The generative basis of natural number concepts

Alan M. Leslie, Rochel Gelman and C.R. Gallistel

Department of Psychology and Center for Cognitive Science, Rutgers University, 152 Frelinghuysen Road, Piscataway, NJ 08854, USA

Number concepts must support arithmetic inference. Using this principle, it can be argued that the integer concept of exactly ONE is a necessary part of the psychological foundations of number, as is the notion of the exact equality – that is, perfect substitutability. The inability to support reasoning involving exact equality is a shortcoming in current theories about the development of numerical reasoning. A simple innate basis for the natural number concepts can be proposed that embodies the arithmetic principle, supports exact equality and also enables computational compatibility with real- or rational-valued mental magnitudes.

Number concepts and integer words

If a language provides any explicit number words, these always denote the first few positive integers. In one school of thought, this is evidence that the first few number words are learned by their repeated mention in connection with their referents. These accounts assume the empirical induction of arithmetic principles from a system for representing objects and sets of objects, in conjunction with language learning [1–5], and for critiques thereof [6–8]. An alternative school of thought, which is here taken for granted, is that there is an innate system of arithmetic reasoning with preverbal symbols for both discrete and continuous quantity. On this view, the preverbal system supports arithmetic operations from the outset and directs the learning of the number words with the strict ordering that is a prerequisite for valid counting. Our concern is with the question, what are the indispensable aspects of the preverbal system of arithmetic reasoning?

There is wide agreement that humans, including infants [9,10], share with non-verbal animals a system that represents discrete and continuous quantity with an underlying analogue system that supports arithmetic reasoning [11–15] (Figure 1). What an analogue system will not support is the notion of exact equality, or perfect substitutability, because the mental magnitudes are noisy; they represent quantity plus or minus some percent uncertainty. Words from many classes refer to a range of discriminably different values – for example, the word ‘red’ refers to a range of different colors. However, a natural number word, such as ‘one’, refers to a single exact quantity and not to a range of quantities.

We are not claiming that words whose meanings are real values are impossible to learn (vide ‘pi’), or that integer words are mandatory in all human languages [16,17]. We do claim that integer-valued words occur commonly in languages and do so because children are disposed to entertain integer-valued hypotheses when learning the meanings of those words.

Constraints on number concepts

For us, it is crucial that potential symbols for numbers support arithmetic inference (Box 1). This means that to be a number concept, a symbol must denote an entity over which arithmetic operations and functions could operate. We do not require that an individual who possesses the symbol or concept in question be able to perform a given arithmetic operation or calculation over that concept and/or produce the correct word for the result. These additional requirements are too strong; performance or developmental or other constraints might prevent the carrying out of a given arithmetic inference. Because it is always possible to find an arithmetic operation that an individual cannot perform, this over-strong requirement would entail that no one possesses number concepts.

Exact equality

One use we make of integers is counting things. A fundamental intuition here is that if three things are counted, then the resulting cardinal value will be exactly equal to the cardinal value that will result from counting them again. It is hard to account for this intuition if the brain represents cardinal values by noisy reals because two exactly equal values will never occur. Although we find compelling the evidence for the existence of an analogue magnitude representation underlying counting and other number tasks [18], exact equality challenges such models. We are left without an account of why our basic number concepts – the ones picked out by language – should be integers rather than reals. Or why, in counting, each discrete value added to a magnitude should equal exactly 1, rather than some real number. Or why two counts of the same set must yield the exact same number. Learning the count words in a language, even in the presence of small sets of objects, will not help if the hypothesis space for possible (number) word meanings is the space of real values. Two real-valued measures of the same entity are, in general, infinitesimally likely to be exactly equal, and infinitesimally likely to have an integer value. Thus, the chance of a child entertaining an integer hypothesis would be infinitesimal. If no child would learn integer values, then no language would contain words for such
Box 1. Insights from the theory of measurement: the importance of symbol processing in constraining representational mappings

There are two aspects to a useful representational system [12,28]: (i) the mapping from the represented entities to the symbols that represent them (heavy arrows in Figure I); (ii) the combinatorial symbol processing operations into which the symbols enter (light arrows). They are strongly interdependent because the validity of the symbol processing depends on the properties of the mapping. If, for example, the mapping from any objective quantity, discrete or continuous, to mental symbols or their brain realization is logarithmic [30], then the process that implements multiplication must have the formal properties of addition; the addition of quantities in the logarithmic (subjective) domain corresponds to the multiplication of the corresponding quantities in the linear (objective) domain (Figure I).

To make this concrete, suppose that number maps logarithmically to the loci of activity in linear arrays of neurons [30]: the larger the number, the more the locus of activity is displaced toward the ‘far’ end of the array, but the displacement increases in proportion to the logarithm of the number mapped, not in proportion to the number. Multiplication in this representational system then becomes the process that maps from two loci of activity (representing the two numbers to be multiplied) to a third locus. The extent of the displacement of the third locus must be the sum of the displacements of the two multiplicands. It is not hard to contrive physical combinatorial processes with this property, as in the case of the slide rule. The problem comes when we try to imagine how addition and subtraction could be implemented. There is no mathematical or physical process that combines logarithmic quantities to produce a quantity that is the logarithm of the sum of their corresponding linear quantities, other than either converting to antilogarithms and adding the no-longer logarithmic symbols, or computing by table look-up. Another problem comes when we try to imagine how the system represents the outcome of operations that yield 0, such as subtracting 1 from 1. The problem is that the logarithm of 0 (−∞) is not physically realizable. There is also the question of how the system is to represent the directed (signed) quantities that naturally arise from subtraction (e.g. 1−2) because only positive quantities have logarithms.

Processing considerations must also be taken into account in evaluating proposals that small numbers (3 or 4, and smaller) have a fundamentally different representation from larger numbers. The processing of the symbols for small numbers must be able to generate the symbols for large numbers (e.g. 2 + 3 = 5) and vice versa (e.g. 9 − 7 = 2).

Next number

We also need a notion of next number. A mechanism for accumulating continuous magnitudes can supply a next magnitude, given some procedure to determine ‘effective’ or stochastic equality and therefore ordering [13]. However, there is more at stake in next number than simply discrete ordering and exactness. Counting is not simply a matter of identifying ‘some’ discrete value ‘minimally’ greater than a value ‘effectively equal’ to a given magnitude (say, the current magnitude in a count). ‘Some’ value will just not cut it. The next value in a count can be obtained only by adding the integer value 1. Accumulator magnitude accounts have to stipulate that the count value to be added is effectively equal to 1 [13]. However, this value is not only unobtainable with exactness, in such accounts it is also ad hoc. Why should the ‘unit’ magnitude in an accumulator count be ~1? Why could it not happen to be ~0.67, say, or ~1.134, or any other real value that would discretely order the magnitudes in the accumulator? Such values would yield the next stochastic magnitude, nicely ordered; but they would not yield the count values, which ‘happen’ to be integers. Furthermore, the exact value 1 is needed as the identity element in multiplication, where, again, no other value will do (Box 2).

Computational compatibility

As we have seen already, the arithmetic principle is useful because it imposes heavy constraints on potential number representations. A further constraint on number
Box 2. The multiplicative identity and natural psychological units

The scaling of the possible mappings from numbers to the mental magnitudes (brain signals) that represent them is constrained by the multiplicative identity (the number one). Suppose that the neurobiologist has access to the inputs to and outputs from the process that is thought to implement the multiplicative combination of the symbols for number. The symbol for one (one) can be inferred from the behavior of the neural multiplier alone, without reference to external world quantities. It is the unique number symbol that, when entered into the multiplier together with any other symbol, generates an output that is exactly the same as the other input (example 3 in Box 2, Figure I). Whenever two symbols for values larger than one are fed to the multiplier, the output of the multiplier will be greater than either of its inputs (example 1 in Box 2, Figure I). When one of the inputs is less than one, the output of the multiplier will be less than the other signal (example 2 in Box 2, Figure I). From results such as these, it is possible to determine completely the mapping from discrete quantity (integers) to corresponding brain signals without knowledge of the objective quantities to which they refer.

Suppose, further, that a neurobiologist can physically measure brain signals that are causally connected to a continuous quantity, such as duration. The mapping to any such continuous quantity is completely specified as soon as one has identified the objective duration to which the unit magnitude in the system of brain signals maps; the neurobiologist will have discovered mentsecs, the units of mental time. Similar remarks apply to the mental units of other physical quantities. In short, arithmetic processing strongly constrains the mappings from both discrete and continuous quantity to the mental symbols for them and to the neurobiological variables that physically realize those mental symbols.

(i) \[ x \times 1 = x \]

(ii) \[ x \times 0 = 0 \]

(iii) \[ x \times -1 = -x \]

Figure 1. One and the fixed properties of the multiplier.

Box 3. Counting and cardinality

Children gradually improve their counting skill throughout the preschool years [26]. An important question is how early in this process they understand that counting yields a word that represents the cardinal numerosity of the counted set. Piaget [31] argued that children did not comprehend cardinality until they understood that two sets had the same cardinality if, and only if, they could be placed in one-to-one correspondence. This view still has adherents [32]. LeCorre and Carey [1] argue that children do not understand the connection between counting and cardinality until they can reliably give the correct number of items in response to a request (e.g. ‘give me five’). However, in their experiments, they prevent counting and discard the subjects who do it, which would seem to render moot the bearing of their experimental results on the child’s understanding of the relationship between counting and cardinality.

Gelman and Gallistel [28] stressed the importance of arithmetic manipulation in the definition of cardinality. Counting yields a symbol that refers to that property of a set – its cardinality – to which the basic arithmetic operations of ordering, addition and subtraction validly apply. Gelman and Gallistel argued that children understand this very early, when they are still poor counters (Figure I). As evidence that young preschoolers understand the connection between cardinality and counting, Gelman and her co-workers cite: (i) the fact that they resort to counting when the cardinality of a set comes into question [27, 33]; (ii) when observing a puppet count (see earlier), they distinguish between departures from convention that do not affect the cardinal reference of the result and those that do [34–36].

A learning mechanism for the natural numbers

We argue that basic number representation in humans is not limited to the reals; it must include a representation of the natural numbers qua integers. The natural numbers are exact values (whose representations are not inherently noisy, vague or fuzzy), ordered by a well-defined notion of next number. Moreover, they are not simply any sequence of well-ordered exact values, such as 0.67, 1.34, 2.01…, they are precisely integer values. Young children access such values when they entertain hypotheses regarding the meaning of the count words (Box 3).

What would the required learning mechanism look like? What properties should a mechanism have in order that it will learn: meanings for count words that designate integer values (as opposed to discrete reals, stochastic functions over reals, vague values, etc.); to order integer values according to a next relation; to support arithmetic reasoning representations is computational compatibility. Suppose there is a computer that represents the numbers 0 to 3 by bit patterns (00, 01, 10 and 11) and represents larger numbers by voltage levels (an analog representation). How could such a device determine that \( 7 - 5 = 2 \) (a difference between noisy voltages somehow becomes the bit pattern 10)? How could it compute \( 5 + 2 \) (sum a voltage and a bit pattern)? It is possible to add and subtract voltages or to add and subtract bit patterns but it is not possible to subtract a bit pattern from a voltage. Bit patterns and voltages are computationally incompatible.

Humans and non-human animals compute rates and proportions [13,19–23]. The computation of a temporal rate of occurrence requires dividing a symbol that represents a discrete quantity (number of observed events) by a symbol that represents a continuous quantity (duration of observation) to obtain a symbol that represents a different continuous quantity (rate). Therefore, we reject proposals that make the symbols for discrete and continuous quantity computationally incompatible. A fortiori, we reject proposals that make the mental symbols for small and large discrete quantities computationally incompatible [1].
(Boxes 1, 2); and to enable related magnitude estimation judgments? What properties should such a learning mechanism have in order that it will complete its task within the finite learning-trial opportunities available to real learners? What is the minimal structure that such a mechanism could have?

We distinguish between a generative symbol system and a realized system of symbols. A generative rule specifies the derivation of an infinite set of symbols – for example, the numbers. A realized system consists of symbols that have already been produced by running a symbol derivation and storing the result in memory. A system of arithmetic reasoning and one are mutually supportive but not because realized symbols for infinitely many numbers pre-exist in memory. What pre-exists is a procedure for generating them.

The well-known accumulator, in which mental magnitudes are generated by the summation of successive magnitudes of approximately unit size, is an example of such a generative mechanism; it realizes mental magnitudes to represent real-valued numerosities, as needed. We now postulate, first, an integer symbol generator that generates mental symbols to represent discrete quantities (that is, numerosities). Second, its symbols are ordered by the ‘next’ relation. Third, the symbols are appropriately linked to positions in the system of mental magnitude, much as the numbers on a speedometer are positioned on the underlying analogue representation (of speed). Fourth, the substitution operations that implement the concept of exact equality (perfect substitutability) are defined with respect to these symbols. Fifth, this symbol system supports algebraic reasoning about quantity (reasoning in which the quantities involved need not be specified) (Figure 1).

A minimal learning mechanism

The five requirements above can be met by the following assumptions:

(i) There is at least one innately given symbol with an integer value, namely, ONE = 1.
(ii) There is an innately given recursive rule \( S(x) = x + \text{ONE} \). The rule \( S \) is also known as the successor function [24].
(iii) Each realized integer symbol is given a corresponding accumulator value. As with all such values, it is a noisy real. The difference between the accumulator values assigned to any two successive integer symbols is always approximately equal to the accumulator value assigned to ONE itself. [This calibrates the system of integer symbols to the system of magnitude symbols (Figure 1).]
(iv) Inference mechanisms operating on these symbols support unbounded substitution. For any such symbol, \( N, \text{N}^*\text{ONE} = \text{N}^*\text{ONE} \).

The integer symbols could be sets of mental hash marks, a system that has been proposed for the representation of small numerosities [1]. For example, the symbol | represents 1, || represents 2, ||| represents 3. However, this type of notation has a property that severely limits its usefulness. As the \( n \) to be represented grows in size, the ‘physical size’ or length of the symbols themselves grows linearly with \( n \). It is as if the word for elephant had to be thousands of times bigger than the word for bacteria – not a welcome property. The accumulator magnitude representation has this same unwelcome property. In both cases the problem is, how can an unbounded, or even a large bounded, set be represented? This suggests that in addition to the grid there should be a compact notation, in which the physical size of the symbols grows as the logarithm of the quantities represented, as it does in place-value number notation systems [25].

To provide a compact notation, an unbounded set of symbols must be generated, with a one-to-one correspondence between symbols and integer values, so that each symbol functions as a unique identifier for some unique integer value. Each symbol is linked to a unique position on the mental number line, the ordinal position of which determines the meaning of the symbol (i.e. its specific integer value). Finally, these compact symbols can compose in centrally constructed and centrally processed thoughts.

Whence the compact integer notation?

Minimally, the concept of, and symbol for, 1 must be innately realized because the recursive rule \( S \) that generates the integer symbols requires that concept and symbol. The \( S \) also requires that the addition operation +, the identity relation =, algebraic variables and a recursive capacity (minimal algebraic notions) also be innately realized.

Many variants on this proposal are possible; for example, variant 1: the integer notation also includes an innately realized symbol TWO (= 2); variant 2: the integer notation also includes an innately realized symbol for THREE (= 3); and so on. However, given that the set of natural numbers is unbounded, not all of them can be represented by realized symbols, and thus not all can be innately realized. Nonetheless, the entire set can be represented in the sense that it is generated recursively by \( S \). One way to think about \( S \) is that it generates the meanings for the infinite set of integers, using finite means. Because the means for generating the set are finite (namely \( S \) \( S \) can be innately realized. However, for realizing an unbounded, or even a large bounded, set of symbols, where each symbol uniquely carries an individual integer meaning, a notation is required whose symbol length does not grow monotonically with the magnitude of the value represented (as does a grid or an accumulator representation). Notably, the count word list in a natural language such as English has a notational system for integers with just this property; for example, English uses just two words to represent 1 000 000.

Conceivably, the brain might have an innate compact integer notation – for example, produced recursively or by a cascading notation for orders of magnitude [25]. Alternatively, the notation for values larger than ONE (>TWO, >THREE...) might simply co-opt natural language itself and the brain acquire that compact notational system. In this case, the detachable unique identifiers designating integer values larger than 1 (>2, >3...) will be drawn from a learned notation – namely, the natural language expressions of the learner’s first language. Under this proposal, natural language provides a (compact) notation
for prior integer concepts. The integer concept space must be available to supply the hypotheses required to learn what this lexical subsystem encodes. The calibrated and ordered grid and the specific rung on the grid are realized internally by running the recursive function \( S \). The ordinal position of a given rung in the grid fixes the meaning of the unique identifier bound to it. Whereas the meanings themselves are not learned from input, the notation for their unique identifiers is.

A related proposal assumes some innate algebraic principles that mediate or govern reasoning about discrete numerosities [26]. This system, too, is distinct from the accumulator system and has symbols that do not enter into arithmetic operations that determine numerical values. Unlike the mental magnitudes, they are not used for arithmetic computation, thus circumventing the problem of computational compatibility. Rather, they are used to draw conclusions about the outcomes of computations, by licensing symbolic substitutions. Three- and four-year-old children have little difficulty switching between an approximate and an exact system, and prefer the latter when the task is an arithmetic one [27,28].

Summary
The basis of our natural number concepts is hypothesized to be the innate representation \( S \) that recursively defines the positive integers and the concept next number. The basis of these concepts cannot be a system of continuous magnitude representation, accumulator or connectionist, noisy or not, without a system that can represent exactly the value 1. Moreover, the integer representation becomes calibrated to accumulator magnitudes, enabling integer calculation and magnitude estimation. The brain might generate its own compact code for representing integer values and then learn the appropriate mapping from that internal compact code to the corresponding compact code in natural language. Alternatively, the brain might simply co-opt the compact code of a natural language. This latter account would afford an important role to language learning without embracing the Whorfian claim that natural language teaches de novo the meanings of integer concepts. These meanings are known in an important sense, not by learning but innately: namely, as generated by \( S \). It is hard to make the case that \( S \) can be empirically induced from experience [6]. Instead, humans possess an inbuilt learning mechanism in the form of the successor function that employs a little piece of algebra [8]. This account raises new outstanding questions that have yet to be resolved (see Box 4).

Acknowledgements
The work reported herein was in part supported by grant NSF BCS-0725169 to A.M.L.

References
30 Dehaene, S. (2001) Subtracting pigeons: logarithmic or linear? Psychol. Sci. 12, 244–246

33 Gelman, R. (1972) Logical capacity of very young children: number invariance rules. Child Dev. 43, 75–90