how sub-vocal counting could yield a power-function relation between the target count and the mean number of presses made that has a slope different from 1. Presumably with vocally based counting, the subject articulates each successive count word coincident with each successive press until the word articulated matches the articulation of the target number. In their haste some subjects might skip (fail to count) some percentage of the presses they make, so that they systematically undercount their presses, but this would lead to scalar error not an error in the exponent. The plot of the log of the mean number of presses against the log of the target number should still be a straight-line with slope 1, because any scalar relation looks like this on a log-log plot.

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.

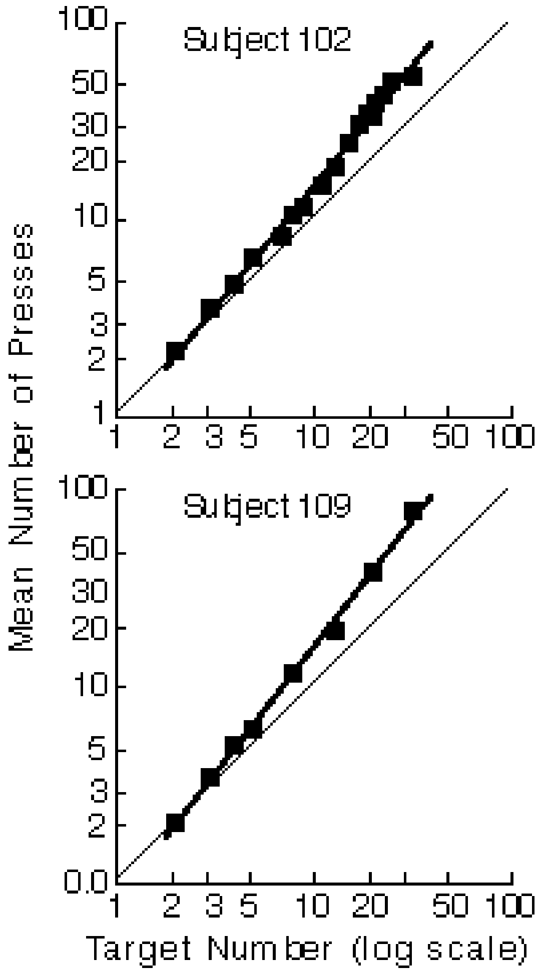

Figure 6. *The mean number of presses is plotted against the target number for the two of the four subjects in Cordes, et al who responded for targets ranging from 2 to 32 (double logarithmic coordinates). For these two subjects, the points fall on straight regression lines (thick lines) whose slopes are slightly but very significantly greater than 1 (thin lines). Note that this deviation from a scalar relation (a line with a slope of 1) is evident even in the small number range (2- 5). Data from Cordes, et al. (in preparation).*

Where the Integers Come From

On our hypothesis, it is the real numbers, not the integers, that are the primitive foundation of numerical reasoning. The integers are a special case whose prominence in the cultural history of the numbers derives from the discrete character of language. When a discrete system like language attempts to represent quantity, it will find it much easier to represent countable (discrete) quantity than to represent uncountable (continuous) quantity.

Gallistel and Gelman (1992)argue that the non-verbal domain of numerical estimation and arithmetic reasoning in animals is operative in the very young child and provides the foundation on which the child's understanding of verbally mediated arithmetic estimation and reasoning is based. The non-verbal mechanism for establishing reference is the accumulator counting mechanism, which produces mental magnitudes that represent numerosities. By "represent", we mean that the symbols generated (the mental magnitudes) both refer to numerosities and enter into arithmetic reasoning operations.

The situations that trigger verbal counting also trigger non-verbal counting. Importantly, the verbal counting process is homomorphic to the non-verbal counting process. In particular, both processes have effective procedures for defining successor symbols. Each step in the verbal process summons the next word from the list of count words. Each count in the non-verbal process defines a next magnitude. Thus, the products of both processes are discretely ordered, that is, the magnitude that results from the next step is always one fixed increment greater than the magnitude that results from the previous step, and the word used in the next verbal count is always one item later in the list than the preceding word. Finally, the product of the verbal process as a whole, that is, the word used in the last step, represents the numerosity of the set being counted, as does the magnitude produced by the final increment in the non-verbal counting process.

Gallistel and Gelman (1992) argue that the child perceives the homomorphism between the non-verbal and the verbal counting process, and this leads assumes that the words used in the counting process represent the same aspect of the world as do the mental magnitudes obtained from the non-verbal counting process. This means two things: First, it means that the child thinks that the counting words refer to the same things in the world that the mental magnitudes representing numerosity refer to, namely, countable quantities. Second, it means that the rules of inference (the rules of operation) governing the magnitudes will be the rules of inference that govern the use of number words. This is the structural mapping (homomorphism) between the verbal and the non-verbal systems.

In short, we suggest that the integers are picked out by language because they are the magnitudes that represent countable quantity. Countable quantity is the only kind of quantity that can readily be represented by a system founded on discrete symbols, as language is. It is language that makes us think that God made the integers, because the learning of the integers is the beginning of linguistically mediated mathematical thinking about both countable and uncountable quantity.

The Problem of Exact Equivalence

This hypothesis raises interesting questions about where some of our intuitive convictions about quantity come from. One of these concerns our concept of exact equivalence. Empirically, there is no such thing as exact equivalence among uncountable quantities, as every dressmaker knows. Two measured quantities are never exactly the same. If we relied simply on our experience of uncountable quantity, we would not have a concept of exact equivalence, because no two experienced lengths and no two experienced durations are ever exactly the same.

Nonetheless, we believe that when equals are added to equals, the results are equal. We believe this despite the fact that it may or may not be true of mental magnitudes, depending upon whether the non-verbal system for reasoning with magnitudes recognizes equivalence (substitutability) and how it decides whether two magnitudes are equivalent (substitutable). On the face of it, the mental magnitude generated by adding a unit magnitude (the counting increment) to the remembered magnitude corresponding to "5" on one occasion will not be exactly the same mental magnitude obtained by adding a unit magnitude to this "same" remembered magnitude on another occasion, because the remembered magnitude itself varies from occasion to occasion. This realization has been an obstacle to the more general acceptance of the hypothesis that the real numbers are the psychologically primitive system for representing both uncountable and countable quantity (Carey, 1998, in press; Carey & Spelke, in press; Hauser & Carey, 1998; Leslie, in press; Leslie, Xu, Tremoulet, & Scholl, 1998).

At least three answers to the problem of equivalence suggest themselves. First, a notion of quantitative equivalence based on real numbers must be essentially a statistical notion: two noisy mental magnitudes must be judged to represent equivalent quantities only if they are not decideably different, that is, if they are not reliably orderable. A system that works with real numbers cannot determine equivalence by exact comparison, because, given the noise in the representational system itself, no two mental magnitudes ever match exactly. Two noisy magnitudes can, however, be so close that repeated attempts to determine their ordering leads to the conclusion that they cannot be reliably ordered. This can be taken as equivalent to the conclusion that they are substitutable one for the other. In other words, if a and b are magnitudes, then $a = b$ just in case neither $a > b$ nor $b > a$, or, just in case $a \geq b$ and $b \geq a$. On this account, we believe that when equals are added to equals, the results are equal because it is a truth about our nonverbal processing of mental magnitudes, whose processing of order and equivalence is adapted to the noisiness of the symbols being processed.

A slightly different answer that suggests itself is that the preverbal system of arithmetic reasoning makes use of computational shortcuts that have implicit in them principles about the outcomes of arithmetic processing. A system that reasoned arithmetically with noisy magnitudes might implicitly assume--rather than empirically test for--the equivalence of adding a unit magnitude to equivalent remembered magnitudes. When dealing cards, for example, it might not compute through each round the numbers of cards the players have and whether those numbers are equal. It might take care only to deal one and only one card to every player on each round. If it operates as if this care guaranteed numerical equality, then it operates in accord with the principle that when equals are added to equals the results are equal. This principle forestalls the need to actually do the sums and compare them at the end of every round. The system does not have to keep four running sums because it knows that the sums remain equal so long as it continues to add equal increments to each on every round.

As the above example shows, implicit principles about the outcomes of arithmetic procedures could effect significant computational and mensurational economies. They may also guide verbally mediated reasoning about quantity.

A third answer that suggests itself is that the discrete nature of the verbal representation of countable quantity is the origin of our notion of exact equivalence. Words are discrete entities: the word "three" is not confusable with an infinite number of other words that are not really "three" but that are arbitrarily "close" to "three." In fact, it is unclear what "close" could mean when it comes to words. Thus, the outcomes of two carefully done verbal counts of the same set will yield the same count word to represent the set, as every bank teller knows. As already suggested, the outcomes of two different non-verbal counts of the same set will not contradict this, because their ordering will not be reliably decideable. Thus, language might not only pick out the integers, it may also highlight a discrete notion of exact equivalence. In the absence of language, exact equivalence may not be an issue. It is not easy to think of a non-algebraic context in which exact equivalence if of any consequence.

In any case, we do not believe that uncertainty about where our notion of exact equivalence comes from should blind us to the experimental evidence that human numerical reasoning uses noisy magnitudes to determine such fundamental things as numerical order--even when it is given linguistic symbols that represent very small numerosities. If our underlying non-verbal symbols for twoness, threeness and fourness, are fundamentally discrete, and if the "4", "3" and "2" acquire their meaning by reference to these discrete non-verbal symbols, then it is hard to see why it takes us longer to decide that fourness is bigger than threeness than it does to decide that fourness is bigger than twoness.

Is There a Discrete Foundation for the Integers in the Perception of Small Numerosities?

It is widely accepted that the symbolic size and distance effects imply that adult humans map linguistic symbols for number to mental magnitudes and that they rely on the comparison of those noisy magnitudes to determine numerical order. At least at present, there is no other explanation for these experimentally well established effects. As already explained, these effects imply that our underlying representation of numerosity has the continuous character of the real numbers rather than the discrete character of the integers. Nonetheless, it is commonly argued that our concept of an integer originates in a preverbal system for representing numerosity that itself uses countable (discrete) rather than uncountable (continuous) sets of symbols (Carey, in press; Carey & Spelke, in press; Leslie, in press; Leslie et al., 1998; Simon, 1999). It is generally assumed that this discrete system only represents small numerosities (four or less) and that the mapping from small numerosities to the discrete mental entities that represent them--the so-called subitizing process--does not employ any form of counting. On this hypothesis, the numerosity of small sets--oneness, twoness, threeness, and fourness-- is directly perceived--like orangeness, cowness, treeness, and forkness. Alternatively, It has been suggested that the numerosity of small sets is implicitly represented by the numerosity of the set of object files that the perceptual system opens, but that numerosity is not explicitly symbolized (Carey, in press; Carey & Spelke, in press).

We see several empirical and theoretical problems for this hypothesis. The existence of a subitizing process for the direct and unvarying apprehension of small numerosities has often been argued for on the basis of empirically shaky claims about the form of the reaction time function for rapid numerical estimation (for example Davis & Pérusse, 1988; Siegler & Robinson, 1982; Trick & Pylyshyn, 1994). It has been claimed either that this function is flat for numerosities between 1 and 4 or 5, or that there is a discontinuity in the slope of this function somewhere at 4 or 5. Neither claim is consonant with the results from several careful determinations of this function (Balakrishnan & Ashby, 1992; Folk, Egeth, & Kwak, 1988), which show that the reaction time to judge a numerosity is longer, the greater the numerosity, and that each increment in reaction time is greater than the preceding increment (that is, the function accelerates over the range from 1 to 4). The reaction time function for adult judgments of numerosity is at least as consistent with a counting model as it is with a direct perception model (Gallistel & Gelman, 1991). If fourness is like forkness and twoness like cowness, as some versions of the discrete-origins hypothesis maintain, then one needs to explain why it takes so much longer to perceive fourness than it does to perceive twoness?

Secondly, if small numerosities were represented differently from large numerosities, then one would expect to see a discontinuity in the psychophysically measurable properties of number representations at the point where one form of representation gives way to the other. The data from the above mentioned experiments by Cordes, et al. (in preparation) with adults counting key presses while saying "the" coincident with each press are relevant here. For four of the subjects, the target numbers included 2, 3, 4 and 5, as well as several larger numbers unequivocally beyond the range of the putative subitizing process. The coefficient of variation was the same for these small numbers as for the large numbers (see Figure 5, filled squares). Moreover, in the two subjects whose mapping from linguistically represented numerosity to mental magnitudes was systematically distorted, this distortion was continuous between the small and large number ranges (Figure 6).

Proving continuity experimentally is like proving the null hypothesis, it cannot be done. However, there is no evidence of discontinuity in the psychophysical evidence from adult humans, and this is hard to reconcile with the hypothesis that small and large numbers are represented in fundamentally different ways.

We think the theoretical problems with the hypothesis are at least as great as the empirical problems. If there are discrete non-verbal representatives of numerosity for only the small numbers--either non-verbal symbols/percepts for oneness, twoness, threeness and fourness or implicit representations by sets of mental entities whose numerosity equals the objective numerosity--then the first question that arises is whether there is any arithmetic processing of these symbols (or sets of object files). Can the percept of "oneness" be mentally added to the percept of "twoness" to get the percept of "threeness"? Can one set of object files be compared to another set of object files to determine if the two sets have the same numerosity? If so, then to effect the comparison of two sets of three each, the mind would have to have six object files in play at once, and this is, as we understand it, thought to be impossible. Can a set of one object file be combined with a set of two object files to yield a set of three object files? If so, then the assumption that these discrete symbols exist only for a few small numbers leads immediately to problems with closure. This system will not be closed even under addition, because there will be no symbol to represent the results of adding "threeness to "threeness." Thus, if operations with these very limited sets of mental symbols are the foundation of numerical understanding, it is a puzzle how we come to believe in the infinite extentensability of number, in the fact that you can always add one more (Hartnett & Gelman, 1998). On the other hand, if these symbols for numerosity--whether explicit or implicit-- cannot be processed arithmetically, then what justification is there for saying that these symbols constitute a numerical representation? If a representation does not enable any of the processes appropriate for numbers, then one must ask why it may be said to be a representation of number.

The second theoretical problem concerns the relation between the two forms of numerical representation, one discrete and one continuous. The two systems would seem to be immiscible for the same reasons that analog and digital computers cannot be hybridized. Although both do arithmetic, they do it in fundamentally different ways. Thus, there is no way of adding a digitally represented magnitude (for example, a bit pattern) to a magnitude represented by a analogical magnitude (for example, a voltage), because the two forms of representation are immiscible. It is hard to see why this same problem does not arise in the developing human mind, if it represents some numbers discretely and others by means of magnitudes. If oneness is represented discretely but tenness is represented by a mental magnitude, how is it possible to mentally add oneness to tenness?

This question--the question of where the human conception of an integer comes from--is currently one of the most controversial in the field of numerical cognition (Carey, in press; Carey & Spelke, in press; Leslie, in press; Leslie et al., 1998; Simon, 1999). Clearly, our hypothesis about the relation between mental magnitudes and the linguistically mediated concept of a number cannot be more widely embraced until consensus is reached on this central question.

Summary

The evidence from experiments that probe the properties of numerical representations in non-verbal animals and humans suggest that there exists a common system for representing both countable and uncountable quantity by means of mental magnitudes formally equivalent to real numbers. These mental magnitudes are arithmetically processed without regard to whether they represent countable or uncountable quantity.

Adult human appear to rely on a mapping from the linguistic symbols for number to these preverbal mental magnitudes, even for answering elementary verbal or written questions like, "Is 3 > 2?". This has led us to suggest that the non-verbal system for arithmetic reasoning with mental magnitudes precedes the verbal system both phylogenetically and ontogenetically and that the verbal symbols for numerosity are given their meaning by reference to the non-verbal mental magnitudes that represent countable quantity. If this suggestion is correct, then the real numbers are the psychologically primitive system, not the natural numbers. The special role of the natural numbers in the cultural history of arithmetic is a consequence of the discrete character of human language, which picks out of the system of real numbers in the brain the discretely ordered subset generated by the nonverbal counting process, and makes these the foundation of the linguistically mediated conception of number.

References

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

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

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.

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

Carey, S. (in press). Evolutionary and Ontogenetic Foundations of Arithmetic. Mind and Language.

Carey, S., & Spelke, E. (in press). Constructing a new representational resource: The integer list representation of number. In L. Bonati & J. Mehler & S. Carey (Eds.), Conceptual

Development-a Reappraisal. Cambridge, MA: MIT Press.

Catania, A. C. (1963). Concurrent performances: a baseline for the study of reinforcement magnitude. Journal of the Experimental Analysis of Behavior, 6, 299-300.

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.

Folk, C. L., Egeth, H., & Kwak, H. (1988). Subitizing: direct apprehension or serial processing? Perc. & Psychophys., 44(4), 313-320.

Gallistel, C. R. (1990). The organization of learning. Cambridge, MA: Bradford Books/MIT Press.

Gallistel, C. R. (1999). Can a decay process explain the timing of conditioned responses? Journal of the Experimental Analysis of Behavior, 71, 264-271.

Gallistel, C. R., & Gelman, R. (1991). Subitizing: The preverbal counting process. In W. Kessen & A. Ortony & F. Craik (Eds.), Memories, thoughts and emotions: Essays in honor of George Mandler (pp. 65-81). Hillsdale, NJ: Lawrence Erlbaum Associates.

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., Mark, T. A., King, A. P., & 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

Gibbon, J. (1995). Dynamics of time matching: Arousal makes better seem worse. Psychonomic Bulletin and Review, 2(2), 208-215.

Gibbon, J., & Church, R. M. (1981). Time left: linear versus logarithmic subjective time. Journal of Experimental Psychology: Animal Behavior Processes, 7(2), 87-107.

Gibbon, J., Church, R. M., Fairhurst, S., & Kacelnik, A. (1988). Scalar expectancy theory and choice between delayed rewards. Psychological Review, 95, 102-114.

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., & Carey, S. (1998). Building a cognitive creature from a set of primitives: Evolutionary and developmental insights. In E. Denise Dellarosa Cummins & E. Colin Allen & et al. (Eds.), The evolution of mind. (pp. 51-106): New York, NY, USA.

Herrnstein, R. J. (1961). Relative and absolute strength of response as a function of frequency of reinforcement. J. Exp. Anal. Behav., 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.

Heyman, G. M. (1979). A Markov model description of changeover probabilities on concurrent variable-interval schedules. Journal of the Experimental Analysis of Behavior, 31, 41-51.

Heyman, G. M. (1982). Is time allocation unconditioned behavior? In M. Commons & R. Herrnstein & H. Rachlin (Eds.), Quantitative Analyses of Behavior, Vol. 2: Matching and Maximizing Accounts (Vol. 2, pp. 459-490). Cambridge, Mass: Ballinger Press.

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.

Lea, S. E. G., & Dow, S. M. (1984). The integration of reinforcements over time. In J. Gibbon & L. Allan (Eds.), Timing and time perception (Vol. 423, pp. 269-277). New York: Annals of the New York Academy of Sciences.

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

Leslie, A. M. (in press). Where do integers come from? In P. Bauer & N. Stein (Eds.), Festschrift for Jean Mandler.

Leslie, A. M., Xu, F., Tremoulet, P. D., & Scholl, B. (1998). Indexing and the object concept: Developing *What* and *Where* systems. Trends in Cognitive Sciences, 2, 10-18.

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.

Meck, W. H., & Church, R. M. (1983). A mode control model of counting and timing processes. Journal of Experimental Psychology: Animal Behavior Processes, 9, 320-334.

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

Moyer, R. S., & Landauer, T. K. (1973). Determinants of reaction time for digit inequality judgments. Bull Psychonom. Soc., 1, 167-168.

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

Roberts, W. A., Coughlin, R., & Roberts, S. (2000). Pigeons flexibly time or count on cue. Psychological Science, 11(3), 218-222.

Savastano, H. I., & Miller, R. R. (1998). Time as content in Pavlovian conditioning. Behavioural Processes, 44(2), 147-162.

Siegler, R. S., & Robinson, M. (1982). The development of numerical understanding. In H. W. Reese & L. P. Lipsitt (Eds.), Advances in child development and behavior (Vol 16). New York: Academic.

Simon, T. J. (1999). The foundations of numerical thinking in a brain without numbers. Trends in Cognitive Sciences, 3(10), 363-364.

Stevens, S. S. (1970). Neural Events and the Psychophysical Law. Science, 170(3962), 1043-1050.

Trick, L. M., & Pylyshyn, Z. W. (1994). Why are small and large numbers enumerated differently? A limited-capacity preattentive stage in vision. Psychological Review, 101(1), 80-102.

Whalen, J., Gallistel, C. R., & Gelman, R. (1999). Non-verbal counting in humans: The psychophysics of number representation. Psychological Science, 10, 130-137.