PERSPECTIVES: NEUROSCIENCE

Knowledge of Number: Its Evolution and Ontogeny

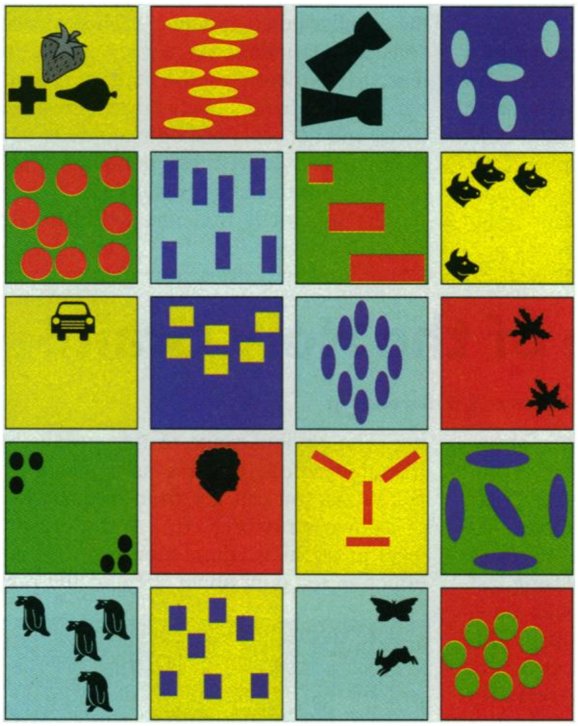
Susan Carey

Mathematical concepts and systems of notation are intertwined with human cultural history. The discovery (or invention) of 0, understanding of negative, rational, and real numbers, and development of the calculus took place over thousands of years. But what of those most fundamental of mathematical objects, the positive integers: 1, 2, 3, 4, 5, …? Until recently, the consensus was that the capacity to represent the positive integers was also the product of culture, dependent on the uniquely human capacity for language. A new study on the represention of number in animals, on page 746 of this issue (*1*), along with studies of the representation of number by prelinguistic infants (*2*), are undermining that consensus.

The consensus was based on two considerations. First, not all human languages contain a list like "one, two, three…" Such a representational system developed slowly, through several stages, as documented by historical language studies (*3*). Second, even in highly numerate cultures, such as ours, children require a two year apprenticeship to master their language's integer list (*4*).

Nevertheless, it is now clear that animals on many branches of the evolutionary tree have the capacity to represent number (*5*). To establish that animals have this capacity, one must show first that they distinguish between sets of individuals on the basis of true numerical differences rather than, say, total volume of the indviduals or spatial density of the scene. Second, one must show that these distinctions between sets carry numerical meaning for the animal. Minimally, the animal should represent that 1 is less than 2, and 2 is less than 3, and so on. It is not enough to know that 1 is different from 2 and 2 is different from 3, and so on.

For those not already convinced that nonhuman animals genuinely represent number, the new data presented by Brannon and Terrace elegantly demonstrate both that rhesus monkeys distinguish between sets of individuals on the basis of number and that they can represent all the ordinal relations that exist among the numbers 1 to 9. Brannon and Terrace build on Terrace's previous demonstrations of rhesus representation of serial order. For example, presented with pictures of a cup, a table, a car, and a flower on a TV touch screen, rhesus monkeys can learn to touch them in any arbitrary order, irrespective of their position on the screen.



How many? Examples of stimuli used to probe numerical representation in rhesus monkeys (*1*).

The initial experiment of the new work begins by showing that rhesus monkeys can learn to order sets of stimuli that consist of 1 element, 2 elements, 3 elements, and 4 elements (see examples of each numerosity in figure). Plainly, each set of stimuli forms an arbitrary list of four items, so how does this experiment go beyond Terrace's previous demonstrations of rhesus learning to order arbitrary lists? In the current report, the question was whether rhesus monkeys would learn the rule, "touch 1, then 2, then 3, then 4," when each list has objects of varying sizes, shapes, and colors, such that no other stimulus dimension covaries with number. The results showed that, yes, they would. Performance on new lists, each seen only once, was every bit as good as on ones at the end of the training series, each seen 60 times.

But which rule did the monkeys learn? "Touch 1, then 2, then 3, then 4" or "touch the items in order of increasing numerosity"? To answer this, Brannon and Terrace gave the monkeys new problems with sets of stimuli having between 1 and 9 elements. They found that the monkeys generalized the rule "touch the lower number of items before the higher one" when presented with stimuli containing set sizes that they had not been taught (for example, 5 versus 7). This demonstrates that the monkeys were capable not just of the representation of numerical order, but also of the abstraction of a numerical rule. More important, it shows that the representations of number underlying success in the first study are not restricted to the subitizing range (1 through 4, the range in which humans can simply look at a set of objects and know how many there are without explicitly counting).

Because it would be very surprising if the abilities that Brannon and Terrace demonstrate for rhesus monkeys were entirely absent in humans, these results challenge the consensus that the capacity to represent numerals is a cultural construction. But to establish whether these representations are the evolutionary source of the human number capacity, we must ask whether they are its ontogenetic source. That is, are these abilities available to infants before they acquire language, and are they the foundation of the culturally constructed integer list representations?

Although the literature on numerical representation by prelinguistic infants contains nothing so impressive as the results of Brannon and Terrace, it is clear that young infants (4.5 to 8 months of age) have some numerical competence. For example, they discriminate between stimuli consisting of 1, 2, and 3 elements, and even expect that 1 item added to an array consisting of 1 hidden item will yield an array of exactly 2 items (that is, $1 + 1 = 2$), as well as that $2 - 1 = 1$, $2 + 1 = 3$, and $3 - 1 = 2$ (*2*). There is as yet no proof, however, that infants discriminate among larger sets of elements on the basis of number, nor have there been stringent controls for possible nonnumeri-

CREDIT: E. BRANNON

The author is in the Department of Psychology, New York University, New York, NY 10003–6634, USA. E-mail: sc50@is6.nyu.edu

cal bases of discrimination, such as total stimulus area. And finally, there has been no convincing demonstration that infants represent the ordinal relations among sets of 1, 2, and 3 elements.

How might nonhuman animals and prelinguistic infants represent number? Two classes of models for nonlinguistic numerical representational systems have received empirical support: object file models and analog magnitude models. In the object file model, the infant or monkey forms a representation with one symbol for each individual in the set and compares representations by computing one-to-one correspondences between sets. Such representations are limited to the number of individuals that can be held in short-term memory at any one time, which is 3 or 4. These representations contain no symbols that function as numerals, and there is no counting process. In analog magnitude models, number is represented by a continuous quantity, akin to a number line. Representations are compared by the same sorts of operations that compare lengths, durations, volumes, and other representations of continuous quantities. The process by which the analog magnitude is incremented for each item in the set is equivalent to counting (*6*), but analog magnitude models differ in many ways from integer list models (*7*).

Brannon and Terrace's data favor an analog magnitude model. Their monkeys represent numbers that exceed the limits of the object file model. Further, analog models correctly predict that number comparisons become easier when the differences between the numbers are greater (the distance effect). By contrast, for infants the evidence favors the conclusion that the object file model underlies the prelinguistic numerical representations in the events studied to date (*7*). There is also considerable indirect evidence that the integer list symbolic representation of number is built from object file representations, and not from analog magnitude representations (*4*), even though human adults certainly use the latter as well (*5*).

The upshot is that one evolutionary source of human number representation—the analog magnitude representations that Brannon and Terrace most probably are tapping in primates—is not the primary ontogenetic source of human symbolic number list representations, either in linguistic evolution or in individual development. Although this conclusion is controversial, our challenge is clear. We must specify the nature of nonlinguistic representations of number (there may be many) and characterize the process by which explicit symbolic representations are constructed in the history of each culture and again by each child.

References

1. E. M. Brannon and H. S. Terrace, *Science* **282**, 746 (1998).
2. K. Wynn, *Math. Cognit.* **1**, 35 (1995).
3. ———, *Cogn. Psychol.* **24**, 220 (1992).
4. J. R. Hurford, *Language and Number* (Oxford Univ. Press, Oxford, 1987).
5. S. Dehaene, *The Number Sense* (Oxford Univ. Press, Oxford, 1997).
6. C. R. Gallistel, *The Organization of Learning* (MIT Press, Cambridge, MA, 1990).
7. C. Uller, S. Carey, G. Huntley-Fenner, L. Klatt, *Cogn. Devel.*, in press.

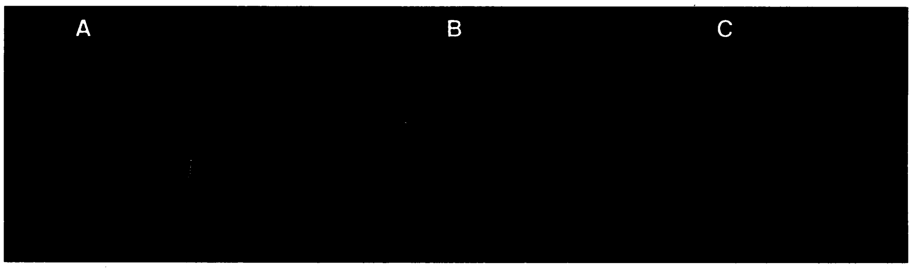
PERSPECTIVES: PROTEIN FOLDING

A Glimpse of the Holy Grail?

Herman J. C. Berendsen

The prediction of the native conformation of a protein of known amino acid sequence is one of the great open questions in molecular biology and one of the most demanding challenges in the new field of bioinformatics. Using fast programs and lots of supercomputer time, Duan and Kollman (*1*) report on page 740 of this issue that they have successfully folded a reasonably sized (36-residue) protein fragment by molecular dynamics simulation into a structure that resembles the native state. At last it seems that the folding of a protein by detailed computer simulation is not as impossible as most workers in the field believe. Or is this an overoptimistic view?

With the number of known gene sequences increasing at an accelerating pace (the complete genomes of 13 bacteria and of yeast are now known, the first multicellular animal will follow soon, three plants and the fruit fly are in the pipeline, and the human genome sequence can be expected at the beginning of the next century), the quest for the structure and function of the coded proteins becomes pressing. The obvious route to that goal is by homology modeling: use as much information as you can get from the database of known structures. But at the present level of sophistication, such methods are effective for only about 25% of the proteins for which the amino acid sequence is known; if sequence homology drops below 25%, the reliability of database-oriented methods drops to nearly zero.



Unfolding is easier than folding. Four snapshots from the simulation of an unfolding protein called HPr (a phosphate-transferring protein): **(A)** native conformation, **(B)** partly unfolded conformations that still contain most of the secondary structure, and **(C)** an unfolded (or randomly folded) structure.

CREDIT: K. ANTON FEENSTRA, GRONINGEN

Still, most small proteins fold spontaneously in seconds into their native conformations; secondary structure elements like α helices or β turns fold in tens of nanoseconds to microseconds. Such folding is thermodynamically downhill and is just a result of the physical interactions between atoms, including those of the solvent. These atomic interactions are elementary and well-known, so why can't we use this knowledge to mimic the native folding process? Well, for two reasons: First, existing computers cannot sample enough configurations in a reasonable time to come up with the thermodynamically stable native structure; second, we are not too sure that the available force field descriptions, which we need to compute the energy of each configuration, are accurate enough to come up with a reliable free energy of a conformation. The trouble resides in the enormously large positive and negative contributions that nearly cancel in the computation of the total energy.

The sampling problem can be summarized as Levinthal's paradox: If we assume three possible states for every flexible dihedral angle in the backbone of a 100-residue protein, the number of possible backbone configurations is 3^{200}. Even an

---

The author is in the Department of Biophysical Chemistry and the BIOSON Research Institute at the University of Groningen, Netherlands. E-mail: berendsen@chem.rug.nl