

Cognition and Behavior

# Traces of Semantization, from Episodic to Semantic Memory in a Spiking Cortical Network Model

Nikolaos Chrysanthis,<sup>1</sup> Florian Fiebig,<sup>1</sup> Anders Lansner,<sup>1,2</sup> and Pawel Herman<sup>1,3,4</sup><https://doi.org/10.1523/ENEURO.0062-22.2022>

<sup>1</sup>Division of Computational Science and Technology, School of Electrical Engineering and Computer Science, KTH Royal Institute of Technology, Stockholm 10044, Sweden, <sup>2</sup>Department of Mathematics, Stockholm University, Stockholm 10691, Sweden, <sup>3</sup>Digital Futures, Stockholm 10044, Sweden, and <sup>4</sup>Swedish e-Science Research Centre, Stockholm 10044, Sweden

## Abstract

Episodic memory is a recollection of past personal experiences associated with particular times and places. This kind of memory is commonly subject to loss of contextual information or “semantization,” which gradually decouples the encoded memory items from their associated contexts while transforming them into semantic or gist-like representations. Novel extensions to the classical Remember/Know (R/K) behavioral paradigm attribute the loss of episodicity to multiple exposures of an item in different contexts. Despite recent advancements explaining semantization at a behavioral level, the underlying neural mechanisms remain poorly understood. In this study, we suggest and evaluate a novel hypothesis proposing that Bayesian–Hebbian synaptic plasticity mechanisms might cause semantization of episodic memory. We implement a cortical spiking neural network model with a Bayesian–Hebbian learning rule called Bayesian Confidence Propagation Neural Network (BCPNN), which captures the semantization phenomenon and offers a mechanistic explanation for it. Encoding items across multiple contexts leads to item–context decoupling akin to semantization. We compare BCPNN plasticity with the more commonly used spike-timing-dependent plasticity (STDP) learning rule in the same episodic memory task. Unlike BCPNN, STDP does not explain the decontextualization process. We further examine how selective plasticity modulation of isolated salient events may enhance preferential retention and resistance to semantization. Our model reproduces important features of episodicity on behavioral timescales under various biological constraints while also offering a novel neural and synaptic explanation for semantization, thereby casting new light on the interplay between episodic and semantic memory processes.

**Key words:** Bayesian–Hebbian plasticity; BCPNN; episodic memory; semantization; spiking cortical memory model; STDP

## Significance Statement

Remembering single episodes is a fundamental attribute of cognition. Difficulties recollecting contextual information is a key sign of episodic memory loss or semantization. Behavioral studies demonstrate that semantization of episodic memory can occur rapidly, yet the neural mechanisms underlying this effect are insufficiently investigated. In line with recent behavioral findings, we show that multiple stimulus exposures in different contexts may advance item–context decoupling. We suggest a Bayesian–Hebbian synaptic plasticity hypothesis of memory semantization and further show that a transient modulation of plasticity during salient events may disrupt the decontextualization process by strengthening memory traces, and thus, enhancing preferential retention. The proposed cortical network-of-networks model thus bridges micro and mesoscale synaptic effects with network dynamics and behavior.

Received February 6, 2022; accepted May 28, 2022; First published July 8, 2022.

The authors declare no competing financial interests.

Author contributions: N.C., A.L., and P.H. designed research; N.C. performed research; N.C. and F.F. contributed unpublished reagents/analytic tools; N.C. analyzed data; N.C. wrote the original draft; N.C., F.F., A.L., and P.H. wrote the paper.

## Introduction

Episodic and semantic memory were originally proposed as distinct systems that compete in retrieval (Tulving, 1972). More recent studies suggest, however, that this division is rather vague (McCloskey and Santee, 1981; Howard and Kahana, 2002; Renoult et al., 2019), as neural correlates of episodic and semantic retrieval overlap (Weidemann et al., 2019). Episodic memory traces are susceptible to transformation and loss of information (Tulving, 1972), and this loss of episodicity can be attributed to semantization, which typically takes the form of a decontextualization process (Viard et al., 2007; Habermas et al., 2013; Duff et al., 2020). Baddeley (1988) hypothesized that semantic memory might represent the accumulated residue of multiple learning episodes, consisting of information which has been semantitized and detached from the associated episodic contextual detail. Extensions of the classical Remember/Know (R/K) behavioral experiment demonstrated that item-context decoupling can occur rapidly (Opitz, 2010). In these experiments, items were presented either in a unique context, or across several contexts. Low context variability improved the recollection rate, whereas context overload led to decontextualization and “Know” type of responses, i.e., recognition of item-only information without any detail about episodic context (Opitz, 2010; Smith and Manzano, 2010; Smith and Handy, 2014). To the best of our knowledge, there have not been any computational hypotheses proposed to offer mechanistic insights into this item-context decoupling effect.

Several computational spiking neural network models of cortical associative memory have previously been developed and used to investigate mechanisms underlying working memory maintenance and recall (Lundqvist et al., 2010, 2011; Herman et al., 2013). A similar model enhanced with a Bayesian–Hebbian learning rule (Bayesian Confidence Propagation Neural Network; BCPNN) representing synaptic and intrinsic plasticity was then used to study one-shot memory encoding (Fiebig and Lansner, 2017), and more recently, it was extended into a multinet-work cortical model to examine a novel “indexing theory” of working memory (Fiebig et al., 2020).

In the present study, relying on a similar spiking neural network model with identical modular architecture we propose and evaluate a Bayesian–Hebbian hypothesis about synaptic and network mechanisms underlying memory semantization and qualitatively match model output to available behavioral data. We show that associative binding between items and contexts becomes

This work was supported by the Swedish Research Council Grant 2018-05360. The simulations were enabled by resources provided by Swedish National Infrastructure for Computing (SNIC) at the PDC Center for High Performance Computing, KTH Royal Institute of Technology.

Correspondence should be addressed to Pawel Herman at [paherman@kth.se](mailto:paherman@kth.se).  
<https://doi.org/10.1523/ENEURO.0062-22.2022>

Copyright © 2022 Chrysanthidis et al.

This is an open-access article distributed under the terms of the Creative Commons Attribution 4.0 International license, which permits unrestricted use, distribution and reproduction in any medium provided that the original work is properly attributed.

weaker when an item is presented across multiple contexts (high context variability). This gradual trace transformation relies on the nature of Bayesian learning, which normalizes and updates weights over estimated presynaptic (Bayesian-prior) as well as postsynaptic (Bayesian-posterior) spiking activity. We compare these findings with an analogous model that features the more well-known spike-timing-dependent plasticity (STDP) instead of the BCPNN learning rule, and demonstrate that no memory semantization effect can be reproduced, regardless of the degree of context variability. Notably, there have been earlier modeling attempts at semantization using STDP or other learning rules but this memory phenomenon has been interpreted differently involving slow memory consolidation (requiring sleep, repeated exposures, or systems consolidation) or extraction of semantic relations (a.k.a. prototype learning) among a group of episodic memories sharing statistical similarities (Deperrois et al., 2021; Remme et al., 2021). We argue that our hypothesis is more generic as it does not assume any statistical structure of the memory object representations. Finally, we also show how selective plasticity neuromodulation of one-shot learning (tentatively modeling effects of attention, emotional salience, and surprise on plasticity) may delay or prevent decontextualization.

In contrast to existing computational models of episodic memory (Norman and O'Reilly, 2003; Wixted, 2007; Greve et al., 2010), our model bridges behavioral outcomes with neural and synaptic mechanisms. It reproduces episodic memory phenomena on behavioral time scales under constrained network connectivity with plausible postsynaptic potentials, firing rates, and other biological parameters.

## Materials and Methods

### Neuron and synapse model

We use adaptive exponential integrate-and-fire point model neurons, which feature spike frequency adaptation, enriching neural dynamics and spike patterns, especially for the pyramidal cells (Brette and Gerstner, 2005). The neuron model is an effective model of cortical neuronal activity, reproducing a wide variety of electrophysiological properties, and offers a good phenomenological description of typical neural firing behavior, but it is limited in predicting the precise time course of the subthreshold membrane voltage during and after a spike or the underlying biophysical causes of electrical activity (Gerstner and Naud, 2009). We slightly modified it for compatibility with the BCPNN synapse model (Tully et al., 2014) by integrating an intrinsic excitability current.

Development of the membrane potential  $V_m$  and the adaptation current  $I_w$  is described by the following equations:

$$C_m \frac{dV_m}{dt} = -g_L(V_m - E_L) + g_L \Delta \tau e^{\frac{V_m - V_t}{\Delta \tau}} - I_w + I_{\text{ext}} + I_{\text{syn}} \quad (1)$$

$$\frac{dI_w}{dt} = \frac{-I_w}{\tau_{I_w}} + b \delta(t - t_{\text{sp}}). \quad (2)$$

Equation 1 describes the dynamics of the membrane potential  $V_m$  including an exponential voltage dependent

**Table 1: Neuron model and synaptic parameters**

Neuron model parameter	Symbol	Value	BCPNN parameter	Symbol	Value
Adaptation current	$b$	86 pA	BCPNN AMPA gain	$w_{gain}^{AMPA}$	0.76 nS
Adaptation decay time constant	$\tau_w$	280 ms	BCPNN NMDA gain	$w_{gain}^{NMDA}$	0.07 nS
Membrane capacitance	$C_m$	280 pF	BCPNN bias current gain	$\beta_{gain}$	40 pA
Leak reversal potential	$E_L$	-70.6 mV	BCPNN lowest spiking rate	$f_{min}$	0.2 Hz
Leak conductance	$g_L$	14 nS	BCPNN highest spiking rate	$f_{max}$	25 Hz
Upstroke slope factor	$\Delta_T$	3 mV	BCPNN lowest probability	$\epsilon$	0.0026
Spike threshold	$V_t$	-55 mV	P trace time constant	$\tau_p$	15 s
Spike reset potential	$V_r$	-60 mV	Regular plasticity	$\kappa_{normal}$	1
Refractory period	$\tau_{ref}$	5 ms	Modulated plasticity	$\kappa_{boost}$	2
Receptor parameter	Symbol	Value	Short-term plasticity parameter	Symbol	Value
AMPA synaptic time constant	$\tau^{AMPA}$	5 ms	Utilization factor	$U$	0.2
NMDA synaptic time constant	$\tau^{NMDA}$	100 ms	Augmentation decay time constant	$\tau_A$	5 s
GABA synaptic time constant	$\tau^{GABA}$	5 ms	Depression decay time constant	$\tau_D$	280 ms
AMPA reversal potential	$E^{AMPA}$	0 mV			
NMDA reversal potential	$E^{NMDA}$	0 mV			
GABA reversal potential	$E^{GABA}$	-75 mV			

activation term. A leak current is driven by the leak reversal potential  $E_L$  through the conductance  $g_L$  over the neural surface with a capacity  $C_m$ . Additionally,  $V_t$  is the spiking threshold, and  $\Delta_T$  shapes the spike slope factor. After spike generation, membrane potential is reset to  $V_r$ . Spike emission upregulates the adaptation current by  $b$ , which recovers with time constant  $\tau_w$  (Table 1). To simplify the model, we have removed subthreshold adaptation, which is part of some AdEx models.

Besides a specific external input current  $I_{ext}$ , model neurons receive synaptic currents  $I_{synj}$  from conductance based glutamatergic and GABAergic synapses. Glutamatergic synapses feature both AMPA/NMDA receptor gated channels with fast and slow conductance decay dynamic, respectively. Current contributions for synapses are described as follows:

$$\begin{aligned}
 I_{synj} &= \sum_{syn} \sum_i g_{ij}^{syn}(t)(V_m - E_j^{syn}) \\
 &= I_j^{AMPA}(t) + I_j^{NMDA}(t) + I_j^{GABA}(t).
 \end{aligned}
 \tag{3}$$

The glutamatergic synapses are also subject to synaptic depression and augmentation with a decay factor  $\tau_D$  and  $\tau_A$ , respectively (Table 1), following the Tsodyks–Markram formalism (Tsodyks and Markram, 1997). We have chosen those time-constants from the plausible range of computational fits made on the basis of electrophysiological recordings of cortical pyramidal cells (Wang et al., 2006). The utilization factor  $u$  represents the fraction of available resources used up by each transmitted spike (a proxy of synaptic release probability), whereas  $x$  tracks the fraction of resources that remain available because of transmitter depletion (synaptic depression):

$$\frac{du_{ij}}{dt} = -\frac{u_{ij}}{\tau_A} + U(1 - u_{ij}) \sum_{sp} \delta(t - t_{sp}^i - t_{ij})
 \tag{4}$$

$$\frac{dx_{ij}}{dt} = \frac{1 - x_{ij}}{\tau_D} - Ux_{ij} \sum_{sp} \delta(t - t_{sp}^i - t_{ij}).
 \tag{5}$$

### Spike-based BCPNN plasticity

We implement synaptic plasticity of AMPA and NMDA connection components using the BCPNN learning rule (Lansner and Ekeberg, 1989; Wahlgren and Lansner, 2001; Tully et al., 2014). BCPNN is derived from Bayes rule, assuming a postsynaptic neuron employs some form of probabilistic inference to decide whether to emit a spike or not. Despite that it accounts for the basic Bayesian inference, it is considered more complex than the standard STDP learning rule (Caporale and Dan, 2008), and as such, it reproduces the main features of STDP plasticity.

The BCPNN synapse continuously updates three synaptic biophysically plausible local memory traces,  $P_i$ ,  $P_j$ , and  $P_{ij}$ , implemented as exponentially moving averages (EMAs) of preactivation, postactivation, and coactivation, from which the Bayesian bias and weights are calculated. EMAs prioritize recent patterns, so that newly learned patterns gradually replace old memories. Specifically, learning implements exponential filters,  $Z$ , and  $P$ , of spiking activity with a hierarchy of time constants,  $\tau_z$ , and  $\tau_p$ , respectively [the full BCPNN model implements additional eligibility  $E$  traces (Tully et al., 2014), which are not used here]. Because of their temporal integrative nature, they are referred to as synaptic (local memory) traces.

To begin with, BCPNN receives a binary sequence of presynaptic and postsynaptic spiking events ( $S_i, S_j$ ) to calculate the traces  $Z_i$  and  $Z_j$ :

$$\begin{cases} \tau_{z_i} \frac{dZ_i}{dt} = \frac{S_i}{f_{max} t_{spike}} - Z_i + \epsilon \\ \tau_{z_j} \frac{dZ_j}{dt} = \frac{S_j}{f_{max} t_{spike}} - Z_j + \epsilon \end{cases},
 \tag{6}$$

$f_{max}$  denotes the maximal neuronal spike rate,  $\epsilon$  is the lowest attainable probability estimate,  $t_{spike}$  denotes the spike duration while  $\tau_{z_i} = \tau_{z_j}$  are the presynaptic and postsynaptic time constants, respectively ( $\tau_z = \tau^{AMPA} = 5$  ms for AMPA, and  $\tau_z = \tau^{NMDA} = 100$  ms for NMDA components; Table 1).

P traces are then estimated from the Z traces as follows:

$$\begin{cases} \tau_p \frac{dP_i}{dt} = \kappa(Z_i - P_i) \\ \tau_p \frac{dP_j}{dt} = \kappa(Z_j - P_j) \\ \tau_p \frac{dP_{ij}}{dt} = \kappa(Z_i Z_j - P_{ij}) \end{cases} \quad (7)$$

The parameter  $\kappa$  adjusts the learning rate, reflecting the action of endogenous modulators of learning efficacy (i.e., activation of a D1R-like receptor). Setting  $\kappa = 0$  freezes the network's weights and biases, though in our simulations the learning rate remains constant ( $\kappa = 1$ ) during encoding (see Results, Semantization of episodic representations in the BCPNN model and Item-context interactions under STDP). However, we trigger a transient increase of plasticity in specific scenarios to model preferential retention of salient events (see Results, Preferential retention; Table 1).

Finally,  $P_i$ ,  $P_j$ , and  $P_{ij}$  are used to calculate intrinsic excitability  $\beta_j$  and synaptic weights  $w_{ij}$  with a scaling factor  $\beta_{gain}$  and  $w_{gain}^{syn}$  respectively (Table 1):

$$\begin{cases} w_{ij} = w_{gain}^{syn} \log \frac{P_{ij}}{P_i P_j} \\ \beta_j = \beta_{gain} \log(P_j) \end{cases} \quad (8)$$

### Spike-based STDP learning rule

In our study, we examine the impact on semantization when the STDP learning rule replaces BCPNN associative connectivity in the same episodic memory task. Synapses under STDP are developed and modified by a repeated pairing of presynaptic and postsynaptic spiking activity, while their relative time window shapes the degree of modification (Ren et al., 2010). The amount of trace modification depends on the temporal difference ( $\Delta t$ ) between the time point of the presynaptic action potential ( $t_i$ ) and the occurrence of the postsynaptic spike ( $t_j$ ) incorporating a corresponding transmission delay ( $\tau_d$ ) from neuron  $i$  to neuron  $j$ :

$$\Delta t = t_j - (t_i + \tau_d). \quad (9)$$

After processing  $\Delta t$ , STDP updates weights accordingly:

$$\Delta w_{ij}(\Delta t) = \begin{cases} \lambda (1 - w)^{\mu_+} e^{-|\Delta t|/\tau_+} & \text{if } \Delta t \geq \tau_d \\ \lambda \alpha w^{\mu_-} e^{-|\Delta t|/\tau_-} & \text{if } \Delta t < \tau_d \end{cases} \quad (10)$$

Here,  $\lambda$  corresponds to the learning rate,  $\alpha$  reflects a possible asymmetry between the scale of potentiation and depression,  $\tau_{\pm}$  control the width of the time window,

**Table 2: STDP model parameters**

Parameter	Symbol	Value
Weight initialization	$w_0$	0 nS
AMPA maximum allowed weight	$w_{max}^{AMPA}$	13.5 nS
NMDA maximum allowed weight	$w_{max}^{NMDA}$	3.5 nS
Learning rate	$\lambda$	0.01
Asymmetry parameter	$\alpha$	1.2
Weight dependence exponent, potentiation	$\mu_+$	1
Weight dependence exponent, depression	$\mu_-$	1
Symmetric time window	$\tau_{\pm}$	20 ms

while  $\mu_{\pm} \in \{0,1\}$  allows to choose between different versions of STDP (i.e., additive, multiplicative; Morrison et al., 2008). Synapses are potentiated if the synaptic event precedes the postsynaptic spike and get depressed if the synaptic event follows the postsynaptic spike (Van Rossum et al., 2000).

Associative weights  $w_{ij}$  are initialized to  $w_0$ , and their maximum allowed values are constrained according to  $w_{max}$  to ensure that synaptic weights are always positive and between  $[w_0, w_{max}]$  (Table 2). The resulting associative weight distributions are generally comparable in strength to the BCPNN model weights, but to make them match, we adjust  $w_{max}$  in conjunction with a reasonably small learning rate  $\lambda$ . The maximum allowed weight ( $w_{max}$ ) is a necessary standard parameter of the default STDP we use in NEST (see below, Code accessibility). To obtain a stable competitive synaptic modification, the integral of  $\Delta w_{ij}$  must be negative (Song et al., 2000). To ensure this, we choose  $\alpha = 1.2$ , which introduces an asymmetry between the scale of potentiation and depression along with a symmetric time window resulting in a ratio of  $\alpha\tau_-/\tau_+ > 1.0$  (Ren et al., 2010). We set  $\mu_{\pm} = 1$  resulting in multiplicative STDP (in-between values lead to rules which have an intermediate dependence on the synaptic strength).

### Two-network architecture and connectivity

The network model features two reciprocally connected networks, the so-called Item and Context networks. For simplicity, we assume that Item and Context networks are located at a substantial distance accounting for the reduced internetwork connection probabilities (Table 3). Each network follows a cortical architecture with modular structure compatible with previous spiking implementations of attractor memory networks (Lansner, 2009; Lundqvist et al., 2011; Tully et al., 2014, 2016; Fiebig and Lansner, 2017; Chrysanthis et al., 2019; Fiebig et al., 2020), and is best understood as a subsampled cortical layer 2/3 patch with nested hypercolumns (HCs) and minicolumns (MCs; Fig. 1A). Both networks span a regular-spaced grid of 12 HCs (Table 3), each with a diameter of 500  $\mu\text{m}$  (Mountcastle, 1997). In our model, items are embedded in the Item network and context information in the Context network as internal well consolidated long-term memory representations (cell assemblies), supported via intranetwork weights derived using prior

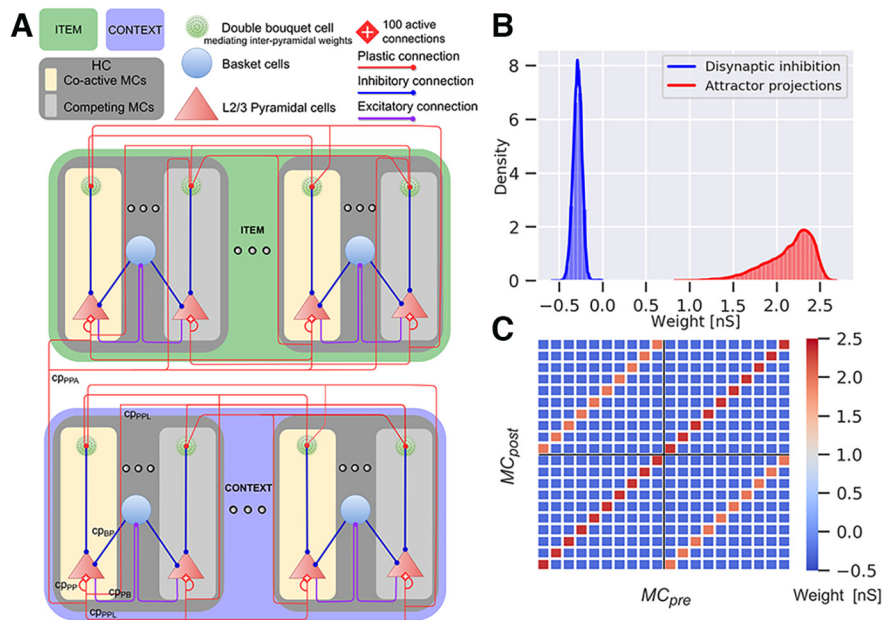
**Table 3: Network layout, connectivity, and stimulation protocol**

Layout	Symbol	Value	Connectivity	Symbol	Value	Stimulation	Symbol	Value
Cortical patch size	$C_{ps}$	$2.0 \times 1.5$ mm	Axonal conduction speed	$V$	0.2 m/s	Background noise PYR (encoding)	$r_{bg}^{PYR-encoding}$	650 Hz
Simulated HCs (each network)	$n_{HC}$	12	Myelinated axonal speed	$V_{myel}$	2 m/s	Background noise PYR (recall)	$r_{bg}^{PYR-recall}$	450 Hz
Simulated MCs (each network)	$n_{MC}$	120	Minimal synaptic delay	$t_{min}^{syn}$	1.5 ms	Background noise BA	$r_{bg}^{BA}$	75 Hz
Simulated MCs per HC	$n_{MC}^{HC}$	10	HC diameter	$d_{HC}$	0.5 mm	Background conductance	$g_{bg}^{PYR,BA}$	$\pm 1.5$ nS
Number of items	$n_{ITEM}$	4 (from 10)	Distance between networks	$d_{CONTEXT}^{ITEM}$	10 mm	Stimulation duration	$t_{stim}$	250 ms
Number of contexts	$n_{CONTEXT}$	10 (from 10)	PYR-PYR recurrent cp	$cp_{PPP}$	0.2	Stimulation rate	$f_{stim}$	500 Hz
Layer 2/3 pyramidal per MC	$n_{MC}^{PYR-L23}$	30	PYR-PYR long-range cp	$cp_{PPL}$	0.25	Cue stimulation length	$t_{cue}$	50 ms
Basket cells per MC	$n_{MC}^{Basket}$	2	PYR-PYR associative cp	$cp_{PPA}$	0.02	Cue stimulation rate	$r_{cue}$	400 Hz
MC grid size (Item + Context)	$G_{MC}^{TOTAL}$	$24 \times 10$	PYR-BA cp, BA-PYR cp	$cp_{PB}, cp_{BP}$	0.7	Stimulation and cue conductance	$g_{stim}$	+1.5 nS
			PYR-BA cc	$g_{PB}$	3 nS	Interstimulus interval	$T_{stim}$	500 ms
			BA-PYR cc	$g_{BP}$	-7 nS	Attractor detection threshold	$r_{th}$	10 Hz

PYR, pyramidal cell; BA, basket cell; cp, connection probability; cc, connection conductance.

BCPNN learning with long time constant (Fig. 1B,C; Table 3). Consequently, these weights were resistant to changes during associative learning of projections between Item and Context networks (see Results). Our item and context memory representations are distributed and nonoverlapping, i.e., with a single distinct pattern-specific (encoding) MC per HC. This results in sparse neocortical activity patterns (Barth and Poulet, 2012). It

should be noted that the model tolerates only a marginal overlap between different memory patterns, i.e., shared encoding MCs (data not shown). Each MC is composed of 30 pyramidal cells (representing the extent of layer 2/3) with shared selectivity, forming a functional (not strictly anatomic) column. In total, the 24 HCs (10 MCs each) of the model contain 7200 excitatory and 480 inhibitory cells, significantly downsampling the number of



**Figure 1.** Network architecture and connectivity of the Item (green) and Context (blue) networks. **A**, The model represents a subsampled modular cortical layer 2/3 patch consisting of MCs nested in HCs. Both networks contain 12 HCs, each comprising 10 MCs. We preload abstract long-term memories of item and context representations into the respective network, in the form of distributed cell assemblies with weights establishing corresponding attractors. Associative plastic connections bind items with contexts. The network features lateral inhibition via basket cells (purple and blue lines) resulting in a soft winner-take-all dynamics. Competition between attractor memories arises from this local feedback inhibition together with disynaptic inhibition between HCs. **B**, Weight distribution of plastic synapses targeting pyramidal cells. The attractor projection distribution is positive with a mean of 2.1, and the disynaptic inhibition is negative with a mean of -0.3 (we show the fast AMPA weight components here, but the simulation also includes slower NMDA weight components). **C**, Weight matrix between attractors and competing MCs across two sampled HCs. The matrix displays the mean of the weight distribution between a presynaptic ( $MC_{pre}$ ) and postsynaptic MC ( $MC_{post}$ ), within the same or different HC (black cross separates grid into blocks of HCs, only two of which are shown here). Recurrent attractor connections within the same HC are stronger (main diagonal, dark red) compared with attractor connections between HCs (off-diagonals, orange). Negative pyramidal-pyramidal weights (blue) between competing MCs amounts to disynaptic inhibition mediated by double bouquet cells.

MCs per HC (~100 MCs per HC in biological cortex). The high degree of recurrent connectivity within (Thomson et al., 2002; Yoshimura and Callaway, 2005) and between MCs links coactive MCs into larger cell assemblies (Stettler et al., 2002; Binzegger et al., 2009; Muir et al., 2011; Eyal et al., 2018). Long-range bidirectional inter-network connections (item-context bindings or associative connections) are plastic (shown in Fig. 1A only for MC1 in HC1 of the Context network), binding items, and contextual information (Ranganath, 2010). On average, recurrent connectivity establishes 100 active plastic synapses onto each pyramidal cell from other pyramids with the same selectivity, because of a sparse inter-network connectivity ( $cp_{PPA}$ ) and denser local connectivity ( $cp_{PP}$ ,  $cp_{PPL}$ ; connection probability refers to the probability that there is a connection between a randomly selected pair of neurons from given populations; in Fig. 1A, connection probabilities are only shown for MC1 in HC1 of the Context network). The model yields biologically plausible EPSPs for connections within HCs ( $0.45 \pm 0.13$  mV), measured at resting potential  $E_L$  (Thomson et al., 2002). Densely recurrent nonspecific monosynaptic feedback inhibition mediated by fast spiking inhibitory cells (Kirkcaldie, 2012) implements a local winner-take-all structure (Binzegger et al., 2009) among the functional columns. IPSPs have an amplitude of  $-1.160$  mV ( $\pm 0.003$ ) measured at  $-60$  mV (Thomson et al., 2002). These bidirectional connections between basket and pyramidal cells within the local HCs are drawn with a 70% connection probability. Notably, double bouquet cells shown in Figure 1A are not explicitly simulated, but their effect is nonetheless expressed by the BCPNN rule. A recent study based on a similar single-network architecture (i.e., with the same modular organization, microcircuitry, conductance-based AdEx neuron model, cell count per MC and HC) demonstrated that learned mono-synaptic inhibition between competing attractors is functionally equivalent to the disinaptic inhibition mediated by double bouquet and basket cells (Chrysanthis et al., 2019). Parameters characterizing other neural and synaptic properties including BCPNN can be found in Table 1.

Figure 1B shows the weight distributions of embedded distributed cell assemblies, representing different memories stored in the Item and Context networks. Attractor projections can be further categorized into strong local recurrent connectivity within HCs, and slightly weaker long-range excitatory projections across HCs (Fig. 1C).

### Axonal conduction delays

Conduction delays ( $t_{ij}$ ) between a presynaptic neuron  $i$  and a postsynaptic neuron  $j$  are calculated based on their Euclidean distance,  $d$ , and a conduction velocity  $V$  (Eq. 11). Delays are randomly drawn from a normal distribution with a mean according to distance and conduction velocity, with a relative SD of 30% of the mean to account for individual arborization differences, and varying conduction speed as a result of axonal thickness and myelination. In addition, a minimal delay of 1.5 ms ( $t_{min}^{syn}$ ; Table 3) is added to reflect synaptic delays because of effects that are not explicitly

modelled, e.g., diffusion of neurotransmitters over the synaptic cleft, dendritic branching, thickness of the cortical sheet and the spatial extent of columns (Thomson et al., 2002). Associative inter-network projections have a 10-fold faster conduction speed than those within each network, reflecting axonal myelination:

$$\bar{t}_{ij} = \frac{d}{V} + t_{min}^{syn}, \quad t_{ij} \sim N(\bar{t}_{ij}, .3\bar{t}_{ij}). \quad (11)$$

### Stimulation protocol

Noise input to pyramidal cells and fast spiking inhibitory basket cells is a zero-mean noise, generated by two independent Poisson generators with opposing driving potentials. Pyramidal cells coding for specific items and contexts are stimulated with an additional specific excitation during encoding and cued recall (all parameters in Table 3). Item-context association encoding is preceded by a brief period of background noise excitation to avoid initialization transients.

### Attractor activation detector

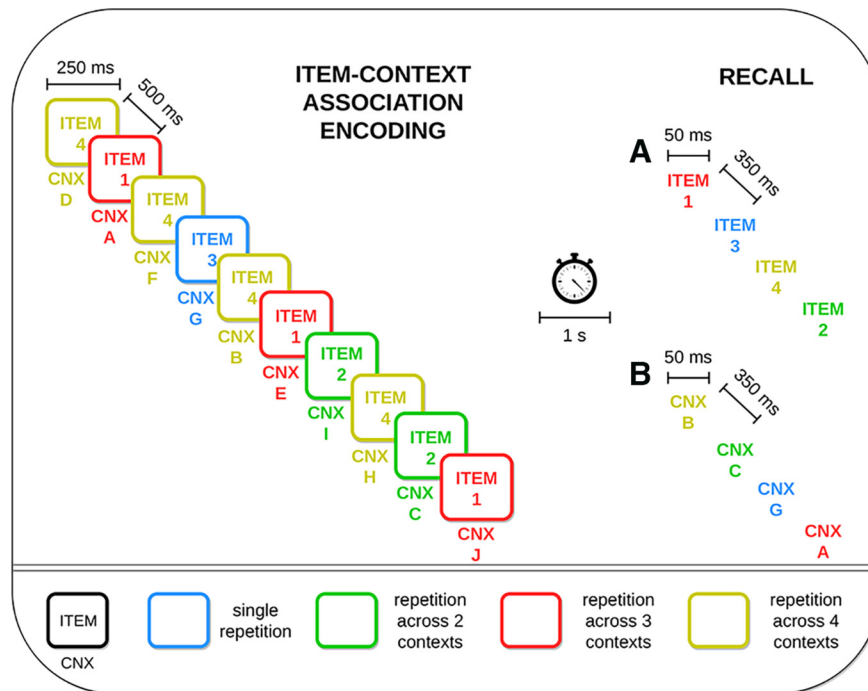
We detect and report cued recall of items or contexts by using an attractor activation detection algorithm based on EMAs of spiking activity. Pattern-wise EMAs are calculated using Equation 12, where the  $\delta$  function  $\delta$  denotes the spike events of a pattern-selective neural population of  $n_{pop} = 30$  pyramidal cells. The filter time constant  $\tau = 40$  ms is much larger than the sampling time interval  $\Delta T = 1$  ms:

$$e_0 = 0, e_t = \frac{\Delta T}{\tau} e_{t-\Delta T} + \delta_t \frac{1}{\tau n_{pop}}. \quad (12)$$

Pattern activations are detected by a simple threshold ( $r_{th}$ ) at about 10-fold the baseline activity with a small caveat: to avoid premature offset detection because of synchrony in fast spiking activity, we only count activations as terminated if they do not cross the threshold again in the next 40 ms. This method is highly robust because of the explosive dynamics of recurrent spiking activity for activated attractors in the network. Any attractor activation that crosses this threshold for at least 40 ms is considered a successful recall.

### Code accessibility

We use Neural Simulation Tool (NEST) version 2.2.2, and a custom-built BCPNN learning rule module (Tully et al., 2014) running on a Cray XC-40 Supercomputer. NEST simulates the dynamics of spiking neural models and features a convenient Python interface (PyNEST) to NEST's simulation kernel (Gewaltig and Diesmann, 2007). The custom-built spiking neural network implementation of the BCPNN learning rule for message passing interface (MPI) parallelized NEST is freely available online at Zenodo (<https://doi.org/10.5281/zenodo.5101626>) and also on Github using a Singularity container platform (<https://github.com/Nikolaos-Chrysanthis/BCPNN-NEST2.2.2-MPICH-SINGULARITY>). The custom NEST BCPNN module is available as Extended Data 1. Further,



**Figure 2.** Trial structure of the two simulated variants of the episodic memory task. Items are first associated with one or several contexts (CNX) during the encoding phase in 250-ms cue episodes, with an interstimulus interval of 500 ms. The colors of the coactivated contexts are consistent with their corresponding associated item. The recall phase occurs with a delay of 1 s and involves different trials with either brief cues (50 ms) of the (A) items or (B) contexts presented during the item-context association encoding phase.

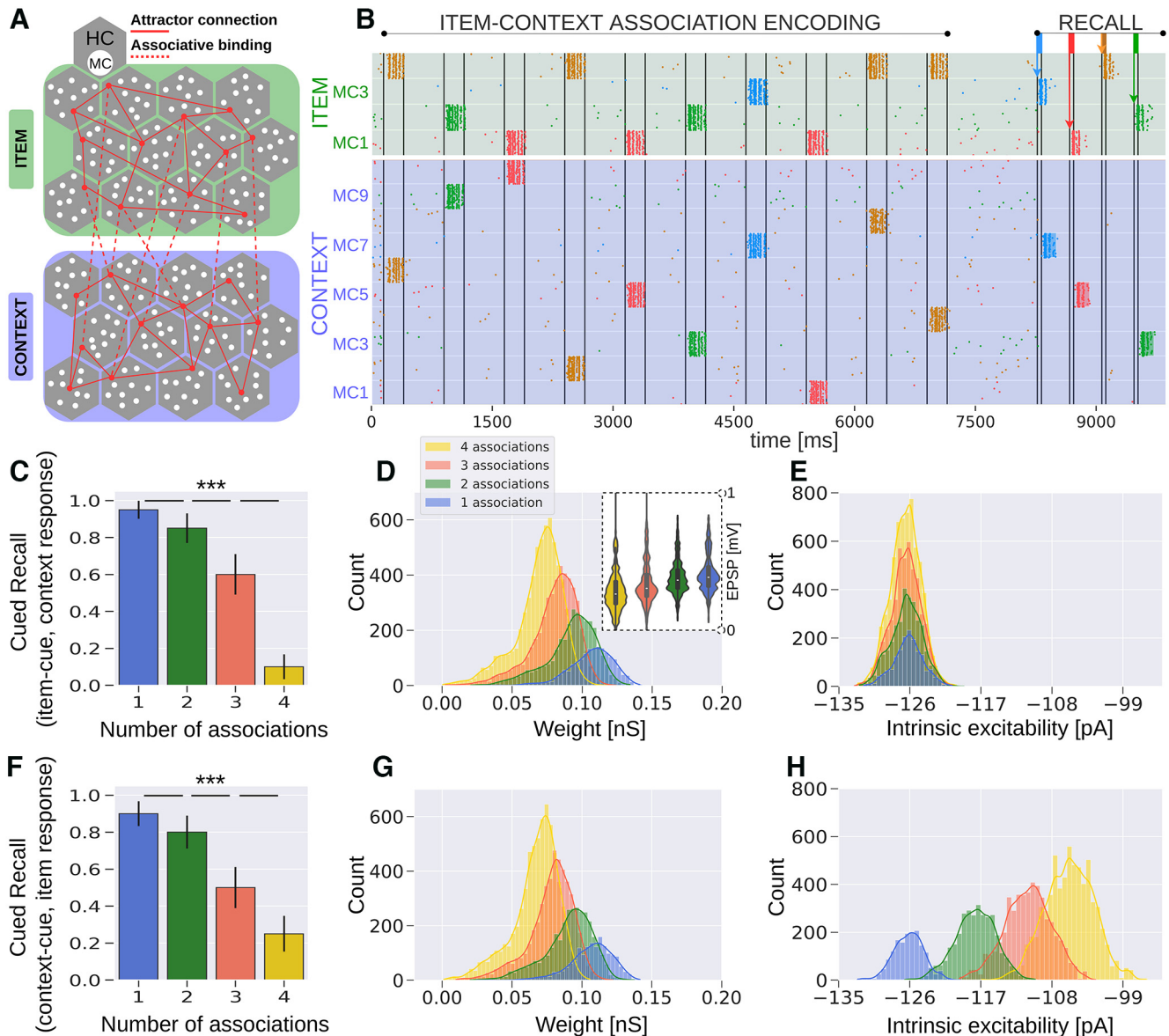
the model is also available on ModelDB (<https://modeldb.yale.edu/257610>).

## Results

### Semantization of episodic representations in the BCPNN model

An episodic memory task simulated in this work is inspired by a seminal memory effect shown in an experimental study by Opitz (2010). We deliberately abstract away some details of Opitz (2010) experimental design to provide a qualitative proof of principle with as few task assumptions as possible. This approach also offers a more generalized computational framework for studying the interplay of synaptic learning and its outcomes. In the same spirit, the systems architecture of our model is reduced to the Item and Context networks storing item and context information, respectively, as internal long-term memory representations (Fig. 1; for details, see Two-network architecture and connectivity and Table 3). We stimulate some items in a single context and others in a few different contexts establishing multiple associations (Fig. 2). Stimulus duration during encoding is  $t_{stim} = 250$  ms with a  $T_{stim} = 500$  ms interstimulus interval, and a test phase occurs after a 1-s delay period, which contains brief  $t_{cue} = 50$ -ms cues of previously learned items (Table 3). Figure 3A illustrates an item-context pair, established by an associative binding through plastic bidirectional BCPNN projections (dashed lines). Item and context attractors (solid red lines) are embedded in each network and remain fixed throughout the simulation, representing well-consolidated long-term memory. We

show an exemplary spike raster of pyramidal neurons in HC1 of both the Item and Context networks reflecting a trial simulation (Fig. 3B). Herein, item-3 (blue) establishes a single association, while item-4 (yellow) is encoded in four different contexts (Figs. 2A, 3B). We find clear evidence for strong item-context decoupling. The yellow item-4 (but not the blue item-3) is successfully recognized when cued but without any corresponding accompanying activation in the Context network (Fig. 3B). Figure 3C demonstrates that this item-context decoupling effect holds true also for the multi trial average as the performance of contextual retrieval when items serve as cues deteriorates with a higher context overload. Successful item recognition without any contextual information retrieval accounts for a “Know” response, as opposed to “Remember” judgments, which are accompanied by context recall. In fact, episodic loss in our network implies that no context is recalled despite the item memory activation. To elucidate this observed progressive loss of episodicity with the higher context variability (Fig. 3C), we sample and analyze the learned weight distributions of item-context binding recorded after the association encoding period (Fig. 3D). The item-context weight distribution in the one-association case is significantly stronger than in the two-association, three-association, or four-association case ( $p < 0.001$ , Mann-Whitney,  $N = 2000$ ). This progressive weakening of weights leads to significantly lower EPSPs for the associative projections ( $p < 0.05$  for one vs two associations;  $p < 0.001$  for two vs three and three vs four associations, Mann-Whitney,  $N = 300$ ; Fig. 3D, see inset). To measure EPSPs, we stimulate individually all the neurons in HC1 of an item which forms one, two, three, or four



**Figure 3.** Semantization of episodic memory traces. **A**, Schematic of the Item (green) and Context (blue) networks. Attractor projections are long-range connections across HCs in the same network and learned associative projections are connections between networks. **B**, Spike raster of pyramidal neurons in HC1 of both the Item and Context networks. Each context/item memory pattern corresponds to the activation of a unique set of MCs in its network. Items and their corresponding context representations are simultaneously cued in their respective networks (compare Fig. 2A). Each item is drawn with a unique color, while contexts inherit their coactivated item's color in the raster (i.e., the yellow pattern in the Item network is repeated over four different contexts, forming four separate associations marked with the same color). The testing phase occurs 1 s after the encoding. Brief 50-ms cues of already studied items trigger their activation. Following item activation, we detect evoked attractor activation in the Context network. **C**, Average cued recall performance in the Context network (20 trials). The bar diagram reveals progressive loss of episodic context information (i.e., semantization) over the number of context associations made by individual cued items (compare Fig. 2A). **D**, Distribution of plastic connection weights between the Item and Context networks (NMDA component shown here). Weights are noticeably weaker for items which participate in multiple associations. The distributions of synaptic weights exhibit a broader range for the items with multiple context associations, as the sample size is larger. The inset displays the distribution of EPSPs for the binding between Item and Context networks. The EPSP distributions follow the trend of the associative weights. The amplitudes (<1 mV) are lower for higher context variability. **E**, The distribution of intrinsic excitability currents of pyramidal cells coding for specific context representations. The intrinsic excitability features similar distributions because each context is activated exactly once, regardless of whether the associated item forms multiple associations or not. **F**, Average cued recall performance in the Item network (20 trials). Decontextualization over the number of associations is also observed when we briefly cue episodic contexts instead (compare Fig. 2B). **G**, Distribution of strength of plastic connections from the contexts to their associated items. Analogously to **D**, synapses weaken once an item is encoded in another context. **H**, Intrinsic plasticity distribution of cells in the Item network. Intrinsic excitability distributions are higher for pyramidal cells coding for repeatedly activated items;  $***p < 0.001$  (Mann-Whitney,  $N=20$  in **C**, **F**). Error bars in **C**, **F** represent SDs of Bernoulli distributions. Distributions of one, two, three, and four associations in **D**, **G**, **H** show significant statistical difference ( $p < 0.001$ , Mann-Whitney,  $N=2000$ ).



associations and record the postsynaptic potential onto their associated context neurons. EPSPs are in general below 1 mV (Thomson et al., 2002; Song et al., 2005), measured at resting potential  $E_L$  (Table 1), after item-context association encoding phase. Therefore, we attribute the loss of episodicity to a statistically significant weakening of the associative weight distributions with the increasing number of associated contexts. The associative weight distributions shown here refer to the NMDA component, while the weight distributions of the faster AMPA receptor connections display a similar trend (data not shown). The gradual trace modification we observe is a product of Bayesian learning, which normalizes and updates weights over estimated presynaptic (prior) as well as postsynaptic (posterior) spiking activity (see below, BCPNN and STDP learning rule in a microcircuit model for details).

Our simulation results are in line with related behavioral studies (Opitz, 2010; Smith and Manzano, 2010; Smith and Handy, 2014), which also reported item-context decoupling as the items were presented across multiple contexts. In particular, Opitz (2010) concluded that repetition of an item across different contexts (i.e., high context variability) leads to item-context decoupling, which is in agreement with our study. Furthermore, Smith and Manzano (2010) demonstrated in an episodic context variability task configuration, that recall deteriorates with context overload (number of words per context). Mean recall drops from  $\sim 0.65$  (one word per context) to 0.50 (three words per context), reaching  $\sim 0.33$  in the most overloaded scenario (fifteen words per context).

In Figure 3E, we show the distribution of intrinsic excitability over units representing different contexts. Pyramidal neurons in the Context network have a similar intrinsic excitability, regardless of their selectivity because all the various contexts are encoded exactly once.

Next, analogously to the previous analysis, we show that item-context decoupling emerges also when we briefly cue contexts rather than items during recall testing (Fig. 2B). In agreement with experimental data (Smith and Manzano, 2010; Smith and Handy, 2014), we obtain evidence of semantization as items learned across several discrete contexts are hardly retrieved when one of their associated contexts serves as a cue (Fig. 3F). We further sample and present the underlying associative weight distribution, between the Context and the Item networks (Fig. 3G). The distributions again reflect the semantization effect in a significant weakening of the corresponding weights. In other words, an assembly of pyramidal neurons representing items encoded across multiple contexts receives weaker projections from the Context network. At about four or more associations, the item-context binding becomes so weak that it fails to deliver sufficient excitatory current to trigger associated representations in the Item network. At the same time, intrinsic excitability of item neurons increases with the number of associated contexts corresponding to how much these neurons were active during the encoding phase (Fig. 3H; cf. Egorov et al., 2002; Tully et al., 2014).

### Item-context interactions under STDP

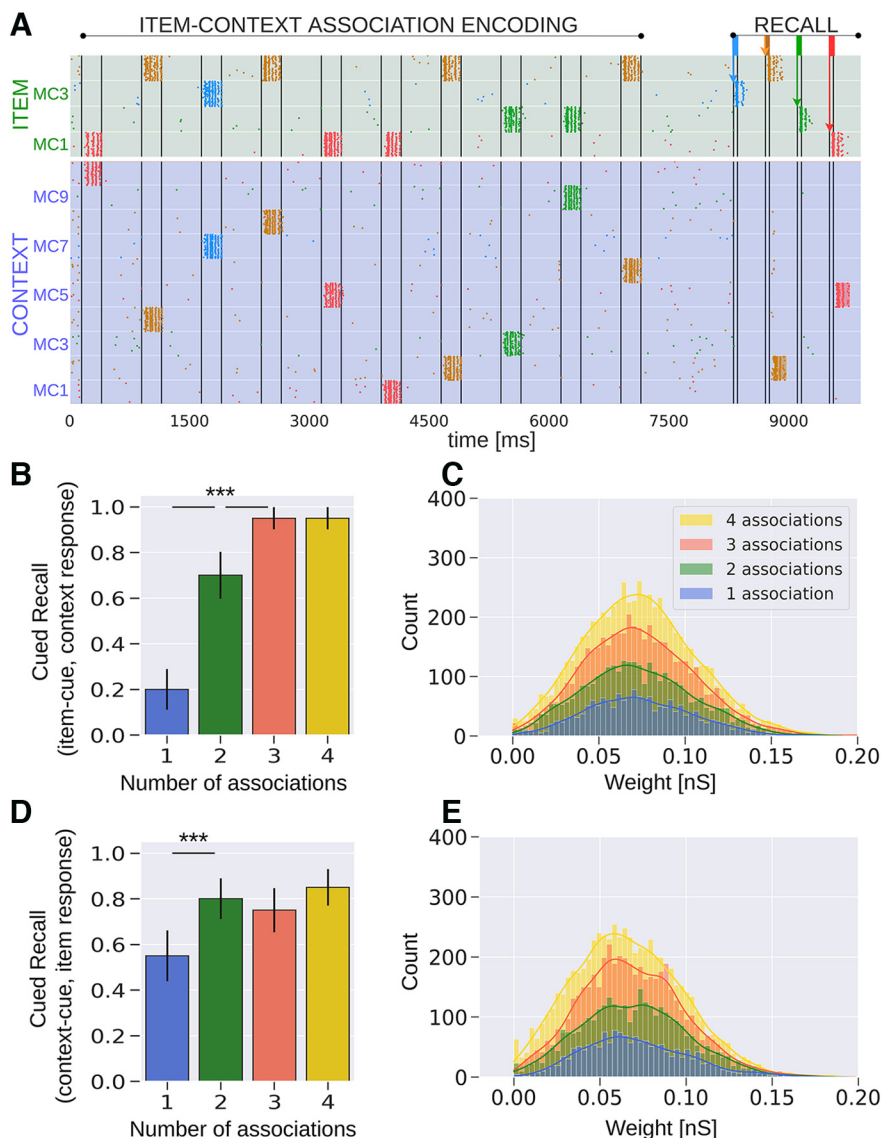
In this section, we contrast the results obtained with the BCPNN synaptic learning rule with those deriving from the more commonly used STDP learning rule in the same episodic memory task (Fig. 2; Spike-based STDP learning rule). The modular network architecture as well as neural properties and embedded memory patterns remain identical, but associative projections between networks are now implemented using a standard STDP synaptic learning rule (Morrison et al., 2008). The parameters of the STDP model are summarized in Table 2.

Figure 4A shows an exemplary spike raster of pyramidal cells in HC1 of both the Item and the Context networks, based on the first variant of the episodic memory task described in Figure 2A. As earlier, items are encoded in a single or in multiple different contexts and they are briefly cued later during recall. A successful item activation may lead to a corresponding activation of its associated information in the Context network. We detect these activations as before (see Materials and Methods, Attractor activation detector) and report the cued recall score over the number of associations (Fig. 4B).

Unlike the BCPNN network, we observe no evidence of semantization for high context variability. Instead, recollection is noticeably enhanced with an increase in the number of associations, which is in fact the opposite of what would be needed to explain item-context decoupling. STDP generates similarly strong associative binding regardless of context variability (Fig. 4C). The enhanced recollection in high context variability cases stems from the multiplicative effect of synaptic augmentation in the Tsodyks–Makram model on the Hebbian attractor weights. Items stimulated multiple times (e.g., four times) have a higher likelihood of being encoded near the end of the task, leading to more remaining augmentation during testing, thus, effectively boosting cued recall (Fig. 5A). This recency effect diminishes after removing synaptic augmentation from the model as attractor weights in the Item network have comparable distributions leading to similar cued recall performance regardless of context variability (Fig. 5B,C). As far as the context-cued variant of the task is concerned (Fig. 2B), there are also no signs of item-context decoupling for high context variability (Fig. 4D). The associative projections between Context and Item networks again have distributions with comparable means over context variability (Fig. 4E). Overall, decontextualization is not evident in either variant of the episodic memory task under the STDP learning rule.

### BCPNN and STDP learning rules in a microcircuit model

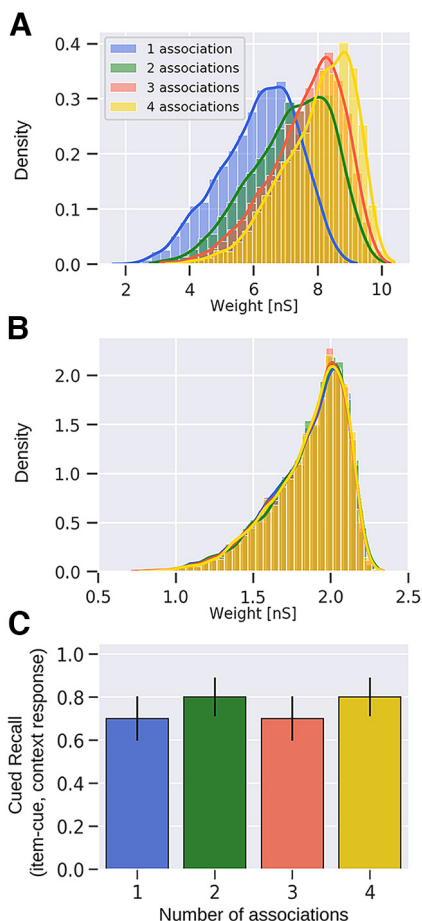
To better elucidate the emergent synaptic changes of the BCPNN and STDP model, we also apply these learning rules in a highly reduced microcircuit of spiking neurons. To this end, we now track the synaptic weight changes continuously. The neural and synaptic parameters (and most importantly all the plasticity parameters) used for the highly reduced BCPNN and STDP model are identical to the ones used for the large scale BCPNN and STDP model, respectively (see Tables 1, 2).



**Figure 4.** Network model where associative projections are implemented using standard STDP synaptic plasticity. **A**, Spike raster of pyramidal neurons in HC1 of both the Item and Context networks. **B**, Average item-cued recall performance in the Context network (20 trials). Episodic context retrieval is preserved even for high context variability (as opposed to BCPNN; compare Fig. 3C). **C**, Distribution of NMDA receptor mediated synaptic weights between the item and context neural assemblies following associative binding. The distributions of item-context weights have comparable means at  $\sim 0.065$  nS regardless of how many context associations a given item forms. Bins merely display a higher count for the four-association case as the total count of associative weights is more extensive compared with items with fewer associations. **D**, Average cued recall performance in the Item network when episodic contexts are cued (20 trials). **E**, Distribution of NMDA component weights between associated context and item assemblies;  $***p < 0.001$  (Mann–Whitney,  $N = 20$  in **B**, **D**). Error bars in **B**, **D** represent SDs of Bernoulli distributions.

First, we apply the BCPNN learning rule to the microcircuit model. We consider two separate item neurons (ID = 1 and 2), which form two or three associations with context neurons (ID = 3, 4, or 5, 6, 7), respectively (Fig. 6A). We display the synaptic strength development of the synapse between item neuron-1 and context neuron-3 (two associations, green), as well as the synapse between item neuron-2 and context neuron-5 (three associations, red) over the course of training these associations via targeted stimulation. BCPNN synapses get strengthened when the item-context pairs are simultaneously active

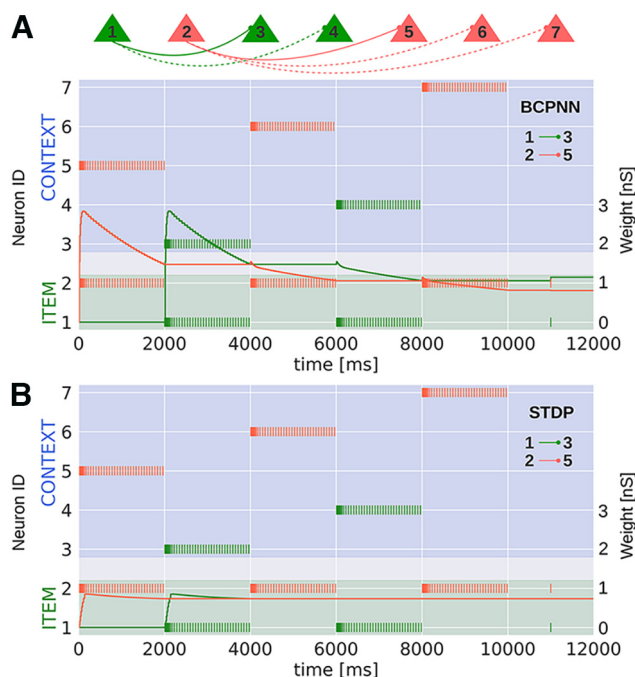
and weaken when the item in question is activated with another context. Therefore, synapses of the item neuron that is encoded in three different contexts converge on weaker weights (Fig. 6A, 12 s), than those of the item neuron with two associated contexts. Weight modifications in the microcircuit model reflect the synaptic alterations observed in the large-scale network. BCPNN weights are shaped by traces of activation and coactivation (Eqs. 7, 8; Materials and Methods), which also get updated during the activation of an item within another context. For example, the item neuron-1 and context neuron-3 are not



**Figure 5.** Removal of the augmentation mechanism in the network model. **A**, Distribution of AMPA component weights of the Item network including synaptic augmentation. The multiplicative effect of synaptic augmentation on the consolidated items features stronger combined synaptic strength for items with higher context variability. Slower NMDA receptor weights follow a similar pattern. Weight distributions of one, two, three, and four associations have statistical difference ( $p < 0.001$ , Mann-Whitney,  $N = 2000$ ). **B**, Distribution of AMPA component weights of the Item network after removing synaptic augmentation. **C**, Cued recall under STDP after removing synaptic augmentation. Average item-cued recall performance in the Context network (20 trials). To compensate for the removal of augmentation, we increased the stimulation rates and the synaptic gain eliciting comparable spiking activity. Error bars represent SDs of Bernoulli distributions.

stimulated together between 6 and 8 s, but neuron-1 and context neuron-4 are. Thus, the P traces of the item activation ( $P_i$ ) increase, while the ones linked to context-3 ( $P_j$ ) decay with a time constant of 15 s (Table 1). Since the item and context neuron (ID = 1, 3) are not stimulated together, their coactivation traces ( $P_{ij}$ ) decay between 6 and 8 s. Overall, this leads to a weakening of the weight and hence to a gradual decoupling (Eq. 8; Materials and Methods).

In the same manner, we keep track of weight change in a microcircuit with the STDP learning rule (Fig. 6B). Unlike the microcircuit with BCPNN presented in Figure 6A, the STDP weights corresponding to the associations made by both item neurons converge to similar values, although



**Figure 6.** Continuous weight recordings in a microcircuit model with plastic synapses under the **A**, BCPNN or **B**, STDP learning rule. Neural and synaptic parameters correspond to those in the scaled model. In both cases, two item neurons (ID = 1, 2) are trained to form two or three associations, respectively (dashed connections are simulated but their weight development is not shown here). During training, neurons are stimulated to fire at 20 Hz for 2 s. We display the developing synaptic weight between specific item-context pairs (ID = 1 and 3 in the 2-association scenario) and (ID = 2 and 5 in the 3-association scenario), and compare the converged weight values between the two-association and three-association case under both learning rules, following a final readout spike at 11 s.

they are associated with different number of contexts. As before, the synapse between an item neuron and an associated context neuron strengthens when this pair is simultaneously active, but remains stable when the item neuron is encoded in another context. For instance, the synapse between item neuron-2 and context neuron-5 strengthens when this pair is encoded (0–2 s), yet remains unaffected when item neuron-2 is activated in another context (i.e., context neuron-6, 4–6 s). This synaptic behavior explains the observed differences between the BCPNN and STDP large-scale model.

### Preferential retention

Several studies propose that one-shot salient events promote learning, and that these memories can be retained on multiple time scales ranging from seconds to years (Frankland et al., 2004; Petrican et al., 2010; Gruber et al., 2016; Panoz-Brown et al., 2016; Eichenbaum, 2017; Sun et al., 2018). Hypothetical mechanisms behind these effects are dopamine release and activation of DR1 like receptors, resulting in synapse-specific enhancement (Otmakhova and Lisman, 1996; Kuo et al., 2008), and systems consolidation (McClelland et al., 1995; Fiebig and

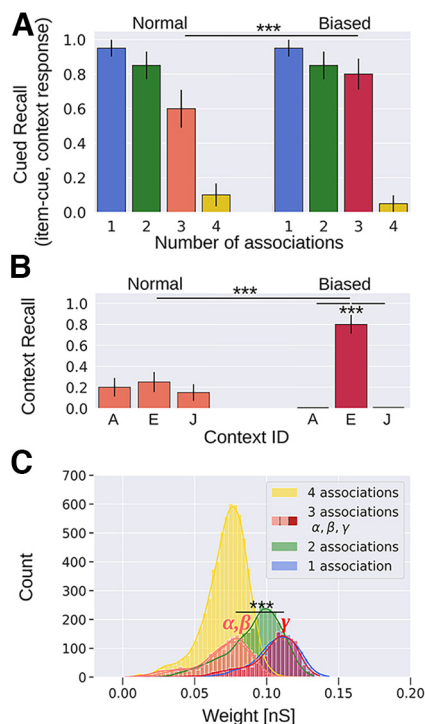
Lansner, 2014). Overall, salient or reward driven events may be encoded more strongly as the result of a transient plasticity modulation. Recall from long-term memory is often viewed as a competitive process in which a memory retrieval does not depend only on its own synaptic strength but also on the strength of other components (Shiffrin, 1970). In view of this, we study the effects of plasticity modulation on encoding specific items within particular contexts, with the aim of investigating the role of enhanced learning for semantization in our model.

Using the same network and episodic memory task as before (Fig. 2A), we modulate plasticity during the encoding of item-1 (red) in context-E via  $\kappa = \kappa_{boost}$  (Eq. 7; Materials and Methods; Table 1). This results in an increased cued recall probability for the item associated with three episodic contexts relative to the unmodulated control (Fig. 7A, Normal vs Biased scenario, three associations). Episodic retrieval improves from 0.6 (Normal; Fig. 7A, left) to 0.8 (Biased, modulated plasticity; Fig. 7A, right) when item-1 is cued, which now performs more similarly to item with just two associated contexts. We further analyze and compare the recall of each context when its associated item-1 is cued (Fig. 7B, three associations). The control scenario (Normal; Fig. 7B, left) without transient plasticity modulation shows that the three contexts (ID = A, E, and J) are all recalled with similar probabilities. In contrast, encoding a specific pair with enhanced learning (upregulated  $\kappa = \kappa_{boost}$ ) yields higher recall for the corresponding context. In particular, the plasticity enhancement during associative encoding of the context-E (with item-1) results in an increased recall score to 0.8 (0.25 control), while the other associated contexts, ID = A and J, are suppressed (Fig. 7B), primarily because of soft winner-take-all competition between contexts (Fig. 1A).

We attribute these changes to the stronger weights because of enhanced learning (Fig. 7C, dark red distribution,  $\gamma$ ). Weights between unmodulated item-context pairs (item-1 and context-A, -J) show mostly unaltered weight distributions ( $\alpha, \beta$ , light red), while the biased associative weight distribution between item-1 and context-E is now comparable to the weight distribution of the one-association case. Performance does not exactly match that case though because of some remaining competition among the three contexts. Overall, these results demonstrate how a single salient episode may strengthen memory traces and thus impart resistance to semantization (Rodríguez et al., 2016).

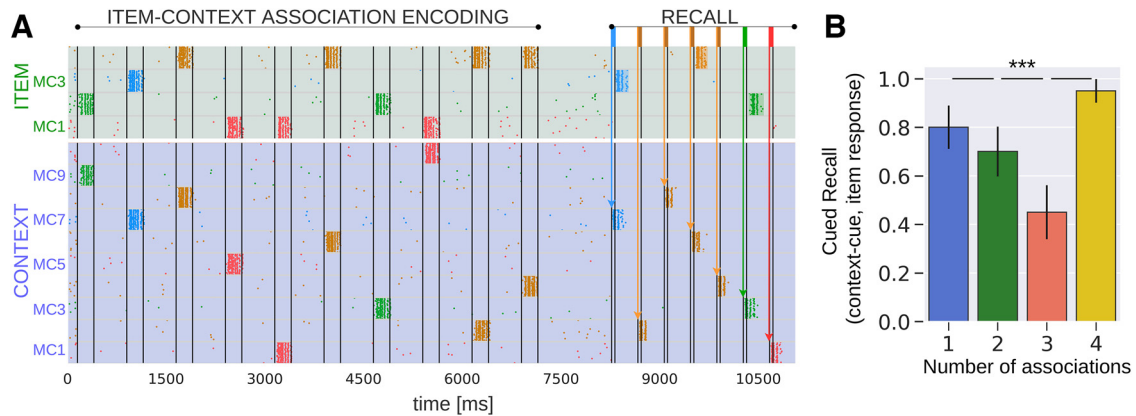
## Discussion

The primary objective of this work was to explore the interaction between synaptic plasticity and context variability in the semantization process. To cast new light on the episodic-semantic interplay, we built a cortical memory model of two spiking neural networks that feature the same modular architecture. The networks are coupled with plastic associative connections, which collectively represent distributed cortical episodic memory. Our results suggest that some forms of plasticity offer a synaptic explanation for the cognitive phenomenon of semantization, thus bridging scales and linking network connectivity and dynamics with



**Figure 7.** Plasticity modulation of a specific item-context pair enhances recollection and counteracts semantization. **A**, Context recall performance. One of the pairs (context-E, item-1) presented in the episodic memory task (compare Fig. 2A) is subjected to enhanced plasticity during encoding, resulting in the boosted recall rate (3 associations, Normal vs Biased, 20 trial average). **B**, Individual context retrieval contribution in the overall recall (3 associations). Retrieval is similar among the three contexts since plasticity modulation is balanced (left: Normal,  $\kappa = \kappa_{normal}$ ; compare Table 1). However, when context-E is encoded with enhanced learning (with item-1), its recall increases significantly (right: Biased,  $\kappa = \kappa_{boost}$ ; compare Table 1). **C**, Weight distributions of the NMDA weight component. Encoding item-1 with context-E under modulated plasticity yields stronger synaptic weights [3 association,  $\alpha, \beta$  (light red, highly overlapping distributions) vs  $\gamma$  (dark red)];  $***p < 0.001$  (Mann-Whitney,  $N = 20$  in **A**, **B**,  $N = 2000$  in **C**). Error bars in **A**, **B** represent SDs of Bernoulli distributions. Weight distributions of one, two, three- $\alpha, \beta$ , and four associations in **C** show significant statistical difference ( $p < 0.001$ , Mann-Whitney,  $N = 2000$ ).

behavior. We use a spiking neuronal network model combined with BCPNN, which allows us to directly compare it with a standard Hebbian STDP learning rule. In particular, we demonstrated that with Bayesian-Hebbian (BCPNN) synaptic plasticity, but not with standard Hebbian STDP, the model can reproduce traces of semantization as a result of learning. Notably, this was achieved with biologically constrained network connectivity, postsynaptic potential amplitudes, and firing rates compatible with mesoscale recordings from cortex and earlier models. Nevertheless, our hypothesis of the episodic-semantic interplay at a neural level requires further experimental study of the synaptic strength dynamics in particular. As mentioned, quantitative data on cortical synaptic plasticity is still quite limited, and while STDP has been shown to offer explanation for



**Figure 8.** Average cued recall performance in the Item network after sequentially cueing all the contexts that are associated with the item that forms four associations. **A**, Spike raster of pyramidal neurons in HC1 of both the Item and Context networks. The cue paradigm during test for the one-association, two-association, and three-association case remains identical to the control case (compare Fig. 2B). However, in particular for the four-association case, we sequentially cue all the four available contexts that share the same target item. **B**, Average cued recall performance in the Item network (20 trials). The bar diagram reveals progressive loss of item information over the number of context associations, but not for the four-association case at which all the available contexts were cued during test. Thus, providing more evidence via different sources boosts retrieval (~95%) recovering a nearly decontextualized item (compare Fig. 3F, four associations, single cue, 25% accuracy score); \*\*\* $p < 0.001$  (Mann–Whitney,  $N = 20$ ). Error bars represent SDs of Bernoulli distributions.

some associative memory phenomena (Pokorny et al., 2020), any specific plasticity rule is insufficiently validated experimentally. Yet our results with BCPNN offer a possible explanation and testable behavioral predictions. The spiking version of this plasticity rule has repeatedly been shown to be compatible with detailed, biologically constrained network activity and structure. Importantly, our simulations clearly demonstrate how cognitive phenomena such as semantization could be produced and thus explained by microscopic plasticity processes. In particular, BCPNN solves the issue of decontextualization by its information-theoretical principle, not by being hand-crafted to do so. Like any Bayesian estimator, BCPNN trades-off synaptic strength (weights) for increased intrinsic excitability (bias) in highly active neurons, thus decreasing synaptic strength of neurons that are highly active outside of a specific spiking correlation. Unlike many other conceivable learning rules that might achieve this effect, BCPNN is working only on locally available information, and is thus also biologically plausible.

Our study conforms to related behavioral experiments reporting that high context variability or context overload leads to item-context decoupling (Opitz, 2010; Smith and Manzano, 2010; Smith and Handy, 2014). These studies suggest that context-specific memory traces transform into semantic representations while contextual information is progressively lost. Traces of item memory representations remain intact but fail to retrieve their associated context. Semantization is typically described as a decontextualization process that occurs over time. However, several experiments, including this study, proposed that exposures of stimuli in different additional contexts (rather than time itself) is the key mechanism advancing semantization (Opitz, 2010; Smith and Manzano, 2010; Smith and Handy, 2014). In fact, simple language vocabulary learning implies that

learners encode words in several different contexts, which leads to semantization and a definition-like meaning of the studied word (Beheydt, 1987; Bolger et al., 2008). Although our network model is limited to a simple item-context decoupling scenario, the proposed plasticity mechanistic explanation for the observed item-context decoupling effect may be generalized to support semantization in more complex scenarios in which other mechanisms may synergistically interact and contribute to decontextualization. Admittedly, our hypothesis does not exclude other seemingly coexisting phenomena and mechanisms supporting memory retrieval that may facilitate semantization over time, e.g., reconsolidation or systems consolidation because of sleep or aging (Friedrich et al., 2015). Further, our model does not feature any higher-order mechanisms allowing a neutral stimulus (lacking prior pairing) to evoke the same contextual memory response as a conditioned stimulus does despite their prior pairing. In other words, each stimulus has to be independently coupled with its context(s).

We also demonstrated (Results, Preferential retention) how a transient plasticity modulation, reflecting known isolation effects, may preserve episodicity, staving off decontextualization. Semantization may also be overcome by accumulating additional evidence regarding an episode. In our simulations, we typically used single context cues to retrieve an item during cued recall (e.g., item-4 forms four associations, but only one of its associated contexts was cued, compare Figs. 2B and 3F). However, an interesting question is whether providing multiple context cues that share the same target item boosts its recall (Fig. 8A). Figure 8B shows the result where we sequentially stimulated all the four different contexts in the four-association case. The 4-fold contextual information considerably increases the likelihood of retrieval of a nearly fully semantized item (compare Figs. 3F and 8B, four

associations). These results are relatively intuitive yet novel from a modeling point of view and in line with behavioral studies reporting enhanced cued recall with multiple cues compared with a single one (Rubin and Wallace, 1989; Broadbent et al., 2020; Pearson and Wilbiks, 2021).

To our knowledge, there is no other spiking computational model of comparable detail that captures the semantization of episodic memory explored here, while simultaneously offering a neurobiological explanation of this phenomenon. Unlike other dual-process episodic memory models, which require repeated stimulus exposures to support recognition (Norman and O'Reilly, 2003), our model is able to successfully recall events learned in "one shot" (a distinctive hallmark of episodicity). We note that the attractor-based theory proposed in this study does not exclude the possibility of a dual-process explanation for recollection and familiarity (Yonelinas, 2002; Yonelinas et al., 2010).

### Related models of familiarity and recollection

Perceptual or abstract single-trace dual-process computational models based on signal detection theory explain episodic retrieval but the potential loss of contextual information is only implied as it does not have its own independent representation (Wixted, 2007; Greve et al., 2010). These computational models often aim to explain traditional R/K behavioral studies. As discussed earlier, participants in such studies are instructed to give a "Know" response if the stimulus presented in the test phase is known or familiar without any contextual detail about its previous occurrence. Conversely, "Remember" judgments are to be provided if the stimulus is recognized along with some recollection of specific contextual information pertaining to the study episode. This results in a strict criterion for recollection, as it is possible for a subject to successfully recall an item but fail to retrieve the source information (Ryals et al., 2013). Numerous studies suggest that recollection contaminates "Know" reports because recalling source information sensibly assumes prior item recognition (Wais et al., 2008; Johnson et al., 2009). Mandler (1979, 1980) and Atkinson and Juola (1973) treat familiarity as an activation of preexisting memory representations. Our results are compatible with this notion because our model proposes to treat item-only activations as "Know" judgments, while those accompanied by the activation of context representations best correspond to a "Remember" judgment. Item activation is a faster process and precedes context retrieval (Yonelinas and Jacoby, 1994), and our model reflects this finding by necessity, as item activations are causal to context retrieval.

We assume that familiarity recognition is simply characterized by lack of contextual information, yet the distinction we make between the Context and Item networks is arbitrary. Memory patterns stored in the Context network are referred to as contexts and those in the Item network as items. From the perspective of the network's architecture, items and contexts have representations of the same nature, nonoverlapping

sparse distributed patterns. While sparse internetwork connectivity is sufficient for our model's function, both networks may just as well be part of the same cortical brain area. The actual physical separation of the two networks (which incurs connection delays commensurate with the axonal conduction speed) is motivated by our assumption that items and contexts are not necessarily represented within the same network. A more specific scenario might assume that items and contexts share part of the same local network. In principle, our model should be capable of replicating similar results in that case.

### Biological plausibility and parameter sensitivity

We investigate and explain behavior and macroscale system dynamics with respect to neural processes, biological parameters of network connectivity, and electrophysiological evidence. Our model consequently builds on a broad range of biological constraints such as intrinsic neuronal parameters, cortical laminar cell densities, plausible delay distributions, and network connectivity. The model reproduces plausible postsynaptic potentials (EPSPs, IPSPs) and abides by estimates of connection densities (i.e., in the associative pathways and projections within each patch), axonal conductance speeds, typically accepted synaptic time constants for the various receptor types (AMPA, NMDA, and GABA), with commonly used neural and synaptic plasticity time constants (i.e., adaptation, depression).

The model synthesizes a number of functionally relevant processes, embedding different components to model composite dynamics, hence, it is beyond this study to perform a detailed sensitivity analysis for every parameter. Instead, we provide insightful observations for previously unexplored parameters that may critically affect semantization. Importantly, a related modular cortical model already investigated sensitivity to important short-term plasticity parameters (Fiebig and Lansner, 2017). After extensive testing we conclude that the model is generally robust to a broad range of parameter changes and its performance only gradually degrades in terms of the effect size. We expect even lower sensitivity to parameter variations in a network approaching biological scales. Further, it is worth reporting that the model's function is preserved across a wide range of sizes of a cortical column as long as the number of pyramidal cells is not excessively low. The same holds for the population of the inhibitory basket cells provided that the rough total inhibitory synaptic current is maintained by controlling feedback synaptic strength ( $g_{BP}$ ; Table 3).

The P trace decay time constant,  $\tau_p$ , of the BCPNN model is critical for the learning dynamics modelled in this study because it controls the speed of learning in associative connections and the resulting weight amplitude. High values of  $\tau_p$  imply long-lasting but weaker memory traces and therefore lead to slower, more inertial learning (more resistance to change and encoding new information as well as forgetting) with overall lower weights and hence weaker binding. Varying  $\tau_p$  by  $\pm 30\%$  does not change

the main outcome, i.e., episodicity still deteriorates with a higher context variability. At the same time, as mentioned, slower weight development results in weaker associative binding and overall lower recall (and vice versa for faster learning). To compensate for this loss of episodicity, an additional increase in the unspecific input is usually sufficient to trigger comparable recall rates. Alternatively, the recurrent excitatory gain can be amplified to complete noisy inputs toward discrete embedded attractors. Unspecific background input during recall plays a critical role as well. In general, we use a low background noise input into the two coupled networks. However, for the enhanced noise by +40% the model operates in a free recall regime with spontaneously reactivating memories without any external cues.

As we explained in the investigation of the reduced microcircuit model, semantization is an inherent property of BCPNN-driven weight dynamics, derived from Bayesian logic. However, countervailing forces in local microcircuits contribute to the generation and maintenance of our associated memories, Bayesian weight development drives semantization while intrinsic plasticity counteracts it. In consequence, it is possible to lessen the relative impact of this synaptic weight-dependent effect by making intrinsic plasticity more prominent: frequently activated items (in varying contexts) become more excitable because of memory recency effect. Conversely, we can maximize the semantization effect by making the bias current weaker, though fully removing it ( $\beta_{gain} = 0$ ) is hard to justify from a biological perspective. By manipulating the strength of intrinsic plasticity ( $\beta_{gain}$ ) to diminish signs of decoupling as described above, the capacity, i.e., the number of retrievable item-context associations, can increase beyond three associations (compare Fig. 3). Other key factors that can enhance model capacity for item-context associations (resistance to semantization over many established episodic associations) are larger network size, higher associative binding connection probability (e.g., increase in  $c_{PPA}$  from 2% to 4%; Table 3), and elevated background unspecific noise during the cue-response association period. Strengthening associative binding by upregulating  $w_{gain}^{syn}$  can also enhance the model's capacity for item-context associations (Table 1). Still, there is an upper limit to  $w_{gain}^{syn}$  as extreme values can lead to implausible EPSPs.

This study also demonstrates how a selective transient increase of plasticity can counteract semantization. The plasticity of the model can be modulated via the parameter  $\kappa$  (Eq. 7; Materials and Methods). Typically,  $\kappa$  is set to 1 ( $\kappa = \kappa_{normal}$ ; Table 1), whereas we double plasticity ( $\kappa = \kappa_{boost}$ ; Table 1), when modeling salient episodic encoding. We notice that by selectively tripling or quadrupling plasticity (relative to baseline) during encoding of a specific pair whose item component forms many other associations, the source recall improves progressively (data shown only for  $\kappa = \kappa_{boost}$  in Results, Preferential retention).

Finally, in Results, BCPNN and STDP learning rule in a microcircuit model, we compare STDP and BCPNN plasticity in a highly reduced model. We bind items with contexts to form different number of associations and keep

track of the weight development per time step. STDP plasticity generates same magnitude item-context binding regardless of how many associations an item forms. A detailed parameter analysis for every critical synaptic parameter ( $\pm 30\%$ ) did not yield any behaviorally significant changes to the converged weights.

### Semantization over longer time scales

Source recall is likely supported by multiple independent, parallel, interacting neural structures and processes since various parts of the medial temporal lobes, prefrontal cortex and parts of the parietal cortex all contribute to episodic memory retrieval including information about both where and when an event occurred (Gilboa, 2004; Diana et al., 2007; Watrous et al., 2013). A related classic idea on semantization is the view that it is in fact an emergent outcome of systems consolidation. Sleep-dependent consolidation in particular has been linked to advancing semantization of memories and the extraction of gist information (Payne et al., 2009; Friedrich et al., 2015).

Models of long-term consolidation suggest that retrieval of richly contextualized memories become more generic over time. Without excluding this possibility, we note that this is not always the case, as highly salient memories often retain contextual information (which our model speaks to). Instead, our model argues for a much more immediate neural and synaptic contribution to semantization that does not require slow multiarea systems level processes that have yet to be specified in sufficient detail to be tested in neural simulations. It has previously been shown, however, that an abstract simulation network of networks with broader distributions of learning time constants can consolidate memories across several orders of magnitude in time, using the same Bayesian-Hebbian learning rule as used here (Fiebig and Lansner, 2014). That model included representations for prefrontal cortex, hippocampus, and wider neocortex, implementing an extended complementary learning systems theory (McClelland et al., 1995), which is itself an advancement of systems consolidation (Squire and Alvarez, 1995). We consequently expect that the principled mechanism of semantization explored here can be scaled along the temporal axis to account for lifelong memory, provided that the plasticity involved is itself Bayesian-Hebbian. Our model does not advance any specific anatomic argument as to the location of the respective networks (Yonelinas, 2002; Diana et al., 2007). The model purposefully relies on a generic cortical architecture focused on a class of synaptic plasticity mechanisms which may well serve as a substrate of a wider system across brain areas and time.

In conclusion, we have presented a computational mesoscopic spiking network model to examine the interplay between episodic and semantic memory with the grand objective to explain mechanistically the semantization of episodic traces. Compared with other models of episodic memory, which are typically abstract, our model, built on various biological constraints (i.e., plausible postsynaptic potentials, firing rates, connection densities, synaptic delays, etc.) accounting for neural processes

and synaptic mechanisms, emphasizes the role of synaptic plasticity in semantization. Hence it bridges micro and mesoscale mechanisms with macroscale behavior and dynamics. In contrast to standard Hebbian learning, our Bayesian version of Hebbian learning readily reproduced prominent traces of semantization.

## References

- Atkinson RC, Juola JF (1973) Factors influencing speed and accuracy of word recognition. In: *Attention and performance IV* (Kornblum S, ed), pp 583–612. New York: Academic Press.
- Baddeley A (1988) Cognitive psychology and human memory. *Trends Neurosci* 11:176–181.
- Barth AL, Poulet JF (2012) Experimental evidence for sparse firing in the neocortex. *Trends Neurosci* 35:345–355.
- Behaydt L (1987) The semantization of vocabulary in foreign language learning. *System* 15:55–67.
- Binzegger T, Douglas RJ, Martin KA (2009) Topology and dynamics of the canonical circuit of cat v1. *Neural Netw* 22:1071–1078.
- Bolger DJ, Balass M, Landen E, Perfetti CA (2008) Context variation and definitions in learning the meanings of words: an instance-based learning approach. *Discourse Process* 45:122–159.
- Brette R, Gerstner W (2005) Adaptive exponential integrate-and-fire model as an effective description of neuronal activity. *J Neurophysiol* 94:3637–3642.
- Broadbent H, Osborne T, Mareschal D, Kirkham N (2020) Are two cues always better than one? the role of multiple intra-sensory cues compared to multi-cross-sensory cues in children's incidental category learning. *Cognition* 199:104202.
- Caporale N, Dan Y (2008) Spike timing-dependent plasticity: a Hebbian learning rule. *Annu Rev Neurosci* 31:25–46.
- Chrysanthis N, Fiebig F, Lansner A (2019) Introducing double bouquet cells into a modular cortical associative memory model. *J Comput Neurosci* 47:223–230.
- Deperrois N, Petrovici MA, Senn W, Jordan J (2021) Memory semantization through perturbed and adversarial dreaming. *arXiv* 2109.04261.
- Diana RA, Yonelinas AP, Ranganath C (2007) Imaging recollection and familiarity in the medial temporal lobe: a three-component model. *Trends Cogn Sci* 11:379–386.
- Duff MC, Covington NV, Hilverman C, Cohen NJ (2020) Semantic memory and the hippocampus: revisiting, reaffirming, and extending the reach of their critical relationship. *Front Hum Neurosci* 13:471.
- Egorov AV, Hamam BN, Fransén E, Hasselmo ME, Alonso AA (2002) Graded persistent activity in entorhinal cortex neurons. *Nature* 420:173–178.
- Eichenbaum H (2017) Prefrontal-hippocampal interactions in episodic memory. *Nat Rev Neurosci* 18:547–558.
- Eyal G, Verhoog MB, Testa-Silva G, Deitcher Y, Benavides-Piccione R, DeFelipe J, De Kock CP, Mansvelder HD, Segev I (2018) Human cortical pyramidal neurons: from spines to spikes via models. *Front Cell Neurosci* 12:181.
- Fiebig F, Lansner A (2014) Memory consolidation from seconds to weeks: a three-stage neural network model with autonomous reinstatement dynamics. *Front Comput Neurosci* 8:64.
- Fiebig F, Lansner A (2017) A spiking working memory model based on Hebbian short-term potentiation. *J Neurosci* 37:83–96.
- Fiebig F, Herman P, Lansner A (2020) An indexing theory for working memory based on fast Hebbian plasticity. *eNeuro* 7:ENEURO.0374-19.2020.
- Frankland PW, Josselyn SA, Anagnostaras SG, Kogan JH, Takahashi E, Silva AJ (2004) Consolidation of CS and US representations in associative fear conditioning. *Hippocampus* 14:557–569.
- Friedrich M, Wilhelm I, Born J, Friederici AD (2015) Generalization of word meanings during infant sleep. *Nat Commun* 6:6004.
- Gerstner W, Naud R (2009) How good are neuron models? *Science* 326:379–380.
- Gewaltig MO, Diesmann M (2007) Nest (neural simulation tool). *Scholarpedia* 2:1430.
- Gilboa A (2004) Autobiographical and episodic memory—one and the same?: evidence from prefrontal activation in neuroimaging studies. *Neuropsychologia* 42:1336–1349.
- Greve A, Donaldson DI, Van Rossum MC (2010) A single-trace dual-process model of episodic memory: a novel computational account of familiarity and recollection. *Hippocampus* 20:235–251.
- Gruber MJ, Ritchey M, Wang SF, Doss MK, Ranganath C (2016) Post-learning hippocampal dynamics promote preferential retention of rewarding events. *Neuron* 89:1110–1120.
- Habermas T, Diel V, Welzer H (2013) Lifespan trends of autobiographical remembering: episodicity and search for meaning. *Conscious Cogn* 22:1061–1073.
- Herman PA, Lundqvist M, Lansner A (2013) Nested theta to gamma oscillations and precise spatiotemporal firing during memory retrieval in a simulated attractor network. *Brain Res* 1536:68–87.
- Howard MW, Kahana MJ (2002) When does semantic similarity help episodic retrieval? *J Mem Lang* 46:85–98.
- Johnson JD, McDuff SG, Rugg MD, Norman KA (2009) Recollection, familiarity, and cortical reinstatement: a multivoxel pattern analysis. *Neuron* 63:697–708.
- Kirkcaldie MT (2012) Neocortex. In: *The mouse nervous system*, pp 52–111. San Diego: Elsevier.
- Kuo MF, Paulus W, Nitsche MA (2008) Boosting focally-induced brain plasticity by dopamine. *Cereb Cortex* 18:648–651.
- Lansner A (2009) Associative memory models: from the cell-assembly theory to biophysically detailed cortex simulations. *Trends Neurosci* 32:178–186.
- Lansner A, Ekeberg Ö (1989) A one-layer feedback artificial neural network with a Bayesian learning rule. *Int J Neur Syst* 01:77–87.
- Lundqvist M, Compte A, Lansner A (2010) Bistable, irregular firing and population oscillations in a modular attractor memory network. *PLoS Comput Biol* 6:e1000803.
- Lundqvist M, Herman P, Lansner A (2011) Theta and gamma power increases and alpha/beta power decreases with memory load in an attractor network model. *J Cogn Neurosci* 23:3008–3020.
- Mandler G (1979) Organization and repetition: organizational principles with special reference to rote learning. In: *Perspectives on memory research*. New York: Psychology Press.
- Mandler G (1980) Recognizing: the judgment of previous occurrence. *Psychol Rev* 87:252–271.
- McClelland JL, McNaughton BL, O'Reilly RC (1995) Why there are complementary learning systems in the hippocampus and neocortex: insights from the successes and failures of connectionist models of learning and memory. *Psychol Rev* 102:419–457.
- McCloskey M, Santee JL (1981) Are semantic memory and episodic memory distinct systems? *J Exp Psychol Hum Learn* 7:66–71.
- Morrison A, Diesmann M, Gerstner W (2008) Phenomenological models of synaptic plasticity based on spike timing. *Biol Cybern* 98:459–478.
- Mountcastle VB (1997) The columnar organization of the neocortex. *Brain* 120:701–722.
- Muir DR, Da Costa NM, Girardin CC, Naaman S, Omer DB, Ruesch E, Grinvald A, Douglas RJ (2011) Embedding of cortical representations by the superficial patch system. *Cereb Cortex* 21:2244–2260.
- Norman KA, O'Reilly RC (2003) Modeling hippocampal and neocortical contributions to recognition memory: a complementary-learning-systems approach. *Psychol Rev* 110:611–646.
- Opitz B (2010) Context-dependent repetition effects on recognition memory. *Brain Cogn* 73:110–118.
- Otmakhova NA, Lisman JE (1996) D1/d5 dopamine receptor activation increases the magnitude of early long-term potentiation at ca1 hippocampal synapses. *J Neurosci* 16:7478–7486.
- Panoz-Brown D, Corbin HE, Dalecki SJ, Gentry M, Brotheridge S, Sluka CM, Wu JE, Crystal JD (2016) Rats remember items in context using episodic memory. *Curr Biol* 26:2821–2826.



- Payne JD, Schacter DL, Propper RE, Huang LW, Wamsley EJ, Tucker MA, Walker MP, Stickgold R (2009) The role of sleep in false memory formation. *Neurobiol Learn Mem* 92:327–334.
- Pearson HC, Wilbiks JM (2021) Effects of audiovisual memory cues on working memory recall. *Vision* 5:14.
- Petrican R, Gopie N, Leach L, Chow TW, Richards B, Moscovitch M (2010) Recollection and familiarity for public events in neurologically intact older adults and two brain-damaged patients. *Neuropsychologia* 48:945–960.
- Pokorny C, Ison MJ, Rao A, Legenstein R, Papadimitriou C, Maass W (2020) Stpd forms associations between memory traces in networks of spiking neurons. *Cereb Cortex* 30:952–968.
- Ranganath C (2010) Binding items and contexts: the cognitive neuroscience of episodic memory. *Curr Dir Psychol Sci* 19:131–137.
- Remme MW, Bergmann U, Alevi D, Schreiber S, Sprekeler H, Kempster R (2021) Hebbian plasticity in parallel synaptic pathways: a circuit mechanism for systems memory consolidation. *PLoS Comput Biol* 17:e1009681.
- Ren Q, Kolwankar KM, Samal A, Jost J (2010) Stpd-driven networks and the *c. elegans* neuronal network. *Physica A* 389:3900–3914.
- Renoult L, Irish M, Moscovitch M, Rugg MD (2019) From knowing to remembering: the semantic–episodic distinction. *Trends Cogn Sci* 23:1041–1057.
- Rodríguez TM, Galán AS, Flores RR, Jordán MT, Montes JB (2016) Behavior and emotion in dementia. In: *Update on dementia*, p 449. London: IntechOpen.
- Rubin DC, Wallace WT (1989) Rhyme and reason: analyses of dual retrieval cues. *J Exp Psychol Learn Mem Cogn* 15:698–709.
- Ryals AJ, Cleary AM, Seger CA (2013) Recall versus familiarity when recall fails for words and scenes: the differential roles of the hippocampus, perirhinal cortex, and category-specific cortical regions. *Brain Res* 1492:72–91.
- Shiffrin RM (1970) Memory search. In: *Models of human memory* (Norman DA, ed), pp 375–447. New York: Academic Press.
- Smith SM, Manzano I (2010) Video context-dependent recall. *Behav Res Methods* 42:292–301.
- Smith SM, Handy JD (2014) Effects of varied and constant environmental contexts on acquisition and retention. *J Exp Psychol Learn Mem Cogn* 40:1582–1593.
- Song S, Miller KD, Abbott LF (2000) Competitive Hebbian learning through spike-timing-dependent synaptic plasticity. *Nat Neurosci* 3:919–926.
- Song S, Sjöström PJ, Reigl M, Nelson S, Chklovskii DB (2005) Highly nonrandom features of synaptic connectivity in local cortical circuits. *PLoS Biol* 3:e68.
- Squire LR, Alvarez P (1995) Retrograde amnesia and memory consolidation: a neurobiological perspective. *Curr Opin Neurobiol* 5:169–177.
- Stettler DD, Das A, Bennett J, Gilbert CD (2002) Lateral connectivity and contextual interactions in macaque primary visual cortex. *Neuron* 36:739–750.
- Sun Q, Gu S, Yang J (2018) Context and time matter: effects of emotion and motivation on episodic memory overtime. *Neural Plast* 2018:7051925.
- Thomson AM, West DC, Wang Y, Bannister AP (2002) Synaptic connections and small circuits involving excitatory and inhibitory neurons in layers 2–5 of adult rat and cat neocortex: triple intracellular recordings and biocytin labelling in vitro. *Cereb Cortex* 12:936–953.
- Tsodyks MV, Markram H (1997) The neural code between neocortical pyramidal neurons depends on neurotransmitter release probability. *Proc Natl Acad Sci U S A* 94:719–723.
- Tully PJ, Hennig MH, Lansner A (2014) Synaptic and nonsynaptic plasticity approximating probabilistic inference. *Front Synaptic Neurosci* 6:8.
- Tully PJ, Lindén H, Hennig MH, Lansner A (2016) Spike-based Bayesian-Hebbian learning of temporal sequences. *PLoS Comput Biol* 12:e1004954.
- Tulving E (1972) 12. Episodic and semantic memory. In: *Organization of memory* (Tulving E, Donaldson W, eds), pp 381–403. New York: Academic Press.
- Van Rossum MC, Bi GQ, Turrigiano GG (2000) Stable Hebbian learning from spike timing-dependent plasticity. *J Neurosci* 20:8812–8821.
- Viard A, Piolino P, Desgranges B, Chételat G, Lebreton K, Landeau B, Young A, De La Sayette V, Eustache F (2007) Hippocampal activation for autobiographical memories over the entire lifetime in healthy aged subjects: an fMRI study. *Cereb Cortex* 17:2453–2467.
- Wahlgren N, Lansner A (2001) Biological evaluation of a Hebbian-Bayesian learning rule. *Neurocomputing* 38–40:433–438.
- Wais PE, Mickes L, Wixted JT (2008) Remember/know judgments probe degrees of recollection. *J Cogn Neurosci* 20:400–405.
- Wang Y, Markram H, Goodman PH, Berger TK, Ma J, Goldman-Rakic PS (2006) Heterogeneity in the pyramidal network of the medial prefrontal cortex. *Nat Neurosci* 9:534–542.
- Watrous AJ, Tandon N, Conner CR, Pieters T, Ekstrom AD (2013) Frequency-specific network connectivity increases underlie accurate spatiotemporal memory retrieval. *Nat Neurosci* 16:349–356.
- Weidemann CT, Kragel JE, Lega BC, Worrell GA, Sperling MR, Sharan AD, Jobst BC, Khadjevand F, Davis KA, Wanda PA, Kadel A, Rizzuto DS, Kahana MJ (2019) Neural activity reveals interactions between episodic and semantic memory systems during retrieval. *J Exp Psychol Gen* 148:1–12.
- Wixted JT (2007) Dual-process theory and signal-detection theory of recognition memory. *Psychol Rev* 114:152–176.
- Yonelinas AP (2002) The nature of recollection and familiarity: a review of 30 years of research. *J Mem Lang* 46:441–517.
- Yonelinas AP, Jacoby LL (1994) Dissociations of processes in recognition memory: effects of interference and of response speed. *Can J Exp Psychol* 48:516–535.
- Yonelinas AP, Aly M, Wang WC, Koen JD (2010) Recollection and familiarity: examining controversial assumptions and new directions. *Hippocampus* 20:1178–1194.
- Yoshimura Y, Callaway EM (2005) Fine-scale specificity of cortical networks depends on inhibitory cell type and connectivity. *Nat Neurosci* 8:1552–1559.