11-22-2022

Are Numerical Symbols Fundamental to Neural Computation?

Mirinda James
University of Missouri-St. Louis, mirindajames@gmail.com

Are Numerical Symbols Fundamental to Neural Computation?

Mirinda L. James

B.A. Philosophy, Iowa State University, 2017

A Thesis Submitted to The Graduate School at the University of Missouri-St. Louis
in partial fulfillment of the requirements for the degree
Master of Science in Philosophy

December
2022

<u>Advisory Committee</u>

Gualtiero Piccinini, Ph.D.
Chairperson

Lauren Olin, Ph.D.

Carrie Figdor, Ph.D.

Abstract: Neuroclassicism is the view that cognition is computation and that core mental processes, such as perception, memory, and reasoning are products of digital computations realized in neural tissue. Cognitive psychologist C. R. Gallistel uses this classical framework to argue that all cognitive information processing is based on symbolic operations performed over quantitative values (i.e. numbers) stored in the brain, much like a digital computer. Assuming this hypothesis, he investigates how the brain stores quantitative information (i.e. the numerical symbols involved in neural computation). He claims that it is more plausible that memories for numbers are stored within molecular mechanisms inside the neuron, rather than within specific patterns of cell connectivity (the substrate for memory storage assumed by the traditional Hebbian plastic synapse model). In this paper, I dissect and critique Gallistel's argument, which I find to be undermined by the findings of contemporary neuroscience.

1. Introduction

Identifying the neural basis of memory remains a central project of contemporary neuroscience. In his 2017 paper titled *The Coding Question*, cognitive scientist C. R. Gallistel investigates the relationship between memory storage and the neural code. He presents us with an experimental study implying that Purkinje neurons (neurons located inside the cerebellum) are able to store specific temporal durations between the onset of two stimuli (i.e. a memory for a specific amount of time). Based on the results of this study, he poses the following questions:

1. How are numbers, that is any sort of numerical information (e.g. "durations, distances, rates, probabilities, amounts, etc.") transmitted in the brain?
2. How are numbers (same information as above) stored in the brain?

He writes, "This raises the general question - how is quantitative information (durations, distances, rates, probabilities, amounts, etc.) transmitted by spike trains and encoded into engrams?" (Gallistel, 2017, p.1).

As a cognitive scientist, Gallistel investigates the physical mechanisms and information processing systems underlying learning and memory in the brain. He approaches these questions from a classicist theory of mind, a version of the computational theory of mind, the origins of which date back to the middle of the 20th century, around the dawn of the modern digital computer. Classicism states that cognition is computation and that cognitive processes, including memory, are a kind of digital computational processes. The mind is literally - not metaphorically - a computing system, and core mental processes (e.g. perception, memory, reasoning, decision-making, problem solving) are products of digital computations realized in neural tissue.

Gallistel's paper addresses the problem of neural coding, which involves understanding how information is represented, stored, and processed as it flows through the nervous system (Johnson, 2000). One of his central beliefs is that numbers are fundamental to how the brain (and any powerful and efficient computational system for that matter) processes information, and that this is a truism that stems from the logic of computation itself (Gallistel & King, 2010; Gallistel, 2017b, Gallistel, 2017c). He vehemently argues that numerical representations of both internal and distal stimuli exist in the brain and are fundamental to computational processing underlying cognition. Moreover, he believes *all* mental phenomena and human behaviors derive from computational operations over numbers encoded and stored in neural tissue.

For this reason, he is especially concerned with understanding the neural code for numbers, and poses the two preceding questions in his paper: how are numbers transmitted and stored in the brain? Most researchers believe that information is transmitted in the brain by the firing rates of action potentials and stored in the patterns of connections between neurons (i.e. *plastic synapse theory*, or *Hebbian Learning*). Although scientists have learned a great deal about the neural code, the precise way in which information is represented and categorized in the brain is still a mystery.

In the paper I am presently critiquing, Gallistel claims that it is more conceivable to implement a neural code based on numbers at the cell-intrinsic molecular level (that is, with molecules inside the cell) than at the level of the synapse (i.e. Hebbian learning). Based on purely theoretical considerations, he argues against the viability of the plastic synapse model for memory storage. In its place, he calls for an investigation into potential intracellular molecular mechanisms of memory, stating that it is easier to conceptualize storing numerical values using molecules inside the cell than at the synaptic junction (i.e. plastic synapses).

Moreover, he writes that he has thoroughly investigated the possibility of storing a number at a synapse, even asking several neuroscientists concerning their thoughts on how a number may be stored at a synapse, and has received only inadequate responses, or no response at all. Thus, he concludes that we must do away with plastic synapse theory altogether as a candidate for the memory storage mechanism, since it could never serve as a numerical storage unit in his classicist cognitive architecture. This line of reasoning in my view is patently unsound.

On the contrary, while it is true that the brain must store quantitative information, it is misleading to say that the brain must store ‘numbers’, and that numerical information is fundamentally the only type of information the brain processes and stores. While quantitative information is indeed instrumental in some cognitive reasoning processes (e.g. navigation), I disagree with Gallistel's statement that all information processing in the brain is based on numbers, in other words, computational operations performed on numerical representations, similar to how computation works in the digital computer. Quantitative information doesn't necessarily have to be stored in the brain using numerical or digital variables (not to mention all types of mental representations as Gallistel insists are numerically based), but rather, can be stored in the brain using whichever variables will work.

In the following pages, I will present Gallistel's argument against the plastic synapse model of memory. I will begin with a review of his sophisticated defense of the classicist theory of mind and explain why he believes the brain processes information in the same manner as a digital computer (computational operations over numerical bit patterns realized in neural tissue). Then I will recapitulate his three-fold attack on plastic synapse theory: 1) it is impossible to store a number at the location of the synapse, 2) synapses don't satisfy important information-theoretic criteria for a powerful and highly effective memory storage devices, 3) it is easier to formulate a neural coding hypothesis by considering a cell-intrinsic molecular mechanism for memory, rather than plastic synapse model. Lastly, I will raise objections to each of these points, and in doing so support the plastic synapse model of memory.

2. **Gallistel's approach to "*The Coding Question*"**

In *The Coding Question*, Gallistel writes that researchers haven't yet identified the molecular mechanism inside the Purkinje cell responsible for encoding the number representing the temporal duration. Nor do they know which specific physical change occurs inside the cell which stores the information for the temporal duration, or in which part of the cell the storage mechanism resides. However, it is apparent that some encoding occurs and based on this knowledge, we must "think seriously about the coding question in neuroscience" (Gallistel, 2017, p. 1).

Numbers all the way down?

To evaluate Gallistel's argument, we must first understand why he believes numbers are fundamental to the neural code. In the section of his paper titled "*Why Number Coding is Fundamental*", he explains that his conviction that numbers are central to neural computation is inspired by Claude Shannon's *Mathematical Theory of Communication* (1948). Shannon was concerned with understanding communication and established the mathematical foundations of communication theory, a set of laws explaining the dynamics of information processing systems. His argument that the plastic synapse is not a good model for memory storage in the brain is based purely on theoretical considerations, not on empirical evidence from neuroscience.

I will paste here two passages from his paper so you may follow his argument. Here is an excerpt taken from Shannon's original 1948 paper, which Gallistel references in his paper. Shannon writes:

> 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 (Gallistel, 2017, p. 4).

The following excerpt from Gallistel's paper explains his interpretation of Shannon's previous passage, and its influence on his theoretical approach to the neural code. He writes:

> What Shannon understood was that all information may be treated as numerical, irrespective of what numbers refer to. This insight simplifies and focuses the coding question in neuroscience – how does neural tissue transmit and preserve numbers? The number-coding question is equally foundational when the brain is approached from a computational perspective. At the foundation of a computing machine is the representation of numbers (Gallistel, 2017, p.4).

This graphic designed by Shannon in his 1948 paper illustrates the basic flow of information in a general communication system (Gallistel and King, 2009, p. 2; Shannon, 1948, p. 2). Gallistel uses this schematic as a framework for his theoretic approach to both information processing and representation in the brain.

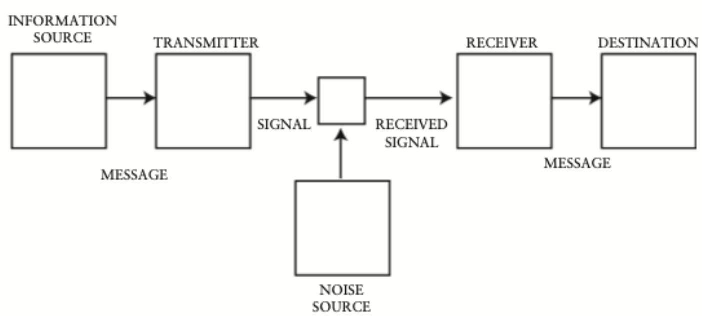

Figure 1.1 Shannon's schematization of communication (Shannon, 1948).

The flow of information begins with a source, which conveys a message encoded in a signal by a transmitter. This signal is received by some receiving apparatus, which decodes the message contained therein. The transmitter is responsible for converting the messages into a signal, that is, fluctuations of some physical medium which travels from the source to the receiver. The receiver must be sensitive to the fluctuations of this physical medium for it to understand the message.

The rules governing the conversion of the message to a signal is the coding language (or *code*). The code serves as representation for the messages coming from the source. Lastly, the mechanism in the transmitter which implements the conversion is the *encoder*, and the mechanism in the receiver which reconstitutes the message is known as the *decoder*.

So, in Gallistel's example of conveying quantitative information to the brain, the source would be some distal stimulus in the external world associated with some quantitative concept (say, distance from some object), the message would be that quantitative value (say, 10 meters), the transmitter is the neural mechanism which receives the message from the stimulus and produces a signal representing that particular quantitative concept (according to the coding rules innately designed in the system), the receiver receives the signal and reconstitutes the original message (10 meters away), and lastly, the destination is the part of the brain which uses the quantitative information in some cognitive or behavioral task (walking towards the object).

Gallistel states that another key insight from Shannon information theory to bear in mind when approaching the neural coding problem is that we must assume there is a finite set of possible messages the receiver could receive. If not, the set of possible messages would be infinite, and the messaging system would be useless. "The most important aspect of Shannon's analysis for our understanding of how meanings arise in brains is the starting point of his analysis: the assumption that there is a set of possible messages" (Gallistel, 2019, p. 392).

When applying this concept to the flow of information from the world to the brain, this naturally leads us to question: 'What determines the set of possible messages in the brain?' (Gallistel, 2019). Gallistel states that the brain stipulates the set of possible messages, since the external world is indifferent to the brain's information processing system. Thus, meaning in the brain is innately constructed, and arises when it creates its own apparatus for receiving and processing external information. He writes that these representational apparatuses are genetically determined and produced during the early years of brain development.

Lastly, another lesson gleaned from Shannon's information theory is that when it comes to conveying a message, the semantic content of the message is irrelevant to the overall engineering of the messaging system (Gallistel, 2017c, p. 3; Gallistel, 2017, p. 4). The message being conveyed only bears meaning when it has been received by the receiver and interpreted according to whatever domain-specific set of rules it has established to process that specific type of information (Gallistel, 2017b). When it comes to conveying information, "it's all bits" (Gallistel, 2017b).

Since the semantic meaning carried in the message is irrelevant until the receiver has received it, the main goal is to understand how to convey bits in the most efficient and reliable manner possible. "Bit patterns are basically numbers, because "numbers" is our word for symbols by which we represent quantities, both discrete – like numerosity – and continuous – like distance and duration (Gallistel, 2017b). Moreover, he notes, computing machines were originally thought of as number processors. These discrete and continuous quantities we experience are naturally represented by numbers because numbers serve as symbols for quantities - any number can refer to any quantity (Gallistel, 2017, p. 3). Thus, it is possible to convey any quantity simply by transmitting a number in a bit pattern.

Numerical abstract symbols

Box 1 below, titled *Numbers in the Brain*, is extracted from Gallistel's 2017 paper and contains a list of numbers in the brain – that is, numerical abstract symbols – Gallistel believes to be central to neural computation and involved in the information processing underlying all psychological phenomena (Gallistel, 2017). These numbers include: amounts, durations, distances, probabilities, rates, navigational vectors, etc. Gallistel believes that these numbers are the vehicles of computation, and that all human behaviors and mental states are the product of computations performed over these numerical symbols realized in the brain (Gallistel & King, 2009).

Box 1. Numbers in the Brain

In the brain the symbols for quantities must represent:

(i) Discrete quantities (numerosities [55–58]).
(ii) Continuous quantities (e.g., durations [1,59] and distances [60,61]).
(iii) Bernoulli probabilities (proportions between numerosities) [62–64] and relative likelihoods (the probabilities of the observed events given various hypotheses about what generated them) [57,65,66].
(iv) Rates (numerosities divided by durations) [65,67].
(v) Directed (that is, signed) quantities (as in dead-reckoning [68] and the **vectors** that compose a cognitive map [12,69,70]).
(vi) Large quantities: bees navigate over mapped environments measured in kilometers ([69], pp. 246ff); the Arctic tern makes annual round trips between the Arctic and Antarctica, covering as much as 80 000 kilometers in a year [71].
(vii) Small quantities, such as the already mentioned probabilities.
(viii) The values of physical variables such as inertia and compliance that are essential to the brain's representation of the physics of our body and the objects it must manipulate, a representation upon which even simple behaviors such as grasping a glass depend [72].
(ix) Views (snapshots, patterned photon catches) [68,73].

Code for numbers

Gallistel believes that the brain processes information in the same manner as a conventional digital computer, by manipulating numerical symbols computationally. He states that any machine that manipulates quantities computationally must have a code for numbers (Gallistel, 2017). Therefore, to understand information processing in the brain, we must understand the code for the numbers it uses, that is, we must discover how numbers are encoded and stored in neural tissue, since they serve as the vehicles of neural computation (Gallistel, 2017).

He writes that there are a variety of different coding schemes to consider. Broadly speaking, there are unary codes and combinatorial codes. An example of a unary code is the hash mark, where the number 5 is represented by '|||||'. This type of code is highly inefficient. For example, say we wanted to convey the number 1,532 in the brain. If the brain uses a unary code like the hash mark system, it would have to send 1,532 action potentials (or bits) to send the number, which would take much too much time and energy.

On the other hand, to send the same number using a combinatorial code, like say the decimal coding system, we would only have to send 4 bits of information (one representing the value in the 1's place, on for the 10's place, 100's place, and 1'000's place = 1,532). Or if we wanted to convey 1,532 in binary code, we would have to send 11 bits of information ('10111111100' = 1,532). Gallistel's point is that combinatorial coding schemes are vastly more efficient, from both a timing and an energy expenditure perspective, than the unary schemes. We ought to bear this in mind when approaching the neural coding problem.

His paper appears to be discussing two different senses of neural code: the neural code for spikes (how numbers are transmitted in spike trains), and what he refers to as the "engram code", or the code for how numbers are written to the engram. I found it easy to confuse these two senses of the word 'code' he uses, but it is important to not conflate the two. He writes, "No known work to me has explicitly considered the engram coding question" (Gallistel, 2017, p.5).

To understand how quantitative information is written to the engram, we must first understand how numbers are embedded in the spike trains which convey them to the engram, and also how the numerical symbol contained in the engram is accessed to be used in neural computations underlying cognition. Most neuroscientists would say that quantitative information is stored in a pattern of connectivity between cells. However, Gallistel claims this is unlikely. If viewed from an information-theoretical perspective, it's easier to conceive of the neural code if numbers are stored in the brain using molecules inside the cell, and not written in a pattern of synaptic connections.

3. Gallistel's intracellular molecular mechanism of memory

In this section, I'll present Gallistel's reasoning behind why he believes the brain's memory storage mechanism will be found inside the cell, and not at the location of the synapse as indicated by the plastic synapse hypothesis. Since the neural code must be numerically based (given the foregoing information-theoretic reasons and that the logic of computation itself requires as such), we must find a storage mechanism for housing the numbers (or bit patterns) involved in neural computation.

Gallistel strongly believes that the memory storage mechanism is more likely to be found inside the cell in the form of an intracellular molecular memory storage mechanism. He refers to this idea as the *cell-intrinsic memory hypothesis*, which states that at least some memories are stored in altered or synthesized molecular structures inside the cell or within the cell membrane (Gallistel, 2017).

He claims that it is much easier to formulate a neural coding hypothesis if we assume that memories are stored using an intracellular molecular mechanism as opposed to specific patterns of connections between cells, as stated by the traditional Hebbian plastic synapse model for memory. Moreover, intracellular molecular mechanisms of memory fully satisfy all his information theoretic criteria characteristic of highly efficient and reliable information storage devices, while plastic synapses do not (I will explain these criteria he lists in detail in the next section). Lastly, he notes that computing with chemistry is "cheaper" than computing with action potentials – that is, more energetically efficient – so it makes sense to have memory devices at the biochemical level inside the cell rather than at the synaptic level.

DNA molecule: exemplar for biological memory mechanism

Gallistel is greatly inspired by the information storage mechanisms found inside the DNA molecule and claims the DNA exemplifies the sort of intracellular molecular mechanism we should be searching for in the brain for storing information into memory, and which serves as an important constituent materially realizing the neural code. He writes that a striking fact about DNA is that both digital computer memory and genetic memory have the same bipartite memory architecture (Gallistel, 2017b). A bit register in the RAM of a conventional digital computer stores both a *coding* portion and an *address* portion. To retrieve a specific memory, the computer sends a probe signal to the specific address location of that memory to access the information it contains in the coding portion.

Gallistel writes that genes also have these two parts for each memory location, or gene, DNA molecule: the coding portion of a gene and the promoter portion. Transcription factor proteins are analogous to probing signals in digital computers – they float around the cell and bind to their respective address on the gene, catalyzing expression of that gene.

Gallistel claims this example of an intracellular molecule (DNA) using an addressable memory structure strengthens his belief that the functional structure of all powerful computing systems, including the brain, are dictated by the inescapable logic of computation itself (Gallistel and King, 2009). The most efficient way to organize information into any computationally useful system is the old fashioned, tried-and-true digital method. If DNA uses this storage method, it is likely to be the case that the brain uses such a system as well. Why would it go all the way back to the drawing board and develop a radically different system?

Moreover, both classical digital computers and DNA compute using digital bit strings. Computers generate information-conveying symbols by building strings of two elements, the '0' state voltage level and the '1' state voltage level. DNA polynucleotide string works with 4 elements, the base nucleotides (A), cytosine (C), guanine (G), and thymine (T).

Lastly, Gallistel notes that the genetic code is truly symbolic, a fact which has been known since a few years after the discovery of DNA. In other words, there is no chemical or physical necessity connecting the molecular structure of the gene with the structure of the protein it encodes – the molecule produced its own innate coding language (Gallistel & King, 2009, p. 125). That is, there is no physically-constrained (i.e. mechanical) reason dictating why one three-letter codon codes for some specific amino acid. The association is arbitrary. “The divorcing of the code from what it codes for is the product of a complex multi-stage molecular mechanism for reading the code (transcribing it) and translating it into a protein structure”, similar to a classical digital system (Gallistel and King, 2009, p. 125).

The information conveyed by a gene is oftentimes abstract. For example, the *Pax6* gene is involved in eye development in vertebrates (Gallistel, 2017b, p. 2). When the *Pax6* gene from a mouse is expressed in *drosophila* (the fruit fly), the insect grows a faceted insect eye, not the lensed vertebrate eye. So, the abstract message contained in the *Pax6*, roughly translates to: “Build an eye here.” Moreover, the message conveyed by the *Pax6* gene doesn’t have specific semantic meaning until it is read by its biological receiver – just as Shannon wrote in his 1948 paper.

Given all of these reasons, Gallistel makes this rather grand claim: “We should aspire to be able to read the engram code and infer from it properties of the animal’s past experience as readily as molecular geneticists read the molecular genetic code and infer from it properties of the built animal” (Gallistel, 2017, p. 2).

DNA example is not generalizable to neural memory

I will argue that Gallistel’s explanation of the dynamics of hereditary information storage in the DNA molecule is not immediately relevant to the neural coding question, or the search for the material basis of neural memory. Gallistel states that the DNA molecule cannot be the brain's actual memory mechanism since it doesn't fulfill several of the necessary functional and behavioral requirements for cognitive memory. Accessing information in DNA takes much too long to be useful on the time scales required for memory recall. Moreover, while it may be true that DNA is a wonderfully efficient memory storage device, the type of information the DNA molecule evolved to store is quite different from the type of information the brain evolved to store.

His explanation of DNA as a computational device is missing the forest for the trees. The DNA molecule is not functionally equivalent (or even similar) to the mechanism for neural memory – the two memory storage mechanisms evolved to store widely dissimilar types of information. DNA stores hereditary information underlying morphological and physiological features of organisms and passes this information down through generations. Neural memory stores experienced information only for up to a lifetime and must be able to both store and recall information on rapid timescales, much quicker than information in DNA can be accessed by the cell (Shamir, M. et al., 2016, p. 1302). It seems reasonable to assume that if we are searching for a mechanism capable of immediately storing newly-acquired information from the environment, we shouldn't be looking to a molecule such that it takes millennia to alter the content of the information which it bears.

With such functional dissimilarities, it is unlikely that the structural and functional components realizing these two different systems will be similar – certainly not similar enough to make the claim that the engram "should look like a gene to which experience can write" (Gallistel, 2017, p. 2). The DNA analogy is misplaced and doesn't strengthen his argument calling for the existence of an intracellular molecular mechanism for neural memory.

Moreover, the computations performed on the information contained in the DNA strand are purely biochemical in nature. Accessing and utilizing the information stored in DNA code is done biochemically, not by means of any electrical signals as we observe in the brain. One important medium for sending information in the brain is electrical impulses. Thus, a mechanism which operates using only biochemical processes (like DNA) would undoubtedly be an incomplete model for the sort of mechanism realizing the types of computations the brain performs. The memory storage mechanism we are looking for must encode its information in such a way that it is at a minimum accessible by means of electrical signals.

Purkinje cell memory not generalizable to all types of neural memory

Furthermore, I also don't find his Purkinje cell example compelling evidence in favor of the claim that the mechanism for memory is likely to be found inside the cell. The Purkinje cell was shown to encode and store the value of an interstimulus interval (duration of time) between the onset of two different stimuli (the cell was trained on three different timing intervals: 150, 200, and 300 milliseconds) (Johansson, et al., 2014). This is a relatively simple type of memory, and one whose value is immediately experienced by the cell (i.e. the cell experienced the 150 ms interval between the two stimuli). This temporal value is a simple, straightforward form of memory, dissimilar to the more complex, abstract concepts and memories we know the brain to store. Some examples include higher-order concepts like freedom, a complicated scientific theory, the memory of what we did last Tuesday, or memories of some vacation we took. All these notions require much richer representations, and are not immediately experienced physical variables, like the simple conditioned memory for the temporal value. I don't find his Purkinje cell example to be generalizable to these other more complicated forms of memory.

4. Arguments against plastic synapse model

Since the dominant theory of memory storage in the brain is the plastic synapse hypothesis, naturally Gallistel targets this mechanism to potentially include in his theory as a memory storage device for housing the numbers involved in neural computation. He writes, "If the engram consists of altered synaptic conductances (the usual assumption), then we must ask how numbers may be written to synapses [altered synaptic conductances simply refers to memory information stored in the patterns of weighted connections between neurons)] (Gallistel, 2017, p. 1).

Can't store numbers at a synapse

In his paper, Gallistel's attack on the plastic synapse model for memory is three-fold. Firstly, since he insists that numbers are fundamental to the neural code, and that neural computation operates over numerical representations, identical to the digital computer, we must understand how numbers are encoded and stored in neural tissue. Since it is difficult, if not impossible, to conceive of how a number may be stored at the synapse, we ought to dispense with the model as a candidate memory storage mechanism, since it won't work in tandem with the processing parts of the cognitive architecture.

To recap, his argument is as follows:

1. Neural computations are digital.
2. We need a memory storage device for housing the numbers involved in digital neural computations.
3. It is difficult to explain how a number can be stored at a synapse.
4. Conclusion: the plastic synapse hypothesis, according to which memory storage occurs at the synaptic level, is false.

I find premise 1 and 2 to be false, thus his conclusion is unsound. His argument stating that since it is difficult to conceptualize how a number may be stored at the synapse, the entirety of the plastic synapse hypothesis ought to be thrown out is in my view a hasty move. His false conclusion is based on false premises as well as the mistaken assumption of truth of classicist theory of mind in which information processing is digital.

To explain plastic synapse theory, and to incorporate it in our understanding of the material basis of the engram, it is not necessary to understand how numbers are written to the synapse. The plastic synapse model would claim that the quantitative information would be stored in the pattern of synaptic connectivity between two or several engram cells, not stored at the location of a single synapse (Mu-Ming et al., 2016). The quantitative information would be accessed by activating a synaptic pathway containing a specific pattern of connectivity between cells, not by accessing a number stored at a single synapse.

Lastly, speaking to premise 1, it is unlikely that the brain computes using only digital variables (Piccinini, 2020). The type of computation the brain implements is *sui generis*, of its own kind. Neural computation doesn't appear to fit into the traditional digital (or analog) models of computation. Therefore, *sui generis* neural computation will require its own mathematical models and cannot be boxed into the mathematical formalisms of traditional digital (or analog) types of computation, as Gallistel appears to be attempting to do.

Synapses don't satisfy important criteria for a good memory mechanism

To bolster his attack on plastic synapse theory, Gallistel leverages some principles derived from classical information theory to construct an argument calling again for the disposal of the synaptic hypothesis as a candidate mechanism for memory. Based on information theoretic laws set by computing theory, he delineates a set of criteria that any reliable and efficient memory storage mechanism must satisfy, and as such, so should the mechanism we are searching for in the brain.

He writes that the function of a memory mechanism is to "carry large amounts of information forward in time in a computationally accessible form" (Gallistel, 2017, p. 7). The mechanism must meet the following criteria:

1. High thermodynamic stability (i.e. lasting a long time; not spontaneously disintegrate thus losing the information).
2. Low or negligible energy costs (not use up a tremendous amount of ATP to maintain the memory in storage).
3. Realizable in a maximally compact volume (store lots of information in the smallest possible volumetric space).

4. Capable of representing a huge range of information (store very small and big numbers)
5. Be addressable on a short time scale, quickly readable on the order of a few milliseconds.

"In the literature on synaptic memory hypothesis, it is difficult to find discussions of how altered synaptic conductances satisfy these desiderata" (Gallistel, 2017, p. 7). He uses this set of criteria to again attempt to discount plastic synapse hypothesis by claiming that the synaptic connections do not possess these qualities.

While I agree that good memory storage mechanisms do possess these traits, I disagree with his statement that plastic synapses don't display these qualities. To recap, Gallistel's argument is as follows:

1. Good memory storage mechanisms must invariably possess this set of criteria, as evident by information theory.
2. Plastic synapses do not possess these important traits.
3. Plastic synapses are not good memory storage mechanisms, thus the plastic synapse hypothesis fails, and we should look elsewhere for the memory storage mechanism in the brain.

I disagree with premise 2, thus I find his conclusion (3) to be invalid. High-functioning and effective memory storage mechanisms do possess the aforementioned properties (thermodynamic stability, negligible energy expenditure, etc.). However, his misapplication of these criteria to the synaptic model of memory (i.e. his claim that synapses don't exhibit these qualities) again leads him to a false conclusion.

On the contrary, the structural changes to the neural circuitry made during the formation of new memories leaves a lasting and stable change to the functioning of the circuit, thus exhibiting high thermodynamic stability (memories have the capability of lasting a lifetime (Mu-ming, P., 2016)). And, as evidenced by the experimental literature, when these patterns of synaptic connectivity break down, the information is lost (Ryan, T., 2021; Roy et al., 2016). There is energy expenditure to form the new synaptic connections, but once the connections are made, the neural architecture in which the memories are contained is stable, requiring only regular maintenance (Luis & Ryan, 2022).

Next, Gallistel says the memory mechanisms should be realizable in a maximally compact volume. The brain is a highly compact organ, and we know from electron microscopy that the connections between the synapses are extremely dense and compact. Lastly, the information contained in synaptic memory is quickly accessible, since we can recall any given memory very quickly, on the order of a few milliseconds - thus fulfilling Gallistel's final criteria.

He writes that the function of a memory mechanism is to carry large amounts of information forward in time in a computationally accessible form. I would argue that the computational accessibility of the information stored in synaptic memory will be evident when we have a better understanding of how these patterns of connectivity relate with the other components of the broader cognitive architecture responsible for accessing the information they contain.

Plastic synapses use inefficient rate codes

Gallistel claims that another reason to discount plastic synapses in favor of an intracellular molecular model for memory is that “rate codes are energetically inefficient and computationally awkward. A combinatorial code is more plausible” (Gallistel, 2017, p.1). The usual assumption among neuroscientists is that spike trains convey information between neurons via a rate coding scheme, that is, the # of spikes per window of time. In a rate code, the only value that matters is the number of spikes sent during the reading window. The interspike intervals, or time between spikes, is irrelevant to the rate code.

Gallistel states that the rate code for quantities is equivalent to a ‘hash mark’ code. Both are primitive unary codes, meaning they are inefficient signal messengers and are extremely costly energetically to send quantitative messages. It is better to consider an intracellular molecular mechanism, since the digital computations underlying cognition will be performed inside the cell instead of between cells. Thus, we don’t have to worry about sending numbers in spike trains between cells using such a poor coding scheme as the rate code.

However, there are several other coding schemes under experimental investigation other than rate codes, many of which could embody a combinatorial code to convey information in spike trains. Some examples include TTFS (time-to-first spike) coding, phase coding, and burst coding (Guo, et al., 2021). I don’t see mention of these alternative spike coding schemes included in his evaluation. If one could determine that it’s possible to use any of these alternative codes to send information (perhaps even quantitative information) more efficiently, in the form of some combinatorial code, his statement that plastic synapses convey information using inefficient coding scheme is called into question.

Are numbers vehicles of computation?

In this last section, I will refer back to the list of numbers in the brain (*Box 1*) Gallistel believes to serve as the numerical symbols underlying the neural computations guiding behavior. Gallistel writes, "On this hypothesis, quantities extracted from experience (*Box 1*) are written to the hypothesized molecular engram, and synaptic inputs read what has been written to control and direct subsequent behavior" (Gallistel, 2017, p. 1).

Firstly, several of the quantities Gallistel claims to be acquired directly from some external stimulus or experience, and which are subsequently written to the hypothesized molecular engram, are products of predictive processing made on the incoming environmental information. Take for example distance, one number in the brain he mentions in *Box 1*. Say you are standing on a shoreline next to a body of water and there's a sailboat exactly 347 meters away in the harbor. It appears that Gallistel is claiming that the brain encodes and stores a precise numerical value representing the distance between you and the sailboat, and this value is used to guide behavior with respect to that sailboat.

Clearly, the brain doesn't acquire the precise distance (347 m), but it makes its best prediction of the distance based on incoming information of the scene. It is questionable whether the brain encodes any precise numerical value at all, but only some vague notion of the distance. It would be more accurate to say that these values are products of neural computation (i.e. some predictive processing of the quantitative value), than to identify them as the symbolic vehicles of computation fundamental to neural computation.

The quantitative values he presents in *Box 1* that ostensibly serve as the stuff of neural computation are so vague and imprecise that it is impossible to encode a precise number for them. Many, if not all, of the quantities he mentions in *Box 1*, are not encoded in the brain as precise numbers at all, but only vague concepts. If numbers are central to neural information processing, these values are not acquired from experience but rather are produced in the brain innately.

If numbers do serve as the basis for the neural code, certainly they are not extracted from experience, since the numerical quantities in *Box 1* are too vague and unreliable to be useful for the neural code. Reliable vehicles of computation cannot be imprecise and vague notions. Without any concrete number being encoded, how are the neural computations supposed to be based on numbers, in the same manner that the binary string 10011010 serves in the computations performed in a digital computer? If reliably precise quantitative values are not acquired from the environment, where do these numbers come from which serve as the basis of neural computation?

Conclusion

In conclusion, I find no clear explanation in Gallistel's text that numbers are fundamental to the neural code. He doesn't make it obvious, or at least the connection remains elusive to me, how to bridge the sizeable gap between the uncontroversial statement that the brain stores quantitative information (which we know to be involved in cognitive processes underlying certain behaviors, like navigation and predation), to the much stronger claim that *all* information processing in the brain is carried out on numerical symbols stored in neural tissue. The vehicles of neural computations don't necessarily have to be numerically based, but whichever variables will work.

His reference to examples of other biological memory mechanisms are not immediately relevant to the problem of neural memory. Moreover, his reference to Shannon information theory doesn't consider that the explanations of the type of computations the brain undertakes may require its own set of mathematical formalisms which don't align with the older, classical models of digital computation. His views don't pay adequate consideration to the experimental evidence in favor of the Hebbian model of memory. Adopting such a heavily theoretical view narrows his perspective of the issue.

This narrative brings to light important questions in the search for the neural code and the material basis of memory. To what extent is a purely theoretical approach to these questions profitable in our search? Granting too much weight to the theory at the expense of careful consideration of the experimental evidence blinds us to other ways of thinking about the problem.

References

Gallistel, C. R. (2017). The Coding Question. *Trends in Cognitive Sciences*, *21*(7), 498–508. https://doi.org/10.1016/j.tics.2017.04.012.

Gallistel, C. R. (2017b). Numbers and brains. *Learning and Behavior* 44(1).

Gallistel CR. (2017c) Finding numbers in the brain. *Phil. Trans. R. Soc.* B 373: 20170119. http://dx.doi.org/10.1098/rstb.2017.0119.

Gallistel, C. R. (2018). Finding numbers in the brain. Philosophical Transactions: Biological Sciences, 373(1740), 1–10. https://doi.org/10.2307/44679408.

Gallistel, C. R. & King, A. P. (2009). Memory and the computational brain: Why cognitive science will transform neuroscience. Wiley and Blackwell (UK).

Guo, W., Fouda, M. E., Eltawil, A. M., & Salama, K. N. (2021). Neural Coding in Spiking Neural Networks: A Comparative Study for Robust Neuromorphic Systems. Frontiers in Neuroscience, 14, N.PAG. https://doi.org/10.3389/fnins.2021.638474.

Johansson, F. et al. (2014) Memory trace and timing mechanism localized to cerebellar Purkinje cells. Proc. Natl. Acad. Sci. U. S. A. 111, 14930–14934.

Johnson, K. (2000) Neural Coding. *Cell Press,* Neuron, Vol. 26, 563-566.

Luis, C. O. S., & Ryan, T. J. (2022). Understanding the physical basis of memory: Molecular mechanisms of the engram. *Journal of Biological Chemistry, 298(5),* 1–23. https://doi.org/10.1016/j.jbc.2022.101866

Maya Shamir, Yinon Bar-On, Rob Phillips, Ron Milo. (2016). SnapShot: Timescales in Cell Biology. *Cell*, 164(6), 1302-1302.e1. ISSN 0092-8674, http://dx.doi.org/10.1016/j.cell.2016.02.058.

Roy, D. S., Arons, A., Mitchell, T. I., Pignatelli, M., Ryan, T. J., & Tonegawa, S. (2016). Memory retrieval by activating engram cells in mouse models of early Alzheimer's disease. *Nature, 531*(7595), 508–512. https://doi.org/10.1038/nature17172.

Ryan, T. J., Ortega-de San Luis, C., Pezzoli, M., & Sen, S. (2021). Engram cell connectivity: an evolving substrate for information storage. *Current Opinion in Neurobiology*, *67*, 215–225. https://doi.org/10.1016/j.conb.2021.01.006.

Shannon, C. E. (1948). A mathematical theory of communication. *The Bell System Technical Journal*, *27*(3), 379-423. doi: 10.1002/j.1538-7305.1948.tb01338.x.

Piccinini, G. (2020). Neurocognitive mechanisms: explaining biological cognition (First edition.). Oxford University Press.

Mu-ming Poo, Pignatelli, M., Ryan, T. J., Susumu Tonegawa, Bonhoeffer, T., Martin, K. C., Rudenko, A., Li-Huei Tsai, Tsien, R. W., Fishell, G., Mullins, C., Gonçalves, J. T., Shtrahman, M., Johnston, S. T., Gage, F. H., Yang Dan, Long, J., Buzsáki, G., & Stevens, C. (2016). What is memory? The present state of the engram. BMC Biology, 14, 1–18. https://doi.org/10.1186/s12915-016-0261-6.