

Numbers and brains

C. R. Gallistel¹

Published online: 14 April 2017
© Psychonomic Society, Inc. 2017

Summary The representation of discrete and continuous quantities appears to be ancient and pervasive in animal brains. Because numbers are the natural carriers of these representations, we may discover that in brains, it's numbers all the way down.

Keywords Memory · Numerical processing · Navigation · Classical conditioning · Matching

Cuttlefish assess prey number and prey size before deciding where and when to attack (Yang & Chiao, 2016). Bees have cognitive maps on which they represent distances and directions (Degen et al., 2016). Mice represent the number of lever presses they have made and the duration of an elapsing interval, and they also represent their uncertainty about those numbers and those durations (Berkay, Cavdaroglu, & Balci, 2016; Kheifets, Freestone, & Gallistel *in press*). The capacity of animal brains to represent in memory such abstract quantities as number, distance, direction, duration, rate, and probability appears to be evolutionarily ancient and extremely widespread. This poses the question, What carries these representations? How are they coded in neural tissue? That is, how are they stored in memory? The purpose of learning is to put information in memory so as to make it retrievable for later use in behavior.

Most treatments of memory are domain specific: spatial memory, associative memory, emotional memory, declarative

Target publication Yang, T. I., & Chiao, C.-C. (2016). Number sense and state-dependent valuation in cuttlefish. *Proceedings of the Royal Society: Series B*, 283, 20161379. doi:10.1098/rspb.2016.1379

✉ C. R. Gallistel
galliste@rucss.rutgers.edu

¹ Rutgers Center for Cognitive Science, Rutgers University, 158 Frelinghuysen Road, Piscataway, NJ 08854, USA

memory, procedural memory, episodic memory, and so on. To be sure, the processes that extract information from experience and those that use it later are domain specific, because predicting what will happen next is not the same computational problem as computing the range and bearing of another location from one's current location. Both computations presuppose memory, because both rely on information acquired piecemeal from experiences distributed over long stretches of the past. But building a cognitive map from scraps in memory is a different enterprise from building a good model of what events predict what other events at what latencies, and using a stochastic model to anticipate the next event requires different computations than those required when using a cognitive map for route finding.

Domain-specific treatments of memory focus on how memories are created and how they are utilized, not on memory itself. Memory is the medium through which the past communicates with the future. Its function is to convey information extracted from past experience forward in time in a manner that makes this information accessible to computation whenever it is needed to inform behavior.

A key insight at the foundation of information theory and modern information technology is this: When it comes to information conveyance, what the information is about is irrelevant. In the article that laid the foundations of information theory, Shannon (1948, p. 379, emphasis in original) wrote:

Frequently the messages [conveyed] have *meaning*; that is they refer to or are correlated according to some system with certain physical or conceptual entities. These semantic aspects of communication are irrelevant to the engineering problem. The significant aspect is that the actual message is one *selected from a set* of possible messages. The system must be designed to operate for each possible selection, not just the one which will actually be chosen since this is unknown at the time of design.

It is of course possible that the brain uses many different memory media. However, the only biologically realized memory media we know—DNA and RNA—are universal. Polynucleotide strings convey heritable information from generation to generation in every living thing—even in viruses. The information conveyed directs the formation of wildly different organic forms on wildly different time scales.

The information conveyed by polynucleotide strings is often abstract. The *Pax6* gene, for example, is critical for eye development in vertebrates. A close homologue triggers eye development in insects. When the *Pax6* gene from a mouse is expressed in a fruit fly, however, it triggers the development of the faceted insect eye, not the lensed vertebrate eye. Thus, the message carried by *Pax6* roughly translates as “build an eye here”; *Pax6* does not carry information about how to build an eye nor, a fortiori, the form the eye should take. The machinery coded for by other parts of the fruit fly genome responds to the command issued by a reading of the mouse *Pax6* by building an eye appropriate to the organism in which the eye command has been read.

Another way of phrasing the lesson the *Pax6* gene teaches is that a conveyed message does not acquire semantics until it is read. The semantics arise from the interpretation that the reading machinery—the domain-specific arrangement of computational operations—imposes. This lesson is everywhere in computer science. If I write out for you a bit pattern in a register I observed in my laptop’s memory, you have no way of knowing what it means. It may encode a name, the value of a pixel in a photograph, an instruction in a program, the distance from one city to another, the address of another bit pattern elsewhere in the computer’s memory—whatever. Its meaning arises from the codes by which the world and the computer itself write to its memory, and from the computational operations into which the bit pattern enters when it is read. Shannon’s point may be rephrased as: When it comes to conveying information—that is, when it comes to memory—it’s all bits. Because the meaning does not matter, the only question is how to convey bits as reliably and enduringly as possible, while using as little energy and as little material, as possible.

A second striking fact about both computer and genetic memory is that they have the same bipartite architecture. A register in RAM has two parts, a coding portion and an address portion. Memories are retrieved (read) by a probe signal that specifies the location by “binding” to the address portion of the register. Genes also have two parts, a coding portion and a promoter. Transcription factors are the address probes: They identify the location of the information in the coding portion by binding to the promoter portion, thereby initiating the reading of the coding portion. This architectural similarity is either an astonishing coincidence or a hint that there is only one good way to build a memory system.

Bit patterns are basically numbers, because “numbers” is our word for the symbols by which we represent quantities, both discrete—like numerosity—and continuous—like distance and

duration. Computing machines were originally conceived as number processors. The two’s-complement fixed-point number code was chosen for the felicity with which it represents a wide range of signed numbers and the economy with which it may be processed. The computational hardware was optimized for performance of the basic operations of logic and arithmetic. Those choices have not changed. Oddly enough, machines so conceived and so optimized have proved to be able to represent everything else. There is nothing we know how to represent that we cannot represent in a computer, usually in less space and with much faster processing than in any other way. The reason for that is that everything we know how to represent we can represent with numbers. We may someday find a way to represent the ineffable somethings we deeply believe cannot be represented by numbers, but that day is not here yet.

The conclusion I am tempted to draw from these reflections is that there is a sense in which, when it comes to memory, the Pythagoreans were right: It’s numbers all the way down. First and foremost, the numbers in memory represent such behaviorally important abstract quantities as numerosity, duration, distance, location, rate, probability, size, and so on. That is what shaped their evolution. But these numbers can also be used as names, as on athletic jerseys. Another way to think about the *Pax6* gene is that it encodes the name of the program that builds an eye. The name for the eye-building program has not changed since the Cambrian, but the programs for building eyes have diverged spectacularly.

At least as odd, to my mind, is the recent discovery that the names for the small numbers—the number words from “two” to “five”—are the most slowly changing words in the world’s major language families. The words for “two” in most branches of the huge Indo-European family tree are recognizable cognates. The half life of the low number words is estimated to be on the order of 80,000 years, whereas the half lives of the words for such common things as birds are orders of magnitude shorter. Even the pronouns have shorter half lives than the low number words.

References

- Berkay, D., Cavdaroglu, B., & Balci, F. (2016). Probabilistic numerical discrimination in mice. *Animal Cognition*, *19*, 251–265.
- Degen, J., Kirbach, A., Reiter, L., Lehmann, K., Norton, P., Storms, M., & Menzel, R. (2016). Honeybees learn landscape features during exploratory orientation flights. *Current Biology*, *26*, 2800–2804. doi:10.1016/j.cub.2016.08.013
- Kheifets, A., Freestone, D., & Gallistel, C. R. (in press). Theoretical implications of quantitative properties of interval timing and probability estimation in mouse and rat. *Journal of the Experimental Analysis of Behavior*.
- Shannon, C. E. (1948). A mathematical theory of communication. *Bell Systems Technical Journal*, *27*(379–423), 623–656. doi:10.1002/j.1538-7305.1948.tb01338.x
- Yang, T. I., & Chiao, C.-C. (2016). Number sense and state-dependent valuation in cuttlefish. *Proceedings of the Royal Society: Series B*, *283*, 20161379. doi:10.1098/rspb.2016.1379