



**HAL**  
open science

# Visceral Signals Shape Brain Dynamics and Cognition

Damiano Azzalini, Ignacio Rebollo, Catherine Tallon-Baudry

► **To cite this version:**

Damiano Azzalini, Ignacio Rebollo, Catherine Tallon-Baudry. Visceral Signals Shape Brain Dynamics and Cognition. Trends in Cognitive Sciences, 2019, 23 (6), pp.488-509. 10.1016/j.tics.2019.03.007 . hal-02355098

**HAL Id: hal-02355098**

**<https://hal.science/hal-02355098>**

Submitted on 21 Dec 2021

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons Attribution - NonCommercial 4.0 International License

# Testing models of human declarative memory at the single-neuron level

Ueli Rutishauser<sup>1-4</sup>

<sup>1</sup> Department of Neurosurgery, Cedars-Sinai Medical Center, Los Angeles, CA

<sup>2</sup> Department of Neurology, Cedars-Sinai Medical Center, Los Angeles, CA

<sup>3</sup> Center for Neural Science and Medicine, Department of Biomedical Sciences, Cedars-Sinai Medical Center, Los Angeles, CA

<sup>4</sup> Division of Biology and Biological Engineering, California Institute of Technology, Pasadena, CA

Contact Author: Ueli Rutishauser, [ueli.rutishauser@cshs.org](mailto:ueli.rutishauser@cshs.org)

## Abstract

Deciphering the mechanisms of declarative memory is a major goal of neuroscience. While much theoretical progress has been made, it has proven difficult to experimentally verify key predictions of some foundational models of memory. Recently, single-neuron recordings in human patients have started to provide direct experimental verification of some theories, including mnemonic evidence accumulation, balance-of-evidence for confidence judgments, sparse coding, contextual reinstatement, and the VTA-Hippocampus loop model. Here, we summarize the cell types that have been described in the medial temporal lobe and posterior parietal cortex, discuss their properties, and reflect on how these findings inform theoretical work. This body of work exemplifies the scientific power of a synergistic combination of modelling and human single-neuron recordings to advance cognitive neuroscience.

## Keywords:

Declarative Memory, human single-neuron, hippocampus, dopamine, memory retrieval, medial temporal lobe

## 1 Towards a Circuit-Level Understanding of Human Declarative Memory

2       **Declarative memories** (see Glossary) of events and facts are central to human behavior and are  
3 part of what defines each of our individual identities [1]. Consequently, deciphering the mechanisms that  
4 allow humans to form, maintain, and retrieve declarative memories is a major topic in cognitive  
5 neuroscience. Brain mapping and lesion techniques have started to reveal the brain networks that are  
6 involved in declarative memory formation, maintenance, consolidation, and retrieval [1]. This work has  
7 highlighted the distinct roles of different brain areas, such as the hippocampus and other parts of the  
8 medial temporal lobe, cortical areas such as the posterior parietal cortex, and the basal ganglia in  
9 declarative memory. Theoretical work, on the other hand, has proposed both abstract cognitive and  
10 circuit-level mechanisms on how different types of neurons within these areas interact, and how their  
11 response changes as a function of learning to encode new memories (**'engrams'**) [2-10]. Many such  
12 models make predictions that are testable using behavioral or non-invasive neuroimaging-based  
13 experiments, allowing for close interaction between theory and experiment that has been highly fruitful.  
14 Other model predictions, however, require either higher spatial and/or temporal resolution or high  
15 quality single-trial measurements to be tested directly. Due to this, it is challenging to test such  
16 predictions with non-invasive techniques, which have relatively low temporal and spatial resolution and  
17 typically require across-trial averaging. Invasive brain mapping techniques such as **electrocorticography**  
18 (**ECoG**) and **intracranial electroencephalography (iEEG)** performed in patients have significantly higher  
19 resolution and signal quality, an advantage that has already led to significant advances in our  
20 understanding of memory (see for example [11-14]). Nevertheless, there remain a considerable number  
21 of model predictions that are best tested at the level of individual neurons.

22       Invasive recordings in animals have provided an invaluable body of detailed knowledge of the  
23 circuits involved in memory at the systems, circuit, cellular, and molecular level. But, if our goal is to  
24 understand human memory, this animal work has to be complemented by human experiments at similar  
25 levels of resolution. To achieve this, invasive recordings in humans are critical for three reasons. First,  
26 findings from animals have to be validated in humans to build a bridge between humans and animal  
27 model systems and thereby establish the validity and limitations of the model system. Second, aspects  
28 of memory that are uniquely developed in humans or which cannot be practically studied in animals have  
29 to be investigated in humans. Third, some aspects of cognitive models of human memory can best be  
30 validated and refined using data recorded at the scale of individual neurons or small groups thereof.  
31 Here, we review what has been discovered from single-neuron recordings in humans (see Box 1) and

1 illustrate how such recordings have been able to bridge between theory and experimentation and  
2 thereby provide new mechanistic insight into human memory.

### 3 4 **Single-Trial Learning: Assessing the Hippocampus VTA/SN Loop Model**

5 Learning theory is largely focused on repetitive learning that requires many repetitions for a  
6 robust memory to be formed. In contrast, consider the kind of learning that allows a human being to  
7 recognize an image with high accuracy that was previously only seen once as part of a sequence of 10'000  
8 images, each shown once for 1s [15]. How the declarative memory system achieves such rapid and high-  
9 capacity learning is a key open question.

10 A model that has motivated much research on the underlying mechanisms is the Hippocampus-  
11 **Ventral Tegmental Area (VTA)/Substantia nigra (SN)** loop model [2, 3, 16]. This model proposes that the  
12 rapid encoding of novel stimuli is facilitated by transient **dopamine (DA)** release that is triggered by novel  
13 stimuli (also see Box 2). Anatomically, the loop model (Fig. 1A) hypothesizes that, first, circuits within the  
14 hippocampus detect that a stimulus is novel. It has long been known that the hippocampus, along with  
15 upstream areas such as perirhinal cortex [17, 18], is crucial for detecting novelty [19]. This novelty signal,  
16 in turn, is thought to activate DA neurons, which then release DA in the hippocampus. DA release  
17 strengthens synaptic plasticity and enables late long-term potentiation (LTP) [3]. Among the predictions  
18 of the loop model are: First, subsets of DA neurons should be novelty sensitive, i.e. they should be  
19 activated by stimuli which have not been seen before. Secondly, this novelty signal should be conditional  
20 on a novelty signal being present in the hippocampus. Thirdly, the novelty signal should appear first in  
21 the hippocampus, followed later by DA neurons. Forth, following novelty-dependent activation, transient  
22 increases of DA occur in the hippocampus and these increases strengthen late LTP. Fifth, the model  
23 makes specific hypothesis on the anatomical pathways by which hippocampal novelty signals reach the  
24 SNc/VTA.

25 Single-neuron recordings have started to provide experimental support in humans for a subset  
26 of the above predictions. There are two groups of relevant findings: novelty responses in the  
27 hippocampus and other parts of the **medial temporal lobe (MTL)**, and novelty responses of DA neurons.  
28 In the MTL, a subset of neurons preferentially increase their firing rate only the first time a stimulus is  
29 presented [20, 21]. Comparing the response of such novelty-signaling neurons between the first and  
30 second time a stimulus is presented reveals a strong difference, supporting the notion that their  
31 response is different after a single exposure. This is in contrast to habituation, which leads to a gradual



1 reduction in response strength of some human MTL neurons over many trials [22]. Habituation develops  
2 over tens of trials and responses recovers if other stimuli are shown intermittently. In contrast, novelty-  
3 neurons in the MTL do not respond again to the same stimulus even after long periods of time [20].

4 DA neurons in the VTA/SN are intermingled with other cell types, making it difficult to measure  
5 their activity with non-invasive methods. With extracellular recordings, however, it is possible to  
6 differentiate between different cell types due to differences in action potential waveforms and spike  
7 train statistics. This way, individual DA neurons in the human SN can be observed in patients undergoing  
8 **deep brain stimulation (DBS)** implantation for treatment of movement disorders, typically Parkinson's  
9 disease (PD) [23-25] (see Fig. 1D). Note that while PD patients have dopamine cell loss, this loss typically  
10 begins in the ventral tier of the SNc, leaving the neurons in the dorsal tier of the SNc intact [26]. It is the  
11 dorsal tier that projects to the hippocampus, ventral striatum and cortex [26] and which is typically  
12 recorded in humans, making it suitable to investigate the DA system in human PD patients. Such  
13 experiments have revealed two sets of DA neuron responses with respect to novelty: one that increases  
14 its firing rate to novel stimuli, and one that does so for familiar stimuli [24, 27] (Fig. 1B-C shows the  
15 novelty type). Human DA neurons also signal aspects of rewards [23], but whether the same or different  
16 DA neurons respond to both reward and novelty remains unknown. The response properties of the  
17 novelty subtype of DA neurons are compatible with the loop model: their response diminishes after a  
18 single learning trial, the extent of response change between the first and second presentation is  
19 predictive of behaviorally assessed memory strength, and the latency of this novelty response is later  
20 than that of MTL novelty neurons [24]. Interestingly, this work also revealed a familiarity subtype of DA  
21 neurons, a type of response not predicted by the loop model. These neurons increased their firing rate  
22 only for previously seen stimuli, a response which might have a role in strengthening or consolidating  
23 existing memories. Together, this data illustrates the extent to which the loop model has guided ongoing  
24 work and the richness of insight that can be obtained by such model-guided experimental work in  
25 humans.

26

## 27 **Representations of memory content: Assessing sparse and distributed coding models**

28 What are the features that are used to encode and retrieve a declarative memory? Declarative  
29 memory has two subsystems: **semantic memory** and **episodic memory**, which represent knowledge of  
30 concept/facts, and details of autobiographical events, respectively [1, 28]. A theoretical question that  
31 has motivated much work is the level of granularity at which aspects of these two kinds of memory are

1 represented. There are two opposing views [8-10, 29, 30]: on the one hand are models that indicate that  
2 a sparse and highly selective representation is optimal for fast learning [8, 9]. On the other extreme are  
3 fully distributed coding views [31, 32], in which only the pattern of activity across large groups of neurons  
4 can differentiate between different concepts and experiences. Under this second view, listening to a  
5 single neuron in the MTL would not be useful to make a high-level semantic decision such as 'am I viewing  
6 an animal'. In between the two extremes are models that propose a sparse but distributed form of  
7 coding; under this view, a given neuron has high response sparsity (it only responds to a small subset of  
8 all possible stimuli) but there are many neurons that respond to the same subset of stimuli [8, 10, 33].

9 A second, orthogonal, question that is motivated by the semantic vs. episodic memory subsystem  
10 view is whether the neurons encoding concepts/facts also encode aspects of episodic memory (such as  
11 memory strength, where, when) or whether the substrate for these two aspects of declarative memory  
12 are separate. Lesion studies indicate that they can be separate in the case of remote memories: whereas  
13 patients with hippocampal lesions are unable to form both episodic and semantic memories, they can  
14 retrieve remote but not recent semantic memories [34].

15 Compatible with the theoretical prediction of sparse but distributed coding, single-neuron  
16 recordings in humans reveal that concepts and facts are encoded in a sparse and invariant manner by  
17 single MTL neurons [35-39]. The response of these neurons is highly selective (sparse) but at the same  
18 time invariant: a given neuron only responds to a small subset of all tested stimuli, but the subset it  
19 responds to is closely related (i.e. all images of animals). At the same time, the code is distributed  
20 because there are many neurons with the same or highly similar selectivity [38]. Such neurons have  
21 become known as 'category' or 'concept' cells (Fig. 2A-C shows examples), a kind of neurons that are  
22 also commonly referred to as 'gnostic units' and 'grandmother cells', respectively [40]. Here, we refer  
23 to this group of neurons as '**visually selective**' (VS) cells. VS cells are reproducible across tasks, recording  
24 techniques, and laboratories [35-39, 41-43] (see [41] for a review) and their responses are abstract and  
25 multimodal. For example, a cell tuned for a particular individual responds to a variety of images showing  
26 the person, the written name of that person as well as to auditory input saying the individual's name [44]  
27 (Fig. 2B). In addition to sensory input, VS cells can be activated by thought (including free recall and  
28 maintenance in working memory) [39, 45-48]. Also, for identical visual input, the response of VS cells  
29 varies as a function of whether the stimulus entered conscious awareness [42, 49, 50] and if so is  
30 indicative of subjective decisions made about the stimulus [51, 52]. These properties of VS cells are also  
31 properties of declarative memories, supporting the view that VS cells are part of the representation of

1 such memories. VS cell responses are plastic: they are more likely to represent concepts of personal  
2 relevance [53, 54] and can change their tuning by associative pairing [55]. Together, this data motivates  
3 the hypothesis that VS cells in the human MTL represent sparse, abstract, and selective features of  
4 declarative memories, a view compatible with the sparse but distributed coding model. In contrast, this  
5 data is not compatible with fully distributed coding models. One hypothesis motivated by this data is that  
6 VS cells represent semantic memories, and thereby constitute an integral part of an engram [56].

7 A second type of response in the human MTL that has been characterized are **memory selective**  
8 **(MS) cells**. The activity of MS cells correlates with aspects of episodic memory and their activity exhibits  
9 single-trial plasticity [20] because their response is conditional on whether a stimulus has been seen  
10 before. Two types of MS cells (Fig. 2D) have been described [20, 21, 57, 58]: one that increases its firing  
11 only for novel stimuli, and one that increases its firing only for previously seen stimuli (similar cells exist  
12 in macaque perirhinal cortex [17]). The magnitude of MS cell activity in response to a previously seen  
13 stimulus is indicative of memory strength, with stronger memories resulting in larger changes both  
14 during **recognition memory** [58] and spatial cued recall tasks [59]. The firing rate of MS, but not VS, cells  
15 is indicative of the subjective confidence of recognition memory decisions. This constitutes a remarkably  
16 trial-by-trial correlation with the declarative aspect of episodic memories with only one type of cell [58].  
17 Jointly, this indicates that MS cells represent aspects of episodic memory. Similar to VS cells, the code  
18 formed by MS cells is sparse and highly invariant, with the large majority of MS cells not also qualifying  
19 as VS cells [58]. Notably, MS cells respond later than VS cells, with an average response latency that is  
20 ~180ms longer relative to that of VS cells (a delay attributed to the theta oscillation by hippocampal  
21 models [5]). Together, this data shows that VS and MS cells in the human MTL together form an  
22 orthogonal code in both feature space and time. This result supports a sequential model of declarative  
23 memory, in which high-level sensory categorization first leads to reactivation of related concepts in  
24 semantic memory, followed by recognition whether the currently experienced stimulus has been seen  
25 before or not, and if so recall of related attributes and reinstatement of context.

26 What aspect of the highly sparse and abstract responses of VS and MS cells in the human MTL  
27 are computed locally and which are inherited from upstream areas such as perirhinal or inferotemporal  
28 cortex? One important insight on this question comes from analysis of response latencies of individual  
29 neurons relative to stimulus onset. The response latencies of VS cells (which respond earlier than the MS  
30 cells) in the human amygdala, hippocampus, and entorhinal cortex are ~300-400ms [35, 58, 60], a delay  
31 that is considerably longer than in higher visual areas and which would be too slow for many perceptual

1 processes [61] (but see [62] for an argument that the MTL in addition also has a critical role in visual  
2 perception). Also, the latencies of VS cells are inversely proportional to selectivity, with more selective  
3 responses occurring later [35]. Intriguingly, VS cells recorded in parahippocampal cortex are less selective  
4 and respond about 100ms earlier than those in entorhinal cortex, hippocampus, and amygdala,  
5 indicating that as information propagates through the MTL, responses become more and more selective  
6 and sparse. In macaques, perirhinal cortex neurons and the BOLD-fMRI signal differentiate familiar from  
7 novel stimuli, suggesting that they encode mnemonic aspects of the stimuli shown rather than their  
8 physical attributes [17, 63]. Also, selective perirhinal lesions impair recognition memory [64], indicating  
9 that the perirhinal cortex might be the first anatomical area downstream to high level sensory areas in  
10 which responses become contingent on experience [65]. However, its exact contribution to declarative  
11 memory remains unclear [18, 66, 67]. To our knowledge, no single-neuron responses have been  
12 performed in human perirhinal cortex, leaving their latency, novelty-selectivity, and sensitivity to explicit  
13 declaration of memories unknown. Such recordings will be critical to determine what specific  
14 transformations in both selectivity and latency signals undergo as they propagate from perirhinal cortex  
15 into the hippocampal system.

16

### 17 **Retrieving existing Memories: Assessing the Reinstatement of Temporal Context Model**

18 What allows the episodic memory system to differentiate similar memories that happened at  
19 different points of time and to selectively retrieve temporal clusters of memories? While temporal  
20 clustering of retrieval is ubiquitous [68, 69], the aspect of memory search that give rise to this effect  
21 remain poorly understood. A class of theoretical models for studying these questions are temporal  
22 context models (TCM) [70-73]. TCMs propose that neural activity in a subset of cells drifts as time  
23 progresses, leading to identical stimuli perceived at different times being accompanied by different  
24 neuronal states. This contextual information is then combined with sensory information and encoded  
25 into memory. Within this framework, successful retrieval of a memory results in re-instatement of the  
26 neuronal state at the time of encoding (“contextual reinstatement”). As a result, a ‘jump back in time’ or  
27 ‘mental time travel’ occurs, a defining feature of episodic memory [74]. As part of reinstatement, the  
28 context serves as a cue to then retrieve other associated attributes of a memory, a process leading to  
29 the subjective impression of ‘recollecting’ [74]. Similarly, the reinstated context serves as a cue to recall  
30 other nearby memories, thereby explaining effects of temporal encoding order on recall [70-72]. A  
31 second prediction of TCM is that strong memories, which result in recollection, are accompanied by

1 reinstatement, whereas weaker memories are not. This is because the former is associated with a 'jump  
2 back in time' [75], a feature of reinstatement sensitive to MTL damage [76]. Intracranial recordings have  
3 revealed direct evidence for these predictions of TCM.

4         Neural activity in the MTL drifts slowly over both short (seconds)- and long (minutes) timescales  
5 at both the single-neuron [77, 78] and field potential level [79, 80]. This effect is visible as auto-  
6 correlation of activity at the single-neuron level over minutes [77, 78]. This temporal drift was not  
7 restricted to neurons which do not respond to sensory inputs because it is apparent event in neurons  
8 that are visually tuned [77]. This data shows that items which are seen at similar periods of time will be  
9 accompanied by more similar neural activity, thereby resulting in a more similar memory representation.

10         The neural state at encoding is reinstated, that is the patterns of activity during encoding and  
11 retrieval of the same item are similar. Reinstatement is visible at the single-neuron level in several areas  
12 of the temporal lobe, including the hippocampus, amygdala, and middle temporal gyrus (MTG) [77, 78,  
13 81] (Fig. 2E-G). What remains poorly understood is what information is reinstated – is it specifically  
14 temporal context or rather other internal or external features? One notable exception is spatially tuned  
15 cells, which reveal specific reinstatement of spatial location [82], indicating that reinstatement can be  
16 specific. Remarkably, reinstatement is sufficiently powerful to be visible at the aggregate field potential  
17 [79, 80] and **blood-oxygen level dependent (BOLD)** functional magnetic resonance imaging (fMRI) [83]  
18 level. Reinstatement of activity has been seen both during recall (cued [80, 81] and free [79] recall) as  
19 well as recognition memory [77, 78]. In the former, such reinstatement occurs 0-2s before a subject  
20 indicates recall of an item [80, 81], whereas in the later reinstatement occurs following stimulus onset.  
21 Third, the extent of reinstatement is related to the quality of memory retrieval: it is predictive of whether  
22 cued recall occurs correctly or incorrectly [81] (Fig. 2E) and whether an item is recognized with high or  
23 low confidence [77].

24         In the case of recognition memory, the same items are shown during both encoding and  
25 recognition, making it necessary to disambiguate between activity representing stimulus-specific sensory  
26 input from reinstated activity. However, the reinstated context changes only gradually, and should  
27 thereby still be similar to different encoded items which were shown closely (in time) before or after the  
28 recognized item. Indeed, both forward-and backward **contiguity effects** are visible at the single-neuron  
29 level [77] in recognition memory (Fig. 2F-G). Uniquely, in recognition memory tasks, recency effects can  
30 be differentiated from contiguity effects [57]. Using this approach, it has been shown that contiguity  
31 effects (due to reinstatement) are present at the single-neuron level in the absence of recency effects

1 [77]. While forward-and backward contiguity effects remain to be shown during cued or free recall at  
2 the single-neuron level, field potential studies show robust reinstatement during recall as well [79, 80].  
3 Together, there is thus evidence at the single-neuron level for several predictions of TCMs, making them  
4 excellent candidates for continued theoretical study. Of note, TCM models make predictions that relate  
5 to the process and experience of recollection and memory search during free recall. While important,  
6 here our focus is only on the aspects of TCM that have been studied in the context of recognition  
7 memory.

### 9 **Converting Memories into Decisions: Assessing the Mnemonic Evidence Accumulation Model**

10 Many of the decisions we make in daily life depend on previous experience. Take, for example,  
11 deciding whether a person you see on the bus is the same person you met last night at a party. This  
12 decision relies on memory retrieval, the integration of different kinds of information, and meta-cognitive  
13 processes to assess certainty. A model that encapsulates one point of view of how the nervous system  
14 makes decisions is the drift diffusion model (DDM) [84]. The DDM proposes that **leaky integrators** [85]  
15 accumulate evidence in favor of all possible choices and that the action for a particular choice is initiated  
16 once the total accumulated evidence for a choice exceeds a threshold (Fig. 3A). While the DDM was  
17 originally developed for memory-based decisions [86], much of the experimental work so far has focused  
18 on perceptual decisions. In particular, individual neurons have been found in macaques [87-89] and  
19 rodents [90] whose firing rate reflects integrated sensory evidence and predicts choices [84]. This  
20 discovery has led to an unprecedented mechanistic understanding of perceptual decision making. It  
21 remains unknown whether a similar mechanism is at work for declarative memory-based decisions and  
22 if so, whether the same or different neurons support this process.

23 Based on prior work on perceptual decision making in macaques [87-89], connectivity analysis  
24 for areas that receive hippocampal output, lesion studies, and neuroimaging, one candidate area in  
25 humans that has emerged for where mnemonic evidence-integrating neurons might be located is the  
26 left posterior parietal cortex (PPC) [91, 92] (Fig. 3B; see [91-93] for detailed reviews). In recognition  
27 memory, different parts of PPC exhibit differences in BOLD signal activation between new vs. old,  
28 recollected vs. recognized, or high vs. low confidence items [91, 92]. Scalp EEG source localization  
29 similarly indicates that PPC is the source of an ERP that differentiates between new and old items [94].  
30 PPC lesions largely do not reduce recognition accuracy (but see [95]), but rather manifest as  
31 impoverished autobiographical recall, reduced likelihood of recollection, and reduced confidence. Thus,

1 while information is present, PPC lesions lead to an inability to properly access this information for  
2 further processing – a kind of ‘memory neglect’ [93]. While it remains unclear what specifically the PPC  
3 contributes to memory retrieval, a hypothesis is that a key contribution is the accumulation of mnemonic  
4 evidence [91, 96].

5 Invasive recordings from human PPC [97-99] provide insight into this prediction. ECoG recordings  
6 from left PPC [97] reveal a striking functional heterogeneity when comparing response patterns of high-  
7 gamma band power (HGP) between different parts of the PPC (Fig. 3B). In the intraparietal sulcus (IPS),  
8 HGP was higher for old compared to new items, whereas in the superior parietal lobule (SPL), HGP power  
9 was higher for new compared to old items. The time course of the signal relative to stimulus onset also  
10 varied between the two areas: HGP first differentiated between new and old stimuli in SPL (200-300ms),  
11 with signals in IPS differentiating only later (300-700ms). Aligned to button-press, signals in IPS  
12 differentiated between old and new stimuli up to ~200ms before button press, but not later. In light of  
13 the mnemonic integration framework, this data suggests that neurons in IPS integrate evidence for old,  
14 but not new, stimuli, and once a threshold is reached a motor action is initiated.

15 Single-neuron recordings in the IPS [100] of two human subjects participating in a brain-machine  
16 interface clinical trial provide further evidence for a role of the PPC in memory retrieval [98]. Within a  
17 small 4x4 mm patch of PPC, two types of neurons with signals relevant for memory retrieval were found:  
18 memory-selective (MS) and confidence-selective (CS) neurons (Fig. 3D). There were two types of MS  
19 neurons: one that increased its firing rate for familiar stimuli, and one that increased its firing rate for  
20 novel stimuli (below referred to as the preferred stimulus). This firing rate increase was graded with  
21 memory strength as measured by the reported confidence (Fig. 3C-D). Notably, this modulation by  
22 confidence was restricted to the preferred stimulus of the MS cell, with the activity during non-preferred  
23 trials not modulated. CS cells, on the other hand, increased their firing rate either for high-or low  
24 confidence retrieval decisions regardless of whether the stimulus has been seen before. In contrast to  
25 recordings during the same task in the MTL [58], there was no evidence in PPC for neurons carrying  
26 information about stimulus identity. Rather, both CS and MS neurons signaled non-stimulus specific  
27 mnemonic information. Errors trial analysis further revealed that these neurons carried choice signals  
28 (Fig. 3E). The recorded neurons started to differentiate between the two choices well before the motor  
29 response but significantly later than the latency at which the MS signal is available in the MTL [58, 97,  
30 99, 100]. This result reveals a single-neuron candidate for mnemonic integrators in human cortex.

1           Although neurons in PPC are remarkably heterogeneous [98], BOLD-fMRI and ECoG studies [97]  
2 show that novelty, memory strength, or recollection is indicated by the average activity across large  
3 subareas of PPC [99]. This raises the important question of how heterogeneous signals at the single-  
4 neuron level can give rise to such differences on a larger scale. One limitation is that the recordings  
5 discussed here were performed at the border between SPL and the supramarginal gyrus (SMG), making  
6 it possible that a transition area exists between the two. A critical open question is how declarative-  
7 memory based information is accessed by neurons in PPC. One putative mechanism is that PPC neurons  
8 transiently phase-synchronize their activity with hippocampal theta oscillations when integrating  
9 hippocampal information but testing this prediction will require simultaneous recordings. While such  
10 recordings have not yet been performed, it is known that during autobiographical recall, theta-  
11 oscillations transiently synchronize between PPC and MTL [101].

12

13 **Meta-Cognitive Confidence Judgments: Assessing the Balance of Evidence Model**

14           The assessment of confidence is a hallmark of declarative memory. Theoretical work has  
15 advanced several potential ways by which **confidence judgments** might be made [102-105]. One class of  
16 models is the balance of evidence model (BEM) [105], which is an extension of the evidence  
17 accumulation model [106] (Fig. 3F). In the case of two possible choices (i.e. old or new), the BEM consists  
18 of two integrators that each accumulate evidence. The “balance of evidence” is the absolute difference  
19 between the two accumulated evidence values and is proportional to the confidence. Importantly, the  
20 two decisions (the actual choice and the confidence) are made at the same time using the same  
21 mechanisms. This contrasts with other models, in which the two decisions are made sequentially [107].  
22 The BEM makes specific predictions about the underlying neural correlates (Fig. 3F). These include: there  
23 are separate neurons that integrate evidence only for old or new stimuli, the neurons that supply the  
24 evidence as well as the neurons that integrate the evidence are modulated by underlying memory  
25 strength for only their preferred stimuli (i.e. new or old), and the RT, accuracy, and confidence is  
26 proportional to the balance of evidence. The latter prediction implies that confidence is influenced by  
27 the accumulated evidence for both the winning and the losing choice.

28           Support for some of the BEM’s predictions have been found in both macaques [87] as well as  
29 humans [58, 98] at the single-neuron level. First, during declarative memory-based decisions, putative  
30 neurons that integrate evidence in PPC [98] or that supply input to the integration process in MTL [58]  
31 are modulated by evidence strength as reported by subjective confidence (Fig. 3G-H). This modulation



1 is only apparent for the preferred, but not the non-preferred stimulus. MS cells in the MTL and memory-  
2 choice cells in the PPC thus represent signals at the input and output stages of decision making that  
3 correlate with confidence as predicted by the BEM model. Second, estimating the balance of evidence  
4 in individual trials from the firing rate of MS neurons correlates well with RT, accuracy, and declared  
5 confidence. Together, this data thus reveals a remarkable predictive power of the theoretical quantity  
6 of balance of evidence [58], indicating that the BEM model has good explanatory power and can bridge  
7 experiment and theory. Note that these experiments make excellent use of the unique ability of humans  
8 to declare their confidence, thereby allowing a direct (albeit subjective) assessment of memory strength.  
9

10 **Concluding remarks**

11 In this review, we summarized what invasive recordings at the single-neuron level in humans have  
12 revealed about the mechanism of human declarative memory. Our emphasis was on illustrating the  
13 power of this approach by demonstrating how it has enabled direct testing of predictions made by five  
14 models of relevance for declarative memory: the hippocampus/VTA loop model, the sparse coding  
15 model, the contextual reinstatement model, the evidence accumulation model, and the balance of  
16 evidence model (see Table 1 for a summary). While there are of course many other important models,  
17 here we focus on this subset because together they provide a good perspective of the power of the  
18 overall approach. In each case, the combination of human behavior and simultaneous single-neuron  
19 recordings has revealed critical new insights into different aspects of human memory. Jointly, this data  
20 now allows us to synthesize a view of the processing elements and the information flow among these  
21 processing elements in the human brain during memory encoding and retrieval (Fig. 4, Key Figure) and  
22 to tentatively map aspect of this circuitry to specific kinds of neurons in particular brain areas (of course  
23 this abstract model is far from being an actual implementable circuit). While here the focus was  
24 specifically on declarative memory, a similar combination of model-driven analysis of intracranial  
25 recordings has started to provide essential new insights into other human cognitive processes. While still  
26 in its early stage, this powerful approach has a bright future and we anticipate much exciting future work  
27 of this kind that will push ahead our understanding of human cognition in ways not possible with other  
28 experimental approaches (see Outstanding Questions).

29  
30 **Acknowledgements**

1 I thank all members of the Rutishauser Laboratory, Adam Mamelak, Ralph Adolphs, Tyson Aflalo,  
2 Richard Andersen, Doris Tsao, and Wolfram Schultz for discussion. The work described here was  
3 supported by NIH (R01MH110831, U01NS103792, P50MH094258), the McKnight Endowment for  
4 Neurosciences, and the NSF (BCS-1554105).

5

1 **Figures legends**

2

3 **Figure 1: The Lisman Hippocampus VTA/SN loop model and novelty signaling human DA neurons.** (A)  
4 Schematic of interactions and flow of novelty signals within the hippocampus-VTA/SN loop. Adopted  
5 from [2, 3, 108]. (B-D) Novelty signaling human DA neurons. (B) Example neuron that increases its firing  
6 when a stimulus is novel (blue) and decreases when the same stimulus is shown again (red). (C)  
7 Population summary. Novelty-sensitive DA neurons change their firing rate between the first (left) and  
8 second (middle) time the same image is seen in a continuous recognition memory task. Each dot is one  
9 neuron. (D) Analysis of extracellular waveforms of neurons recorded in the human SN indicates a  
10 population of wide-and narrow waveform neurons, which are putatively dopaminergic and GABAergic,  
11 respectively. Note that the novelty-signaling neurons (blue) had wide waveforms. (B-D) adjusted from  
12 [24]. Abbreviations: SN – substantia nigra, VTA – ventral tegmental area, PPTg - pedunculo-pontine  
13 tegmental nucleus.

14

15 **Figure 2: Sparse and selective coding of declarative memory content and reinstatement by single neurons.**

16 (A-C) Example of visually selective neurons. (A) Two visually selective neurons, one responding to many  
17 different images showing clothes (top, from hippocampus) and one only responding to a single image of  
18 a food item (bottom, from amygdala). Adapted from [43]. (B) Example of a highly invariant multimodal  
19 concept neuron that responds to images and written and spoken name of an experimenter, but not many  
20 other images (only examples are shown). Adapted from [44]. (C) Visually selective category neuron. Trials  
21 are ordered by visual category from which the images are chosen (all images shown are different).  
22 Stimulus on/offset is shown with dashed lines or a grey box (C). (D) Example memory-selective neuron.  
23 Note that the images shown during novel (green) and familiar (red) trials are the same. Adapted from  
24 [57]. (E) Contextual reinstatement during free recall by middle temporal gyrus neurons. Adapted from  
25 [81]. (F-G) Contextual reinstatement by VS neurons during recognition memory. Adapted from [77].

26

27 **Figure 3: Mnemonic evidence accumulation model and memory-choice cells.** (A) Mnemonic evidence

28 accumulation as a race process, with the choice made as the integrator (EV) that first reaches a preset  
29 threshold. (B) Anatomy of the PPC. Star marks the recording location for the data shown in (C-E). Adapted  
30 from [99]. (C) Example PPC neuron that increases its firing rate for a “new” decision in a graded manner  
31 modulated by confidence. Top shows the response of this neuron in individual trials, bottom the average

1 response of this neuron. (D) Population summary of memory-sensitive neurons in the PPC reveal that  
2 some neurons signal the confidence of the decision (green), whereas others only signal the new vs. old  
3 decision (red). (E) Group PSTH of memory-choice cells in PPC, grouped according to their preferred  
4 stimulus (new or old). The preference is defined according to ground truth. Note that during errors,  
5 neurons increase their firing rate for their preferred stimulus, indicating that they signal choices  
6 regardless of whether they are correct or not. Adapted from [98]. (F) The balance of evidence model.  
7 Note that this model integrates the difference (new-old or old-new), which is not the same due to  
8 rectification. (G-H) Cumulative firing rate of MS neurons, shown separately for high (G) and low (H)  
9 confidence trials. The balance of evidence  $\Delta EV$  scales as a function of confidence. Adapted from [58].

10

11 **Figure 4 (Key Figure): Summary of functional cell types and their putative interactions during recognition**  
12 **memory encoding and retrieval.** Summarized are four cell types (filled colored circles): visually selective  
13 (VS) cells, memory selective (MS) cells, novelty-sensitive dopamine neurons, and choice neurons. There  
14 are two types of MS cells: Novelty and familiarity selective (NS and FS). Arrows indicate direction of  
15 information flow, but do not indicate monosynaptic connections. Anatomical areas are indicated in  
16 dashed boxes. The theoretical concepts/models (gray boxes) discussed are: i) the Hippocampus-VTA/SN  
17 loop model, which proposes that hippocampal novelty signals excite dopamine neurons in the VTA/SN,  
18 which in turn release dopamine in the hippocampus, which leads to long-lasting plasticity. ii) the sparse  
19 coding memory model, which suggests that distinct neurons encode semantic and episodic aspects of  
20 declarative memories in a sparse but distributed manner. iii) the temporal context model, which suggests  
21 that the neural state present at encoding is re-instated at retrieval. iv) mnemonic accumulation, which  
22 suggests that neurons exist that integrate memory signals to make choices. v) the balance-of-evidence  
23 model, which describes how the difference between two mnemonic integrators can be used to make  
24 metacognitive confidence judgments about memories. Jointly, the body of single-neuron experiments  
25 discussed provides evidence for key aspects predicted by these models, including novelty cells that  
26 exhibit rapid single-trial learning, representations of memory strength that predict subjective confidence  
27 judgments, sparse coding of semantic memories, reinstatement of neural state by VS cells during  
28 retrieval, and representations of memory-based choices that are putative mnemonic accumulators.

29

### 30 Tables

31 **Table 1. Models of memory, their predictions and related human intracranial studies.**

Model class	Key predictions	Related studies
Hippocampus VTA/SN loop model	Dopamine neurons respond to novel stimuli, this novelty response is indicative of memory formation success, and exhibits single-trial learning.	[24, 27]
Sparse and separate encoding of semantic and episodic aspects of declarative memories	Highly visually selective cells (response sparsity), separate neurons encoding whether a stimulus has been seen before or not and other episodic aspects,	[35-39, 58]
Temporal context model, contextual reinstatement	Jump-back in time during retrieval, slowly drifting neural state, recollection/high confidence recognition predicted by reinstatement	[77, 78, 81] [79, 80]
Mnemonic evidence accumulation	Separate accumulators for “new” and “old” decisions. Accumulator output signal is modulated by memory strength. Accumulators predict choices rather than ground truth.	[97, 98]
Balance-of-evidence for metacognitive/confidence judgments	Magnitude of difference between the separate integrators for “new” and “old” choices is proportional to confidence, whereas the sign of the difference is indicative of the choice. Both decisions can be made at the same point of time.	[58]

1  
2

## Boxes

### **Box 1. Clinical procedures are a gold mine for cognitive neuroscience at the single-neuron level.**

In a limited number of circumstances, the activity of single neurons in awake behaving humans performing memory tasks can be recorded [109]. There are three separate clinical scenarios where such work has been performed. First, patients with drug-resistant focal epilepsy undergoing monitoring with depth electrodes [109, 110]. Recordings are performed during the 1-3 week-long stay of such patients in an epilepsy monitoring unit [111] (see Figure I). Second, patients undergoing awake brain surgery for implantation of a deep brain stimulation (DBS) device for treatment of the symptoms of movement disorders [23, 24] or psychiatric indications such as depression or OCD [112, 113]. Third, patients participating in brain machine interface trials with invasive electrodes, such as the Utah array [100].

Each of these approaches provides access to a different set of brain areas with different constraints. The strength of recording in epilepsy patients is ability to perform experiments in the relative comfort of the hospital room, performing experiments for several days, execution of relatively complex behavioral manipulations, and simultaneous recording of neurons from different brain areas. Limitations include restriction in accessible brain areas due to microwires exiting at the tip, inability to move electrodes, and caveats posed by the underlying seizure disorder. The strength of recording intra-operatively is ability to move the electrode to search for neurons and the ability to record anywhere along the track, which often includes areas not accessible in other settings such as the basal ganglia and striatum. Challenges of this approach include limited experiment time, impaired behavior due to after-effects of anesthesia and inability to perform complex behavioral procedures. The strength of Utah-array recordings is the ability to record large numbers of neurons from a small 4x4 mm patch of cortex for long periods of time during complex behavior. Limitations include inability to access areas not on the cortical surface and inability to move electrodes.

Notable, across approaches many cortical and subcortical areas of interest to models of memory can be accessed, including hippocampus, amygdala, dopamine neurons in the substantia nigra, posterior parietal cortex, and the frontal lobe. Indeed, single-neuron data that informs our understanding of human memory has already been obtained from all three clinical scenarios described (as summarized in this review), illustrating the power of utilizing several clinical scenarios to investigate the same scientific question.

**Figure I (place inside Box 1): Electrodes and recordings of single-neuron recordings in humans.** (A) Hybrid Depth electrode that is frequently used for recordings in epilepsy patients. Adapted from [114]. (B) Utah array. (C) Example raw recording (0.3-3kHz bandpass filtered) and two isolated clusters on a microwire of the kind shown in (A). Adapted from [111].

### **Box 2. The role of dopamine in declarative memory.**

Does dopamine have a role in the formation of hippocampal-dependent memories? While comparatively little studied in comparison to its role in reward, substantial evidence suggests that it does. First, the extent of DA neuron activity during viewing of novel stimuli correlates with the likelihood that the stimuli will later be remembered. This was originally shown using BOLD-fMRI with ROIs located in areas that contain, among others, DA neurons [115]. But since the BOLD signal is not cell-type selective, it is possible that the changes in BOLD signal observed were not due to changes in DA neuron activity. However, direct recordings from human DA neurons in the SN reveal that DA neurons increase their firing rate in response to novel stimuli and that the degree of this activation is indicative of whether the stimulus will later be remembered or forgotten [24]. Similarly, in macaques, DA neurons are activated by novel stimuli [116].

1 However, to our knowledge, DA neuron activation and its relationship to behavior has not been tested  
2 in a hippocampal-dependent memory tasks in macaques. Second, blocking or otherwise interfering with  
3 DA receptors in the hippocampus impairs declarative memory encoding [117, 118] and modulates  
4 synaptic plasticity as assessed by LTP/LTD [3]. Thirdly, genetic studies in humans show that  
5 polymorphisms in dopamine-related genes explain variance in declarative memory ability and/or results  
6 in differential activation measured by BOLD fMRI or EEG [119]. Together, this data suggests that  
7 dopamine release is critical for encoding new declarative memories. What remains poorly understood,  
8 however, is what activates dopamine neurons in a novelty-dependent manner. How this occurs is what  
9 the hippocampus-VTA/SN loop model is concerned with (see main text). A critical experiment that  
10 remains to be done is to directly measure levels of dopamine in the human hippocampus at a fast time  
11 scale to assess how transient the changes in dopamine are after exposure to novel stimuli and whether  
12 the increase in dopamine is selectively target anatomically or to different cell types. Among the potential  
13 approaches to pursue this question are fast micro dialysis and cyclic voltammetry, both of which have  
14 been utilized in humans for other purposes [110, 120].

## 15 16 17 Glossary

18  
19 **BOLD:** Blood oxygen level-dependent signal, measured using fMRI.

20  
21 **Confidence judgment:** A subjective assessment of the likelihood that a given decision was correct or not,  
22 typically assessed on a 3 to 10-point scale ranging from “guessing” to “very confident”.

23  
24 **Contiguity effect:** Temporal clustering of retrieval at the behavioral (the item retrieved next is most likely  
25 the one studied right before or after the previously retrieved item) and neuronal (reinstated context is  
26 most similar to that present close in time to the retrieved item) level.

27  
28 **Declarative memory:** Memories that can be brought into conscious awareness and that can be described  
29 and assessed verbally. Includes memories for past events (episodic) and facts (semantic memory).

30  
31 **Deep Brain Stimulation (DBS):** A clinical treatment for movement disorders that is also sometimes used  
32 to treat psychiatric disorders. Applies high-frequency extracellular stimulation, which is thought to inhibit  
33 neural activity in the target area.

34  
35 **Dopamine (DA):** A neuromodulator important for memory (see Box 2).

36  
37 **Electrocorticogram (ECoG):** We use the term ECoG only to refer to recordings performed with subdural  
38 strip or grid electrodes that do not penetrate the brain.

39  
40 **Engram:** The physical substrate of a specific memory.

41  
42 **Episodic memory:** Memories of personally experienced events.

43  
44 **Intracranial electroencephalography (iEEG):** We use the term iEEG to refer to recordings with low-  
45 impedance macro electrodes along the shaft of depth electrodes deep inside the brain (see Box 1).

- 1 **Leaky integrator:** An integrator with a decay rate such that inputs that rely further back in time have less  
2 influence on the current value than more recent inputs  
3
- 4 **Memory selective cells (MS cells):** Summary term for cells that respond differently as a function of  
5 whether a stimulus is novel (never seen before) or familiar (seen before at least once).  
6
- 7 **Recognition memory:** The ability to identify a previously seen stimulus as familiar.  
8
- 9 **Semantic memory:** Memories of concepts and facts about the world.  
10
- 11 **Substantia nigra (SN):** A part of the midbrain that contains cell bodies of dopamine neurons.  
12
- 13 **Ventral tegmental area (VTA):** A part of the midbrain that contains cell bodies of dopamine neurons.  
14
- 15 **Visually selective cells (VS cells):** Summary term that includes all variants of category-and concept cells.  
16



## 1 References cited

- 2 1. Squire, L.R., Stark, C.E., and Clark, R.E. (2004). The medial temporal lobe. *Annu Rev*  
3 *Neurosci* 27, 279-306.
- 4 2. Lisman, J.E., and Grace, A.A. (2005). The hippocampal-VTA loop: controlling the entry of  
5 information into long-term memory. *Neuron* 46, 703-713.
- 6 3. Lisman, J., Grace, A.A., and Duzel, E. (2011). A neoHebbian framework for episodic memory;  
7 role of dopamine-dependent late LTP. *Trends Neurosci* 34, 536-547.
- 8 4. Howard, M.W. (2018). Memory as Perception of the Past: Compressed Time in Mind and Brain.  
9 *Trends Cogn Sci* 22, 124-136.
- 10 5. Hasselmo, M.E., Bodelon, C., and Wyble, B.P. (2002). A proposed function for hippocampal  
11 theta rhythm: separate phases of encoding and retrieval enhance reversal of prior learning.  
12 *Neural Comput* 14, 793-817.
- 13 6. Wixted, J.T. (2007). Dual-process theory and signal-detection theory of recognition memory.  
14 *Psychol Rev* 114, 152-176.
- 15 7. Kepecs, A., and Mainen, Z.F. (2012). A computational framework for the study of confidence  
16 in humans and animals. *Philos Trans R Soc Lond B Biol Sci* 367, 1322-1337.
- 17 8. Marr, D. (1971). Simple memory: a theory for archicortex. *Philos Trans R Soc Lond B Biol Sci*  
18 262, 23-81.
- 19 9. McClelland, J.L., McNaughton, B.L., and O'Reilly, R.C. (1995). Why There Are Complementary  
20 Learning-Systems in the Hippocampus and Neocortex - Insights from the Successes and  
21 Failures of Connectionist Models of Learning and Memory. *Psychological Review* 102, 419-  
22 457.
- 23 10. Norman, K.A., and O'Reilly, R.C. (2003). Modeling hippocampal and neocortical contributions  
24 to recognition memory: a complementary-learning-systems approach. *Psychol Rev* 110, 611-  
25 646.
- 26 11. Solomon, E.A., Kragel, J.E., Sperling, M.R., Sharan, A., Worrell, G., Kucewicz, M., Inman,  
27 C.S., Lega, B., Davis, K.A., Stein, J.M., et al. (2017). Widespread theta synchrony and high-  
28 frequency desynchronization underlies enhanced cognition. *Nat Commun* 8, 1704.
- 29 12. Kahana, M.J., Sekuler, R., Caplan, J.B., Kirschen, M., and Madsen, J.R. (1999). Human theta  
30 oscillations exhibit task dependence during virtual maze navigation. *Nature* 399, 781-784.
- 31 13. Raghavachari, S., Kahana, M.J., Rizzuto, D.S., Caplan, J.B., Kirschen, M.P., Bourgeois, B.,  
32 Madsen, J.R., and Lisman, J.E. (2001). Gating of human theta oscillations by a working  
33 memory task. *J Neurosci* 21, 3175-3183.
- 34 14. Johnson, E.L., and Knight, R.T. (2015). Intracranial recordings and human memory. *Curr Opin*  
35 *Neurobiol* 31, 18-25.
- 36 15. Standing, L., Conezio, J., and Haber, R.N. (1970). Perception and Memory for Pictures -  
37 Single-Trial Learning of 2500 Visual Stimuli. *Psychonomic Science* 19, 73-74.
- 38 16. Lisman, J.E., and Otmakhova, N.A. (2001). Storage, recall, and novelty detection of sequences  
39 by the hippocampus: elaborating on the SOCRATIC model to account for normal and aberrant  
40 effects of dopamine. *Hippocampus* 11, 551-568.
- 41 17. Tamura, K., Takeda, M., Setsuie, R., Tsubota, T., Hirabayashi, T., Miyamoto, K., and  
42 Miyashita, Y. (2017). Conversion of object identity to object-general semantic value in the  
43 primate temporal cortex. *Science* 357, 687-692.
- 44 18. Xiang, J.Z., and Brown, M.W. (1998). Differential neuronal encoding of novelty, familiarity  
45 and recency in regions of the anterior temporal lobe. *Neuropharmacology* 37, 657-676.
- 46 19. Knight, R. (1996). Contribution of human hippocampal region to novelty detection. *Nature* 383,  
47 256-259.

- 1 20. Rutishauser, U., Mamelak, A.N., and Schuman, E.M. (2006). Single-trial learning of novel  
2 stimuli by individual neurons of the human hippocampus-amygdala complex. *Neuron* 49, 805-  
3 813.
- 4 21. Viskontas, I.V., Knowlton, B.J., Steinmetz, P.N., and Fried, I. (2006). Differences in mnemonic  
5 processing by neurons in the human hippocampus and parahippocampal regions. *J Cognitive*  
6 *Neurosci* 18, 1654-1662.
- 7 22. Pedreira, C., Mormann, F., Kraskov, A., Cerf, M., Fried, I., Koch, C., and Quiroga, R.Q.  
8 (2010). Responses of human medial temporal lobe neurons are modulated by stimulus  
9 repetition. *J Neurophysiol* 103, 97-107.
- 10 23. Zaghoul, K.A., Blanco, J.A., Weidemann, C.T., McGill, K., Jaggi, J.L., Baltuch, G.H., and  
11 Kahana, M.J. (2009). Human substantia nigra neurons encode unexpected financial rewards.  
12 *Science* 323, 1496-1499.
- 13 24. Kaminski, J., Mamelak, A.N., Birch, K., Mosher, C.P., Tagliati, M., and Rutishauser, U. (2018).  
14 Novelty-Sensitive Dopaminergic Neurons in the Human Substantia Nigra Predict Success of  
15 Declarative Memory Formation. *Curr Biol* 28, 1333-1343 e1334.
- 16 25. Ramayya, A.G., Zaghoul, K.A., Weidemann, C.T., Baltuch, G.H., and Kahana, M.J. (2014).  
17 Electrophysiological evidence for functionally distinct neuronal populations in the human  
18 substantia nigra. *Front Hum Neurosci* 8, 655.
- 19 26. Fearnley, J.M., and Lees, A.J. (1991). Ageing and Parkinson's disease: substantia nigra regional  
20 selectivity. *Brain* 114 ( Pt 5), 2283-2301.
- 21 27. Mikell, C.B., Sheehy, J.P., Youngerman, B.E., McGovern, R.A., Wojtasiewicz, T.J., Chan,  
22 A.K., Pullman, S.L., Yu, Q., Goodman, R.R., Schevon, C.A., et al. (2014). Features and timing  
23 of the response of single neurons to novelty in the substantia nigra. *Brain Res* 1542, 79-84.
- 24 28. Squire, L.R. (2004). Memory systems of the brain: a brief history and current perspective.  
25 *Neurobiol Learn Mem* 82, 171-177.
- 26 29. Rolls, E.T., and Treves, A. (2011). The neuronal encoding of information in the brain. *Prog*  
27 *Neurobiol* 95, 448-490.
- 28 30. Wixted, J.T., Squire, L.R., Jang, Y., Papesh, M.H., Goldinger, S.D., Kuhn, J.R., Smith, K.A.,  
29 Treiman, D.M., and Steinmetz, P.N. (2014). Sparse and distributed coding of episodic memory  
30 in neurons of the human hippocampus. *Proc Natl Acad Sci U S A* 111, 9621-9626.
- 31 31. Rogers, T.T., and McClelland, J.L. (2014). Parallel Distributed Processing at 25: further  
32 explorations in the microstructure of cognition. *Cogn Sci* 38, 1024-1077.
- 33 32. Rumelhart, D.E., McClelland, J.L., and PDP Research Group (1986). Parallel distributed  
34 processing : explorations in the microstructure of cognition, (Cambridge, Mass.: MIT Press).
- 35 33. Kreiman, G. (2017). A null model for cortical representations with grandmothers galore. *Lang*  
36 *Cogn Neurosci* 32, 274-285.
- 37 34. Manns, J.R., Hopkins, R.O., and Squire, L.R. (2003). Semantic memory and the human  
38 hippocampus. *Neuron* 38, 127-133.
- 39 35. Mormann, F., Kornblith, S., Quiroga, R.Q., Kraskov, A., Cerf, M., Fried, I., and Koch, C.  
40 (2008). Latency and selectivity of single neurons indicate hierarchical processing in the human  
41 medial temporal lobe. *Journal of Neuroscience* 28, 8865-8872.
- 42 36. Kreiman, G., Koch, C., and Fried, I. (2000). Category-specific visual responses of single  
43 neurons in the human medial temporal lobe. *Nat Neurosci* 3, 946-953.
- 44 37. Quiroga, R.Q., Reddy, L., Kreiman, G., Koch, C., and Fried, I. (2005). Invariant visual  
45 representation by single neurons in the human brain. *Nature* 435, 1102-1107.
- 46 38. Waydo, S., Kraskov, A., Quiroga, R., Fried, I., and Koch, C. (2006). Sparse  
47 representation in the human medial temporal lobe. *J Neurosci* 26, 10232-10234.

- 1 39. Kaminski, J., Sullivan, S., Chung, J.M., Ross, I.B., Mamelak, A.N., and Rutishauser, U. (2017).  
2 Persistently active neurons in human medial frontal and medial temporal lobe support working  
3 memory. *Nat Neurosci* 20, 590-601.
- 4 40. Coltheart, M. (2017). Grandmother cells and the distinction between local and distributed  
5 representation. *Lang Cogn Neurosci* 32, 350-358.
- 6 41. Quiroga, R.Q. (2012). Concept cells: the building blocks of declarative memory functions. *Nat*  
7 *Rev Neurosci* 13, 587-597.
- 8 42. Reber, T.P., Faber, J., Niediek, J., Bostrom, J., Elger, C.E., and Mormann, F. (2017). Single-  
9 Neuron Correlates of Conscious Perception in the Human Medial Temporal Lobe. *Curr Biol* 27,  
10 2991-2998 e2992.
- 11 43. Reber, T.P., Bausch, M., Mackay, S., Bostrom, J., Elger, C.E., and Mormann, F. (2019).  
12 Representation of Abstract Semantic Knowledge in Populations of Human Single Neurons in  
13 the Medial Temporal Lobe. *PLoS Biol* *in press*.
- 14 44. Quiroga, R., Kraskov, A., Koch, C., and Fried, I. (2009). Explicit encoding of  
15 multimodal percepts by single neurons in the human brain. *Curr Biol* 19, 1308-1313.
- 16 45. Cerf, M., Thiruvengadam, N., Mormann, F., Kraskov, A., Quiroga, R.Q., Koch, C., and Fried, I.  
17 (2010). On-line, voluntary control of human temporal lobe neurons. *Nature* 467, 1104-1108.
- 18 46. Gelbard-Sagiv, H., Mukamel, R., Harel, M., Malach, R., and Fried, I. (2008). Internally  
19 generated reactivation of single neurons in human hippocampus during free recall. *Science* 322,  
20 96-101.
- 21 47. Kornblith, S., Quiroga, R., Koch, C., Fried, I., and Mormann, F. (2017). Persistent  
22 Single-Neuron Activity during Working Memory in the Human Medial Temporal Lobe. *Curr*  
23 *Biol* 27, 1026-1032.
- 24 48. Kreiman, G., Koch, C., and Fried, I. (2000). Imagery neurons in the human brain. *Nature* 408,  
25 357-361.
- 26 49. Kreiman, G., Fried, I., and Koch, C. (2002). Single-neuron correlates of subjective vision in the  
27 human medial temporal lobe. *Proc Natl Acad Sci U S A* 99, 8378-8383.
- 28 50. Nir, Y., Andrillon, T., Marmelshtein, A., Suthana, N., Cirelli, C., Tononi, G., and Fried, I.  
29 (2017). Selective neuronal lapses precede human cognitive lapses following sleep deprivation.  
30 *Nat Med* 23, 1474-1480.
- 31 51. Wang, S., Yu, R., Tyszka, J.M., Zhen, S., Kovach, C., Sun, S., Huang, Y., Hurlemann, R., Ross,  
32 I.B., Chung, J.M., et al. (2017). The human amygdala parametrically encodes the intensity of  
33 specific facial emotions and their categorical ambiguity. *Nat Commun* 8, 14821.
- 34 52. Wang, S., Tudusciuc, O., Mamelak, A.N., Ross, I.B., Adolphs, R., and Rutishauser, U. (2014).  
35 Neurons in the human amygdala selective for perceived emotion. *Proc Natl Acad Sci U S A*  
36 111, E3110-3119.
- 37 53. Viskontas, I.V., Quiroga, R.Q., and Fried, I. (2009). Human medial temporal lobe neurons  
38 respond preferentially to personally relevant images. *Proc Natl Acad Sci U S A* 106, 21329-  
39 21334.
- 40 54. De Falco, E., Ison, M.J., Fried, I., and Quiroga, R. (2016). Long-term coding of personal  
41 and universal associations underlying the memory web in the human brain. *Nat Commun* 7,  
42 13408.
- 43 55. Ison, M.J., Quiroga, R., and Fried, I. (2015). Rapid Encoding of New Memories by  
44 Individual Neurons in the Human Brain. *Neuron* 87, 220-230.
- 45 56. Josselyn, S.A., Kohler, S., and Frankland, P.W. (2015). Finding the engram. *Nat Rev Neurosci*  
46 16, 521-534.
- 47 57. Faraut, M.C.M., Carlson, A.A., Sullivan, S., Tudusciuc, O., Ross, I., Reed, C.M., Chung, J.M.,  
48 Mamelak, A.N., and Rutishauser, U. (2018). Dataset of human medial temporal lobe single  
49 neuron activity during declarative memory encoding and recognition. *Sci Data* 5, 180010.

- 1 58. Rutishauser, U., Ye, S., Koroma, M., Tudusciuc, O., Ross, I.B., Chung, J.M., and Mamelak,  
2 A.N. (2015). Representation of retrieval confidence by single neurons in the human medial  
3 temporal lobe. *Nature Neuroscience* 18, 1041-1050.
- 4 59. Rutishauser, U., Schuman, E.M., and Mamelak, A.N. (2008). Activity of human hippocampal  
5 and amygdala neurons during retrieval of declarative memories. *Proc Natl Acad Sci U S A* 105,  
6 329-334.
- 7 60. Minxha, J., Mosher, C., Morrow, J.K., Mamelak, A.N., Adolphs, R., Gothard, K.M., and  
8 Rutishauser, U. (2017). Fixations gate species-specific responses to free viewing of faces in the  
9 human and macaque amygdala. *Cell Reports* 18, 878-891.
- 10 61. Kirchner, H., and Thorpe, S.J. (2006). Ultra-rapid object detection with saccadic eye  
11 movements: visual processing speed revisited. *Vision Res* 46, 1762-1776.
- 12 62. Murray, E.A., Bussey, T.J., and Saksida, L.M. (2007). Visual perception and memory: a new  
13 view of medial temporal lobe function in primates and rodents. *Annu Rev Neurosci* 30, 99-122.
- 14 63. Landi, S.M., and Freiwald, W.A. (2017). Two areas for familiar face recognition in the primate  
15 brain. *Science* 357, 591-595.
- 16 64. Buffalo, E.A., Reber, P.J., and Squire, L.R. (1998). The human perirhinal cortex and  
17 recognition memory. *Hippocampus* 8, 330-339.
- 18 65. Naya, Y., Yoshida, M., and Miyashita, Y. (2001). Backward spreading of memory-retrieval  
19 signal in the primate temporal cortex. *Science* 291, 661-664.
- 20 66. Squire, L.R., Zola-Morgan, J.T., and Clark, R.E. (2007). Recognition memory and the medial  
21 temporal lobe: a new perspective. *Nat Rev Neurosci* 8, 872-883.
- 22 67. Brown, M.W., and Aggleton, J.P. (2001). Recognition memory: what are the roles of the  
23 perirhinal cortex and hippocampus? *Nat Rev Neurosci* 2, 51-61.
- 24 68. Miller, J.F., Lazarus, E.M., Polyn, S.M., and Kahana, M.J. (2013). Spatial clustering during  
25 memory search. *J Exp Psychol Learn Mem Cogn* 39, 773-781.
- 26 69. Kahana, M.J. (2012). *Foundations of human memory*, (New York: Oxford University Press).
- 27 70. Polyn, S.M., Norman, K.A., and Kahana, M.J. (2009). A Context Maintenance and Retrieval  
28 Model of Organizational Processes in Free Recall. *Psychological Review* 116, 129-156.
- 29 71. Howard, M.W., Fotedar, M.S., Datey, A.V., and Hasselmo, M.E. (2005). The temporal context  
30 model in spatial navigation and relational learning: toward a common explanation of medial  
31 temporal lobe function across domains. *Psychol Rev* 112, 75-116.
- 32 72. Howard, M.W., and Kahana, M.J. (2002). A distributed representation of temporal context. *J*  
33 *Math Psychol* 46, 269-299.
- 34 73. Sederberg, P.B., Howard, M.W., and Kahana, M.J. (2008). A context-based theory of recency  
35 and contiguity in free recall. *Psychol Rev* 115, 893-912.
- 36 74. Tulving, E. (2002). Episodic memory: From mind to brain. *Annual Review of Psychology* 53,  
37 1-25.
- 38 75. Yonelinas, A.P. (2001). Components of episodic memory: the contribution of recollection and  
39 familiarity. *Philos Trans R Soc Lond B* 356, 1363-1374.
- 40 76. Palombo, D.J., Di Lascio, J.M., Howard, M.W., and Verfaellie, M. (2018). Medial Temporal  
41 Lobe Amnesia Is Associated with a Deficit in Recovering Temporal Context. *J Cogn Neurosci*,  
42 1-13.
- 43 77. Folkerts, S., Rutishauser, U., and Howard, M.W. (2018). Human Episodic Memory Retrieval Is  
44 Accompanied by a Neural Contiguity Effect. *J Neurosci* 38, 4200-4211.
- 45 78. Howard, M.W., Viskontas, I.V., Shankar, K.H., and Fried, I. (2012). Ensembles of human MTL  
46 neurons "jump back in time" in response to a repeated stimulus. *Hippocampus* 22, 1833-1847.
- 47 79. Manning, J.R., Polyn, S.M., Baltuch, G.H., Litt, B., and Kahana, M.J. (2011). Oscillatory  
48 patterns in temporal lobe reveal context reinstatement during memory search. *Proc Natl Acad*  
49 *Sci U S A* 108, 12893-12897.



- 1 80. Yaffe, R.B., Kerr, M.S., Damera, S., Sarma, S.V., Inati, S.K., and Zaghoul, K.A. (2014).  
2 Reinstatement of distributed cortical oscillations occurs with precise spatiotemporal dynamics  
3 during successful memory retrieval. *Proc Natl Acad Sci U S A* *111*, 18727-18732.
- 4 81. Jang, A.I., Wittig, J.H., Jr., Inati, S.K., and Zaghoul, K.A. (2017). Human Cortical Neurons in  
5 the Anterior Temporal Lobe Reinstatement Spiking Activity during Verbal Memory Retrieval. *Curr*  
6 *Biol* *27*, 1700-1705 e1705.
- 7 82. Miller, J.F., Neufang, M., Solway, A., Brandt, A., Trippel, M., Mader, I., Hefft, S., Merkow,  
8 M., Polyn, S.M., Jacobs, J., et al. (2013). Neural activity in human hippocampal formation  
9 reveals the spatial context of retrieved memories. *Science* *342*, 1111-1114.
- 10 83. Tompary, A., Duncan, K., and Davachi, L. (2016). High-resolution investigation of memory-  
11 specific reinstatement in the hippocampus and perirhinal cortex. *Hippocampus* *26*, 995-1007.
- 12 84. Gold, J.I., and Shadlen, M.N. (2007). The neural basis of decision making. *Annu Rev Neurosci*  
13 *30*, 535-574.
- 14 85. Usher, M., and McClelland, J.L. (2001). The time course of perceptual choice: the leaky,  
15 competing accumulator model. *Psychol Rev* *108*, 550-592.
- 16 86. Ratcliff, R., Smith, P.L., Brown, S.D., and McKoon, G. (2016). Diffusion Decision Model:  
17 Current Issues and History. *Trends Cogn Sci* *20*, 260-281.
- 18 87. Kiani, R., and Shadlen, M.N. (2009). Representation of confidence associated with a decision  
19 by neurons in the parietal cortex. *Science* *324*, 759-764.
- 20 88. Shadlen, M.N., and Newsome, W.T. (2001). Neural basis of a perceptual decision in the parietal  
21 cortex (area LIP) of the rhesus monkey. *J Neurophysiol* *86*, 1916-1936.
- 22 89. Schall, J.D. (2003). Neural correlates of decision processes: neural and mental chronometry.  
23 *Curr Opin Neurobiol* *13*, 182-186.
- 24 90. Hanks, T.D., Kopec, C.D., Brunton, B.W., Duan, C.A., Erlich, J.C., and Brody, C.D. (2015).  
25 Distinct relationships of parietal and prefrontal cortices to evidence accumulation. *Nature* *520*,  
26 220-223.
- 27 91. Wagner, A.D., Shannon, B.J., Kahn, I., and Buckner, R.L. (2005). Parietal lobe contributions to  
28 episodic memory retrieval. *Trends Cogn Sci* *9*, 445-453.
- 29 92. Sestieri, C., Shulman, G.L., and Corbetta, M. (2017). The contribution of the human posterior  
30 parietal cortex to episodic memory. *Nat Rev Neurosci* *18*, 183-192.
- 31 93. Davis, S.W., Wing, E.A., and Cabeza, R. (2018). Chapter 27 - Contributions of the ventral  
32 parietal cortex to declarative memory. In *Handbook of Clinical Neurology*, Volume 151, G.  
33 Vallar and H.B. Coslett, eds. (Elsevier), pp. 525-553.
- 34 94. Rugg, M.D., and Curran, T. (2007). Event-related potentials and recognition memory. *Trends*  
35 *Cogn Sci* *11*, 251-257.
- 36 95. Ben-Zvi, S., Soroker, N., and Levy, D.A. (2015). Parietal lesion effects on cued recall following  
37 pair associate learning. *Neuropsychologia* *73*, 176-194.
- 38 96. Hutchinson, J.B., Uncapher, M.R., Weiner, K.S., Bressler, D.W., Silver, M.A., Preston, A.R.,  
39 and Wagner, A.D. (2014). Functional heterogeneity in posterior parietal cortex across attention  
40 and episodic memory retrieval. *Cereb Cortex* *24*, 49-66.
- 41 97. Gonzalez, A., Hutchinson, J.B., Uncapher, M.R., Chen, J., LaRocque, K.F., Foster, B.L.,  
42 Rangarajan, V., Parvizi, J., and Wagner, A.D. (2015). Electrocorticography reveals the temporal  
43 dynamics of posterior parietal cortical activity during recognition memory decisions. *Proc Natl*  
44 *Acad Sci U S A* *112*, 11066-11071.
- 45 98. Rutishauser, U., Aflalo, T., Rosario, E.R., Pouratian, N., and Andersen, R.A. (2018). Single-  
46 Neuron Representation of Memory Strength and Recognition Confidence in Left Human  
47 Posterior Parietal Cortex. *Neuron* *97*, 209-220 e203.
- 48 99. Parvizi, J., and Wagner, A.D. (2018). Memory, Numbers, and Action Decision in Human  
49 Posterior Parietal Cortex. *Neuron* *97*, 7-10.

- 1 100. Aflalo, T., Kellis, S., Klaes, C., Lee, B., Shi, Y., Pejsa, K., Shanfield, K., Hayes-Jackson, S.,  
2 Aisen, M., Heck, C., et al. (2015). Neurophysiology. Decoding motor imagery from the  
3 posterior parietal cortex of a tetraplegic human. *Science* 348, 906-910.
- 4 101. Foster, B.L., Kaveh, A., Dastjerdi, M., Miller, K.J., and Parvizi, J. (2013). Human retrosplenial  
5 cortex displays transient theta phase locking with medial temporal cortex prior to activation  
6 during autobiographical memory retrieval. *J Neurosci* 33, 10439-10446.
- 7 102. Pouget, A., Drugowitsch, J., and Kepecs, A. (2016). Confidence and certainty: distinct  
8 probabilistic quantities for different goals. *Nat Neurosci* 19, 366-374.
- 9 103. Yeung, N., and Summerfield, C. (2012). Metacognition in human decision-making: confidence  
10 and error monitoring. *Philos Trans R Soc Lond B Biol Sci* 367, 1310-1321.
- 11 104. Metcalfe, J. (2008). Evolution of Metacognition In *Handbook of Metamemory and Memory*, J.  
12 Dunlosky and R. Bjork, eds. (New York: Psychology Press), pp. 29-46.
- 13 105. Vickers, D. (1979). *Decision processes in visual perception*, (New York ; London: Academic  
14 Press).
- 15 106. Bogacz, R., Brown, E., Moehlis, J., Holmes, P., and Cohen, J.D. (2006). The physics of optimal  
16 decision making: a formal analysis of models of performance in two-alternative forced-choice  
17 tasks. *Psychol Rev* 113, 700-765.
- 18 107. Navajas, J., Bahrami, B., and Latham, P.E. (2016). Post-decisional accounts of biases in  
19 confidence. *Curr. Opin. Behav. Sci.* 11, 55-60.
- 20 108. Hansen, N., and Manahan-Vaughan, D. (2014). Dopamine D1/D5 receptors mediate  
21 informational saliency that promotes persistent hippocampal long-term plasticity. *Cereb Cortex*  
22 24, 845-858.
- 23 109. Fried, I., Rutishauser, U., Cerf, M., and Kreiman, G. (2014). *Single Neuron Studies of the*  
24 *Human Brain: Probing Cognition*, (Boston: MIT Press).
- 25 110. Fried, I., Wilson, C.L., Maidment, N.T., Engel, J., Behnke, E., Fields, T.A., MacDonald, K.A.,  
26 Morrow, J.W., and Ackerson, L. (1999). Cerebral microdialysis combined with single-neuron  
27 and electroencephalographic recording in neurosurgical patients - Technical note. *Journal of*  
28 *Neurosurgery* 91, 697-705.
- 29 111. Minxha, J., Mamelak, A.N., and Rutishauser, U. (2018). Surgical and Electrophysiological  
30 Techniques for Single-Neuron Recordings in Human Epilepsy Patients. In *Extracellular*  
31 *Recording Approaches*. (Springer), pp. 267-293.
- 32 112. Sheth, S.A., Mian, M.K., Patel, S.R., Asaad, W.F., Williams, Z.M., Dougherty, D.D., Bush, G.,  
33 and Eskandar, E.N. (2012). Human dorsal anterior cingulate cortex neurons mediate ongoing  
34 behavioural adaptation. *Nature* 488, 218-221.
- 35 113. Bari, A.A., Mikell, C.B., Abosch, A., Ben-Haim, S., Buchanan, R.J., Burton, A.W., Carcieri, S.,  
36 Cosgrove, G.R., D'Haese, P.F., Daskalakis, Z.J., et al. (2018). Charting the road forward in  
37 psychiatric neurosurgery: proceedings of the 2016 American Society for Stereotactic and  
38 Functional Neurosurgery workshop on neuromodulation for psychiatric disorders. *J Neurol*  
39 *Neurosurg Psychiatry* 89, 886-896.
- 40 114. Fu, Z., Wu, D.J., Ross, I., Chung, J.M., Mamelak, A.N., Adolphs, R., and Rutishauser, U.  
41 (2019). Single-Neuron Correlates of Error Monitoring and Post-Error Adjustments in Human  
42 Medial Frontal Cortex. *Neuron* 101, 165-177 e165.
- 43 115. Wittmann, B.C., Schott, B.H., Guderian, S., Frey, J.U., Heinze, H.J., and Duzel, E. (2005).  
44 Reward-related fMRI activation of dopaminergic midbrain is associated with enhanced  
45 hippocampus-dependent long-term memory formation. *Neuron* 45, 459-467.
- 46 116. Ljungberg, T., Apicella, P., and Schultz, W. (1992). Responses of monkey dopamine neurons  
47 during learning of behavioral reactions. *J Neurophysiol* 67, 145-163.

- 1 117. Clausen, B., Schachtman, T.R., Mark, L.T., Reinholdt, M., and Christoffersen, G.R. (2011).  
2 Impairments of exploration and memory after systemic or prelimbic D1-receptor antagonism in  
3 rats. *Behav Brain Res* 223, 241-254.
- 4 118. Rosen, Z.B., Cheung, S., and Siegelbaum, S.A. (2015). Midbrain dopamine neurons  
5 bidirectionally regulate CA3-CA1 synaptic drive. *Nat Neurosci* 18, 1763-1771.
- 6 119. Medrano, P., Nyhus, E., Smolen, A., Curran, T., and Ross, R.S. (2017). Individual differences  
7 in EEG correlates of recognition memory due to DAT polymorphisms. *Brain and Behavior* 7.
- 8 120. Kishida, K.T., Saez, I., Lohrenz, T., Witcher, M.R., Laxton, A.W., Tatter, S.B., White, J.P.,  
9 Ellis, T.L., Phillips, P.E.M., and Montague, P.R. (2016). Subsecond dopamine fluctuations in  
10 human striatum encode superposed error signals about actual and counterfactual reward. *P Natl*  
11 *Acad Sci USA* 113, 200-205.  
12











