

# A Bayesian context fear learning algorithm/automaton

Franklin B. Krasne<sup>1,2\*</sup>, Jesse D. Cushman<sup>1,2</sup> and Michael S. Fanselow<sup>1,2,3</sup>

<sup>1</sup> Department of Psychology, University of California Los Angeles, Los Angeles, CA, USA, <sup>2</sup> Brain Research Institute, University of California Los Angeles, Los Angeles, CA, USA, <sup>3</sup> Department of Psychiatry and Biobehavioral Sciences, University of California Los Angeles, Los Angeles, CA, USA

Contextual fear conditioning is thought to involve the synaptic plasticity-dependent establishment in hippocampus of representations of to-be-conditioned contexts which can then become associated with USs in the amygdala. A conceptual and computational model of this process is proposed in which contextual attributes are assumed to be sampled serially and randomly during contextual exposures. Given this assumption, moment-to-moment information about such attributes will often be quite different from one exposure to another and, in particular, between exposures during which representations are created, exposures during which conditioning occurs, and during recall sessions. This presents challenges to current conceptual models of hippocampal function. In order to meet these challenges, our model's hippocampus was made to operate in different modes during representation creation and recall, and non-hippocampal machinery was constructed that controlled these hippocampal modes. This machinery uses a comparison between contextual information currently observed and information associated with existing hippocampal representations of familiar contexts to compute the Bayesian Weight of Evidence that the current context is (or is not) a known one, and it uses this value to assess the appropriateness of creation or recall modes. The model predicts a number of known phenomena such as the immediate shock deficit, spurious fear conditioning to contexts that are absent but similar to actually present ones, and modulation of conditioning by pre-familiarization with contexts. It also predicts a number of as yet unknown phenomena.

**Keywords:** fear conditioning, hippocampus, amygdala, model, Bayesian, context, computational neuroscience

## OPEN ACCESS

### Edited by:

Jeansok J. Kim,  
University of Washington, USA

### Reviewed by:

June-Seek Choi,  
Korea University, South Korea  
Brian Witgen,  
University of California, Davis, USA

### \*Correspondence:

Franklin B. Krasne,  
Department of Psychology, University  
of California Los Angeles, Box  
951563, Los Angeles, CA  
90095-1563, USA  
krasne@psych.ucla.edu

**Received:** 14 January 2015

**Accepted:** 16 April 2015

**Published:** 27 May 2015

### Citation:

Krasne FB, Cushman JD and  
Fanselow MS (2015) A Bayesian  
context fear learning  
algorithm/automaton.  
*Front. Behav. Neurosci.* 9:112.  
doi: 10.3389/fnbeh.2015.00112

## Introduction

During Pavlovian fear conditioning, animals become afraid of both specific cues that predict the imminent onset of aversive events such as foot shock, and the situation or “context” in which the shock occurred. Fear conditioning provides one of neuroscience’s most promising and active arenas for analyzing neural mechanisms of learning, generally. Both cue and context fear conditioning seem to be due to plastic change at synapses within the amygdala (reviewed by Fanselow and LeDoux, 1999; Blair et al., 2001), but there are striking differences in the phenomenology and the neural circuitry of cue and context fear conditioning. Anatomically, cued fear seems to involve pathways from thalamus and cortex directly to the lateral nucleus of amygdala (LeDoux and Clugnet, 1990; LeDoux et al., 1990; Romanski and LeDoux, 1992; Boatman and Kim, 2006), whereas context fear utilizes a pathway from cortex to hippocampus to basal amygdala (Young et al., 1994; Maren and Fanselow, 1995;

Bast et al., 2003; Matus-Amat et al., 2004; Calandreau et al., 2005; Parsons and Otto, 2008; Schenberg and Oliveira, 2008; Onishi and Xavier, 2010), though the exact route of this latter pathway remains under investigation (Fanselow and Dong, 2010). Moreover, development of context fear depends not only on synaptic plasticity within amygdala but also within hippocampus (Kiernan and Cranney, 1992; Fanselow, 2000; Rudy and O'Reilly, 2001; Matus-Amat et al., 2004; Stote and Fanselow, 2004; Rudy, 2009). Phenomenologically, fear becomes conditioned to a cue that warns of shock any time that the cue occurs. But fear of context does not develop if shocks occur too soon after an animal is introduced into a situation; the shock must be delayed by at least some 30 s, and full conditionability does not develop for several minutes (the so-called "immediate shock deficit") (Blanchard et al., 1976; Fanselow, 1986, 1990; Landeira-Fernandez et al., 2006).

That there should be substantial differences between cue and context fear conditioning is not surprising, because learning to fear a context presents the nervous system with problems that do not occur during conditioning to simple warnings of a US. Contexts are defined by stable configurations or "conjunctions" of elementary cues (referred to here as "attributes") that are present over extended periods of time. It can be expected that brains come to a learning situation pre-wired to represent the simple cues that, at least in the laboratory, are usually used as CSs for imminent USs. Such cues may be made to evoke fear simply by strengthening synapses between these representation neurons and amygdala fear-producing cells (see Blair et al., 2001).

However, it cannot be expected that neurons representing novel contexts, qua contexts, pre-exist. Neurons that innately represent the simple attributes which compose a context might of course come to evoke fear, but conditioning based on them would have unsatisfactory properties: Animals can presumably perceive only a small subset of a context's attributes at any one moment, and at each new entry to a given context, somewhat different attributes are likely to be sampled. Conditioning established to the small set of attributes that happened to be represented neurally at the moment of US occurrence would not produce fear during a different sampling of the same context. Therefore, stable contextual fear conditioning would seem to require the prior creation of a neural representation of the to-be-conditioned context that is constant despite variable samplings of the context's attributes. Also, there are likely to be many attributes of a context that are common to multiple contexts; if context fear were evoked by attributes *per se*, the existence of common attributes might well promote non-adaptive over-generalization. These problems can be avoided if conditioning occurs to a representation that is activated by a conjunction of contextual attributes rather than the individual attributes themselves. According to a variety of lines of current thinking, the role of the hippocampus in learning is to rapidly create conjunctive representations of combinations of simpler cues that then act as the neural stimulus for learnt contextual fear as well other kinds of learning (Fanselow, 1986, 2000; Kiernan and Cranney, 1992; Rudy and O'Reilly, 2001; Matus-Amat et al., 2004; Stote and Fanselow, 2004; Rudy, 2009).

Current theoretical thinking about the role of the hippocampus in learning stems from the work of Marr (1971). He proposed that, for a variety of reasons, the cortex

learns slowly. But obviously animals must be able to learn some things rapidly. He argued that the intrinsic connectivity of the hippocampus and its connections to the cortex were consistent with the hypothesis that the hippocampus rapidly stores information about cortical patterns of activity. It does this, he suggested, in such a way that any sufficient fraction of the cortical activity pattern that was present at learning can evoke both the complete hippocampal pattern and, via that, the full cortical pattern that led to the learning. The fixed hippocampal patterns of Marr and followers theories (Marr, 1971; Skaggs and McNaughton, 1992; O'Reilly and McClelland, 1994; McClelland et al., 1995; O'Reilly and Rudy, 2001; Scharfman, 2007; Myers and Scharfman, 2011; Ketz et al., 2013) are just what are needed for representations to drive contextual fear responses. They are constant despite partial and variable sensory input, and since they are relatively non-overlapping ("orthogonal"), even for quite similar contexts, they would not lead to untoward generalization.

However, the fact that the nervous system does not receive information about all the attributes of a context at once means that over time there must be a continual evaluation and re-evaluation of the evidence as to whether the set of contextual cues so far sampled is from a familiar context or from a new context for which a representation should be established. This is essentially a Bayesian inference problem. The present model assumes a cortical-hippocampal circuit fairly similar to those of Marr and his followers. But it uses the Bayesian weight of evidence (Kass and Rafter, 1995; Gallistel, 2009) that the current context is or is not a known one to control learning within the hippocampal/cortical circuit and the hippocampus-amygdala pathway. We thus refer to the model as the Bayesian Context Fear Algorithm/Automaton (BACON).

## Methods

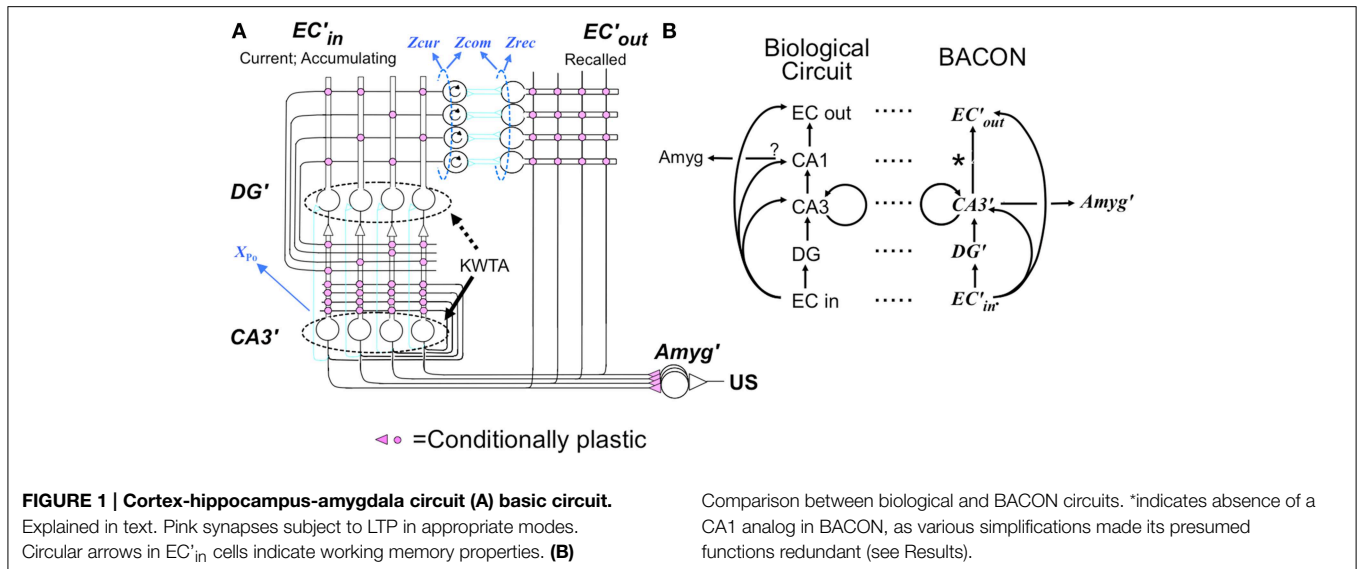
Most details not needed to understand Results are deferred to Supplementary Material.

### Scope

The goal of this work was to create a rational and neurally plausible algorithm for context fear conditioning. The development over time of hippocampus-independent engrams (systems type consolidation), was not considered but it is planned to address it at a later time.

### BACON Neurons

The resting potential of BACON neurons is taken as zero, and membrane potentials ( $V$ ) are specified relative to this baseline. Active synapses are thought of as causing post-synaptic conductances for specific ions. Excitatory input moves  $V$  toward an excitatory reversal potential  $E$ . Active synaptic inputs to a neuron cause post-synaptic conductances given by  $G = A \cdot W$  where  $G$  is conductance relative to membrane leakage conductance,  $W$  the synapse's strength or "weight," which can vary from zero to any positive value, and  $A$  the activity level (firing rate, or "activation") of the presynaptic neuron. The net depolarization of a neuron receiving both excitatory and inhibitory input is given by:



**TABLE 1 | Properties of BACON neurons.**

Cell type	Properties
EC'in	Binary. Working memory properties-see Results
EC'out	Binary. Thrsh = ThrshCtx
DG'	KWTA activation rule in creation and update modes; rate coded in recall mode
CA3'	KWTA activation rule
Amyg'	Rate coded

$$V = \left( \frac{Ge \cdot E}{1 + Ge + Gi} \right) \tag{1}$$

where Ge is total excitatory and Gi total inhibitory conductance. Some BACON neurons are binary, firing at a maximum rate of unity when V is above some threshold value. For the remainder, activation rises linearly from zero starting at some threshold voltage (thrsh) and reaching a maximum of unity at a higher voltage (mxat). We refer to this as a “linear sigmoid activation function,” which we write as follows:

$$Activation = linsig(V|thrsh, mxat). \tag{2}$$

The basic cortical-hippocampal-amygdala circuit of BACON is shown in **Figure 1**, which will be explained in Results. **Table 1** gives salient properties of the various classes of neurons. As explained in Results, algorithm neuron types are named for their intended biological parallels but with a prime mark added to distinguish them from their biological counterparts.

Most of the computations done by the algorithm are explained in logical or mathematical terms. Actual neural circuitry that could underlie them is unknown, but circuitry that could compute the Bayesian Weight of Evidence, which is central to the model, and circuitry that could use it to control hippocampal and amygdala function are sketched out to demonstrate that

they could be carried out by neurons. Ways of carrying out arithmetical operations such as subtraction, multiplication, exponentiation, etc. assumed done within this hypothetical circuitry and ways of implementing K-Winners-Take-All control of neural firing, used extensively in BACON, are discussed in the Supplementary Material (Topic H).

The values and definitions of BACON’s parameters are summarized for reference in **Table 2**.

### Computations

A flow diagram of the computations done by BACON is provided in Supplementary Material (Topic G).

## Results

The goal of this work was to construct a biologically plausible algorithm/automaton for context fear conditioning. It is assumed that upon entering a context its attributes are observed (“sampled”) serially in random order without replacement over an extended period of time and are held in a working memory. As attributes are sampled, we wanted the automaton to compare them to the recalled attributes of previously experienced contexts and use the comparison to decide whether the context is a known one, in which case existing associations to it should be expressed and new ones made, or a novel one, in which case a neural representation of it should be created and further associations made to that. The algorithm we have created deals with these matters using analogs of cortical, hippocampal, and amygdala structures; we name these for their biological counterparts *appended with a prime (') mark*.

### The Basic Cortical-Hippocampal-amygdala Circuit Structure

**Figure 1A** presents BACON’s basic circuit, which as indicated in **Figure 1B**, is formally similar to its biological counterpart.

**TABLE 2 | BACON parameter definitions and values.**

Numbers of EC <sub>in</sub> and EC <sub>out</sub> cells	$N_{Ctx}$	1000
Numbers of DG and CA3' cells	$N_{Hipp}$	10,000
Number of attributes per context	$N_A$	100
Number of attributes common to all contexts ("General" attributes)	$N_{Gen}$	50
Number of EC <sub>in</sub> neurons innervating each DG neuron	$F$	60
Number of winners of hippocampal KWTA calculations	$K$	60
Minimum number of cells that must be active in CA3 pattern $CA3'Ptn_o$ for recurrent collateral input to be computed	$K_0$	15
Minimum number of EC <sub>in</sub> cells that must be active for a hippocampal representation to be created	$Z_0(minZ0)$	45
Number of effective inputs needed for a EC <sub>out</sub> cell to fire	$ThreshCtx$	42
Minimum Bayesian weight of evidence for conditioning to occur	$B_{Old}$	3
Negative Bayesian weight of evidence sufficient for representation creation	$B_{New}$	-3
Minimum Bayesian weight of evidence for addition of attributes to an existing hippocampal representation	$B_{add}$	15
Bayesian weight of evidence making representation "probably valid"	$B_{pv}$	3
Amygdala learning rate parameter	$\alpha$	0.025
Factor multiplying excitation produced in CA3' via direct pathway	$d_{pf}$	Varied

As in most models inspired by Marr's original work, BACON's hippocampal circuitry is designed to construct a cortical "image" of previously experienced events when probed by a sufficient fraction of the image. It does this by creating a sparse, random hippocampal representation of the event that can be fully activated by fragmentary cortical images, and that can then in turn reactivate all those cortical neurons that were active at the time that the hippocampal representation was created. In designing BACON's hippocampus, we attempted to implement the presumed basic operational principles of the biological hippocampus in as clear and simple a form as possible. We did not include features in the name of fidelity if they did not seem useful to the task at hand.

### Entorhinal cortex

Cortical input is presented to BACON's hippocampus' by an entorhinal cortex (EC') that consists of  $N_{Ctx}$  (=1000) input (EC<sub>in</sub>) neurons, each of which innately represents one of the possible contextual attributes that the automaton can detect. Hippocampal output back to cortex is via a corresponding set of output (EC<sub>out</sub>) cells. Each context is composed of  $N_A$  (=100) attributes of which a number,  $N_{Gen}$  (=50) are general to all contexts. When an attribute is sampled the EC<sub>in</sub> cell that represents it becomes active and remains so for the duration the visit to that context. Neurons that remain active during a working memory task have in fact been found in entorhinal cortex (e.g., Suzuki et al., 1997; Young et al., 1997; Schon et al., 2015).

### Dentate (DG')

DG' consists of a large number,  $N_{Hipp}$  (=10,000), of projecting neurons that during representation creation recode the EC<sub>in</sub> pattern of activity into a sparse, relatively non-overlapping form (as in O'Reilly and McClelland, 1994). Each DG' cell is innervated, via Hebbian synapses, by  $F$  (=60) EC<sub>in</sub> neurons. *All of BACON's Hebbian synapses are totally ineffective until potentiated.* During representation creation, firing of the projecting DG' neurons is subject to regulation by an inhibitory network that allows only a specified number of the most excited

cells to fire ("K-Winners Take All" (KWTA) behavior). When operative, this allows  $K$  =60 of DG's 10,000 cells to fire. At other times, DG' cells fire at a rate proportional to their excitation.

### CA3'

As in the biological hippocampus, CA3' cells receive input from EC<sub>in</sub> both via DG' (the "indirect" pathway) and via a direct pathway from EC<sub>in</sub>. As was the case for DG', each CA3' cell is innervated by  $F$  randomly chosen EC<sub>in</sub> neurons, but these are chosen independently of the DG' innervation. In the biological hippocampus each CA3 cell is innervated by a small number of dentate cells, and the synapses, which appear from morphology to be strong, show some non-Hebbian sort of plasticity of uncertain function (Bortolotto et al., 2003; Reid et al., 2004). In BACON this arrangement has been simplified (as considered in Discussion); each CA3' cell is innervated by a single DG' neuron via a non-plastic synapse whose strength can be modulated. CA3' cells, like DG' cells, are subject to KWTA regulation of their firing; the number of cells allowed to fire is the same as in DG'. In the biological hippocampus, each CA3 cell recurrently innervates other CA3 cells via Hebb synapses. This is thought to form an auto-associative memory that binds together the member-neurons of each representation so that when a sufficient subset is activated the rest will follow (e.g., Gardner-Medwin, 1976; McNaughton and Morris, 1987; Treves and Rolls, 1994). In the biological hippocampus each CA3 cell innervates only a portion of others; in BACON we have for simplicity allowed full innervation.

### CA3', CA1', and EC<sub>out</sub>

Entorhinal cortex receives from hippocampus representations of contexts that consist of the activity of essentially random sets of cells. It is thought that these are able to recreate entorhinal output replicas of the entorhinal input patterns that they encode because during representation creation potentiation develops between cells of the 'random' hippocampal representation and EC output cells corresponding to EC input cells that are active (e.g., Treves and Rolls, 1994).



It seems as if this function would be best served if each CA3 cell were to directly innervate each entorhinal output neuron via a Hebb synapse, which is exactly how BACON is constructed. However, fan-ins and fan-outs of most biological neurons are too low for each real CA3 neuron to completely innervate every EC output cell. The interposition of CA1 between CA3 and EC in the biological Hippocampus is thought to mitigate this limitation by recoding the CA3 output into a less sparse form in CA1 so that every EC output cell can receive some innervation that is specific to each encoded pattern (e.g., Kesner and Rolls, 2015). A second function of biological CA1 is thought to be to help complete representations that, because of incomplete innervation in the recurrent collateral network, were incompletely reconstructed by CA3 auto-associative mechanisms (Treves, 1995). However, because in BACON we allow complete innervation within the CA3 recurrent collateral system, this function is also unneeded. Since a CA1' would therefore have served no useful function in BACON, it was for simplicity omitted.

### Amygdala'

Although we assume that hippocampal circuitry evolved to allow reconstruction of the neocortical activity that occurred during past events, *it is to the hippocampal code proper and not to the cortical reconstruction of a context's attributes* that fear becomes conditioned in BACON's amygdala'. We constructed BACON in this way for three reasons: (i) We believe that it is consistent with the weight of available evidence (as discussed in Krasne et al., 2011). In particular, it appears from the immediate shock deficit that fear does not readily become conditioned to the attributes of a context *per se*. (ii) Hippocampal representations of very similar contexts overlap much less than do cortical ones, which makes discriminations easier to learn. (iii) Hippocampal representations are (by construction) constant whereas cortical ones change as more is learned about a context; changing representations would lead to unwanted variations in expression of previously conditioned fear.

The actual (biological) pathway via which information about active hippocampal representations reaches the amygdala from the dorsal hippocampus is a matter of current research. Based on various lines of anatomical, and more recently labeling and optogenetic data, ventral hippocampus as well as various cortical structures have been suggested as possibly being on the relevant pathway Maren and Holt, 2004; Fanselow and Dong, 2010, Cowansage et al., 2014; Tanaka et al., 2014; Jin and Maren, 2015. Whatever the final resolution of this matter, BACON assumes (as said above), that it is the dorsal hippocampal representation (or downstream, possibly cortical, versions of it), and not cortical attribute-representing neurons themselves, to which contextual fear responses become conditioned.

CA3' innervates amygdala' fear-producing cells via Hebb synapses, and conditioned fear develops when synapses between active CA3' cells and amygdala' neurons that are depolarized by an unconditioned stimulus become potentiated.

### Operational Modes

The above features of BACON are all equivalent to, or slight modifications of, conventional ideas about the role of the

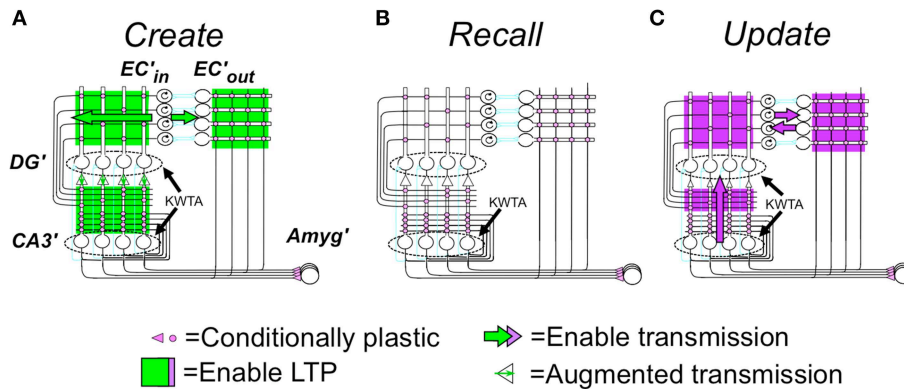
hippocampus in creating and utilizing representations of complex stimuli and in context fear conditioning. However, BACON departs from most conventional thinking in that it includes circuitry conceptualized as outside the hippocampus that makes evidence-based decisions as to whether the automaton's current context is or is not a familiar one, and depending on its conclusions, configures the hippocampus and amygdala appropriately for representation *creation* or *recall* as well as controlling *conditionability* and the *expression* of previously conditioned fear. This control is exerted by enabling or disabling the possibility of Hebbian potentiation, modulating the strength of certain synapses, and turning on or off the KWTA capability of those regions that are potentially subject to it. The idea that hippocampal circuitry is subject to task-based reconfiguration has precedents in the work of O'Reilly, McClelland, Rudy, and Hasselmo (e.g., O'Reilly and McClelland, 1994; Hasselmo et al., 1995; O'Reilly and Rudy, 2001; Hasselmo and Sarter, 2010).

### Representation creation

When BACON decides it is in a novel context, it configures the circuitry of **Figure 1** for representation *creation* (**Figure 2A**). This includes (1) enabling the plasticity of all the Hebbian synapses of hippocampus and cortex, (2) enabling transmission at all EC'<sub>in</sub>-DG', EC'<sub>in</sub>-CA3', and EC'<sub>in</sub>-EC'<sub>out</sub> synapses (irrespective of the state of potentiation of the plastic synapses), and (3) enabling the KWTA circuitry of DG' (the KWTA circuitry of CA3' is always operative). Finally, the strength of transmission at DG' to CA3' synapses is up-modulated sufficiently so that the *K* cells that will fire in CA3' will be the partners of the *K* winners of the DG' KWTA competition and not be affected by direct EC'<sub>in</sub> input.

Under these conditions the *K* DG' cells that are most strongly excited by the active EC'<sub>in</sub> attribute cells fire and drive their CA3' partners. The synapses between all active EC'<sub>in</sub> cells and the active DG'-CA3' partners then become potentiated. It should be noted that this method of determining which DG' and CA3' cells come to represent a context ensures that the representations will be composed of those DG'-CA3' "dyads" whose DG' member is most strongly *innervated* by the EC'<sub>in</sub> cells representing the attributes that had been sampled up to the moment of the representations creation.

Once the *K* most excited CA3' neurons fire, the activity of each active CA3' cell propagates into its dendrites and collaterals, and the recurrent synapses between one active CA3' cell and another potentiate. The *permanent* representation of the current context has at that point been established as the active set of DG'-CA3' dyads. It is the potentiated recurrent collateral synapses that bind together the set of cells that comprise a representation, and it is the potentiated synapses in the direct and indirect pathways that are responsible for the ability of a given context's attributes to call out its proper representation. Finally, the synapses between the active CA3' representation neurons and the set of active EC'<sub>out</sub> attribute cells that are being driven by the active EC'<sub>in</sub> cells become potentiated so that in the future when this representation becomes active it will cause those same EC'<sub>out</sub> cells to fire.



**FIGURE 2 | (A) Creation, (B) Recall, and (C) Update modes.** Explained in text but some details follow: **(A)** Creation mode details: (1) Noise in the system prevents any ties of excitation of DG' cells during representation creation, so *exactly*  $K$  DG' cells fire. (2) Potentiation of synaptic weights of EC'in synapses on each DG' or CA3' cell occurs in such a way that the weights of potentiating synapses sums to one; such normalization of weights is needed so that the representations of contexts created with larger numbers of EC'in neurons active do not dominate in later KWTA competitions. Other synapses potentiate in one step to fixed weights. **(B)** Recall mode details: (1) During production of the pre-recurrent CA3' activity pattern  $CA_3P_{trn_0}$ , noise in the system prevents any ties of excitation, so *at most*, *exactly*  $K$  DG' cells fire. If the number of active neurons in  $CA_3P_{trn_0}$  is less than  $K_0$ , CA3' activity is extinguished until the next attribute is sampled. (2) The excitation evoked by recurrent collateral activity causes the  $K$  most excited CA3' cells to fire to produce  $CA_3P_{trn_1}$  and higher order patterns with ties *not* being broken, so more than  $K$  cells may fire; allowance of more than  $K$  cells is necessary for convergence to the correct pattern (see Supplementary Material, Topic B). (3) When evidence for two patterns is exactly equal,  $CA_3P_{trn_{Fnl}}$  will contain more than  $K$  active cells, and further downstream actions will be suppressed. **(C)** Update mode specifics: (1)

EC'in-EC'out transmission is enabled bi-directionally (purple top block arrows) so that the representations of all currently sampled and recalled attributes will be active in both EC'in and EC'out. (2) Transmission in a powerful "backward" pathway from CA3' to DG' (up-pointing purple block arrow) and KWTA control of DG' firing are enabled. As a result, DG' activity mirrors  $CA_3P_{trn_1}$ , which will be the complete representation with no other cells active. (3) EC'in, DG', CA3' and CA3'-EC'out synapses become LTP-susceptible, and all coactive EC'in-DG' and CA3' synapses and all CA3'-EC'out synapses become potentiated if they are not already so. As in representation creation, synaptic weights of active EC'in synapses on each active DG' and CA3' cell are normalized to a sum of one. \*Information flow in the biological dentate-CA3 pathway is usually considered to be only in the forward direction, and therefore the above use of "backward" transmission may seem biologically implausible. However, a backward pathway from CA3 to dentate has in fact been described (Scharfman, 2007; Myers and Scharfman, 2011), and there is also some evidence for retrograde chemical signaling across synapses in the forward pathway. Contractor et al. (2002) that could imaginably trigger LTP at presynaptically active synapses on dentate neurons innervating active CA3 cells. We therefore do not consider this computationally motivated feature biologically implausible.

It should be noted that, as will be explained below, representation creation generally occurs when only a portion of a context's attributes have been sampled. Therefore, even if two contexts are extremely similar, their representations will be created with somewhat different sets of EC'in attribute cells active.

In some biological experiments to be discussed, DG' was inactivated during encoding. If this is done in BACON, it will be the KWTA properties of CA3' operating on input via the direct path that will determine which CA3' neurons comprise the new representation.

### Recall

In the resting state, the cortical and hippocampal parts of **Figure 1** are configured for *recall*. In this mode (**Figure 2B**) no synaptic plasticity is enabled, and plastic synapses operate according to their current state of potentiation. When in recall mode, the ratio of the strength of direct to indirect pathway synapses on CA3' cells is a parameter  $d_{pf}$  (direct path factor), the effect of whose value we study below. KWTA control is disabled in DG' during recall, with the result that cells there fire at rates proportional to their excitation; however KWTA control is always operative in CA3'.

Under these conditions active EC'in neurons excite DG' and CA3' cells via any previously potentiated synapses. The DG'

cells fire in proportion to their excitation, exciting CA3' cells proportionately, and the direct input from EC'in sums with the excitation from DG'. The KWTA rule then determines CA3' firing. The pattern of activity at this stage we call  $CA_3P_{trn_0}$ .

Once sampling of a context's attributes is fairly far advanced, the strongly excited cells included in  $CA_3P_{trn_0}$  will mostly be ones that were active when a context's representation was created. However, earlier in sampling, when only a modest fraction of the total set of a context's attributes have been sampled, it can easily occur that representation cells that do *not* represent the current context will be part of  $CA_3P_{trn_0}$ . This happens because, as mentioned above, when two contexts have attributes in common (which they always do, since  $N_{Gen}$  is 50% of  $N_A$ ), which of these get encoded when their representations are created is usually somewhat arbitrary. Therefore, during recall in one of the contexts, shared attributes that were encoded for the other context may well be the ones first sampled and may cause representation cells unique to this (incorrect) context to be activated.

Next,  $CA_3P_{trn_0}$  will generate a new pattern of CA3' excitation via the potentiated synapses of the recurrent collateral system, and the  $K$  most excited cells will again fire. This process repeats a specified number of times until a final pattern referred to as  $CA_3P_{trn_{Fnl}}$  results. In BACON, where each CA3' cell recurrently

innervates each other one, a  $CA3P_{trn_o}$  in which a plurality of the cells are part of some representation will, after two iterations generate a  $CA3P_{trn_{Fnl}}$  in which the cells active are the  $K$  representation cells for that context and no others (so long as not too many extremely similar contexts have been encoded – see Supplementary Material, Topic B). If no context's representation cells are in a plurality,  $CA3P_{trn_{Fnl}}$  will consist of more than  $K$  cells (a summation of more than one representation), and in that case no downstream activity will be produced (see **Figure 2**, caption and Supplementary Material, Topic B).

Finally,  $CA3P_{trn_{Fnl}}$  excites  $EC'_{out}$  cells via any potentiated synapses, and those  $EC'_{out}$  neurons that get excited by a sufficient proportion ( $ThrshCtx$ ) of the active  $CA3'$  cells fire.

### Representation updating

If, after sampling a new attribute, BACON determines that it has correctly identified the context it is in, or it has created a new representation during the current session, it adds the newly sampled attribute to the currently active contextual representation. It does this by going into *update* mode (**Figure 2C**). In this mode, once  $CA3P_{trn_{Fnl}}$  has been activated, backward transmission from  $CA3'$  to  $DG'$  is enabled so that  $DG'$  mirrors  $CA3'$  activity. Transmission between  $EC'_{in}$  and  $EC'_{out}$  is also bi-directionally enabled so that the  $EC'_{in}$ - $EC'_{out}$  pairs corresponding to all currently sampled and recalled attributes are active. And plasticity is enabled at  $EC'_{in}$ - $DG'$  and  $-CA3'$  and  $CA3'$ - $EC'_{out}$  synapses. These actions associate the newly sampled attribute with the current representation, while keeping constant the set of  $DG'$ - $CA3'$  pairs that constitute the representation.

### Context fear conditioning

Contextual fear conditioning in BACON is due to potentiation of  $CA3'$ -amygdala' synapses and is explained at the end of the next section.

### Bayesian Control of Operation Mode of Hippocampo-Cortico-Amygdala Circuitry

Based on comparisons of current  $EC'_{in}$  activity, the status of which depends on the outcome of BACON's sampling of the current context so far, and  $EC'_{out}$  activity, which depends on the automaton's cumulative past experience, BACON determines the Bayesian weight of evidence (Kass and Rafter, 1995; Gallistel, 2009, here denoted as  $B_{Rep}$ ) that an active hippocampal representation really is that of the current context. It then uses this value to decide whether to create a new representation or update the current one, whether to permit fear conditioning if a US occurs, and whether to allow expression of conditioned fear.

### Evaluating Representation Validity

The calculation of  $B_{Rep}$  is based on (1) the number of attributes of the current context that have been sampled and have active representations in  $EC'_{in}$  (referred to as  $Z_{cur}$  –for “ $Z_{current}$ ”), (2) the number of attribute representations activated in  $EC'_{out}$  by the current hippocampal representation ( $Z_{rec}$ –for “ $Z_{recalled}$ ”), and (3) the number of active attribute representations in common between  $EC'_{in}$  and  $EC'_{out}$  ( $Z_{com}$ –for “ $Z_{common}$ ”).  $B_{Rep}$  is defined as:

$$B_{Rep}(Z_{com}|Z_{cur}, Z_{rec}) = \log_{10} \frac{P[Z_{com}|Same, Z_{cur}, Z_{rec}]}{P[Z_{com}|Diff, Z_{cur}, Z_{rec}]} \quad (3)$$

where “Same” means that the active contextual representation is that of the actual context and “Diff” means that it is not.

**Figure 3A** graphs  $B_{Rep}$  as a function of  $Z_{com}$  for a range of values of  $Z_{cur}$  and  $Z_{rec}$ ; the combinatorial formulas used to calculate these values are given in Supplementary Material (Topic C). High positive values mean that the specified degree of communality between  $EC'_{in}$  and  $EC'_{out}$  activity makes it very likely that the actual current context is in fact the one that was present when the currently active hippocampal representation was created. A large negative value indicates that the actual current context is not the one currently represented by the hippocampus. A value of zero means that there is no basis for believing either hypothesis. We note that  $B_{Rep} = 0$  if  $EC'_{out}$  remains silent, as it does when the number of cells active in the initial  $CA3'$  pattern of the computation cycle ( $X_{p_0}$ ) is less than  $K_0$  (the number active cells in  $CA3P_{trn_o}$  needed to activate the recurrent collaterals).

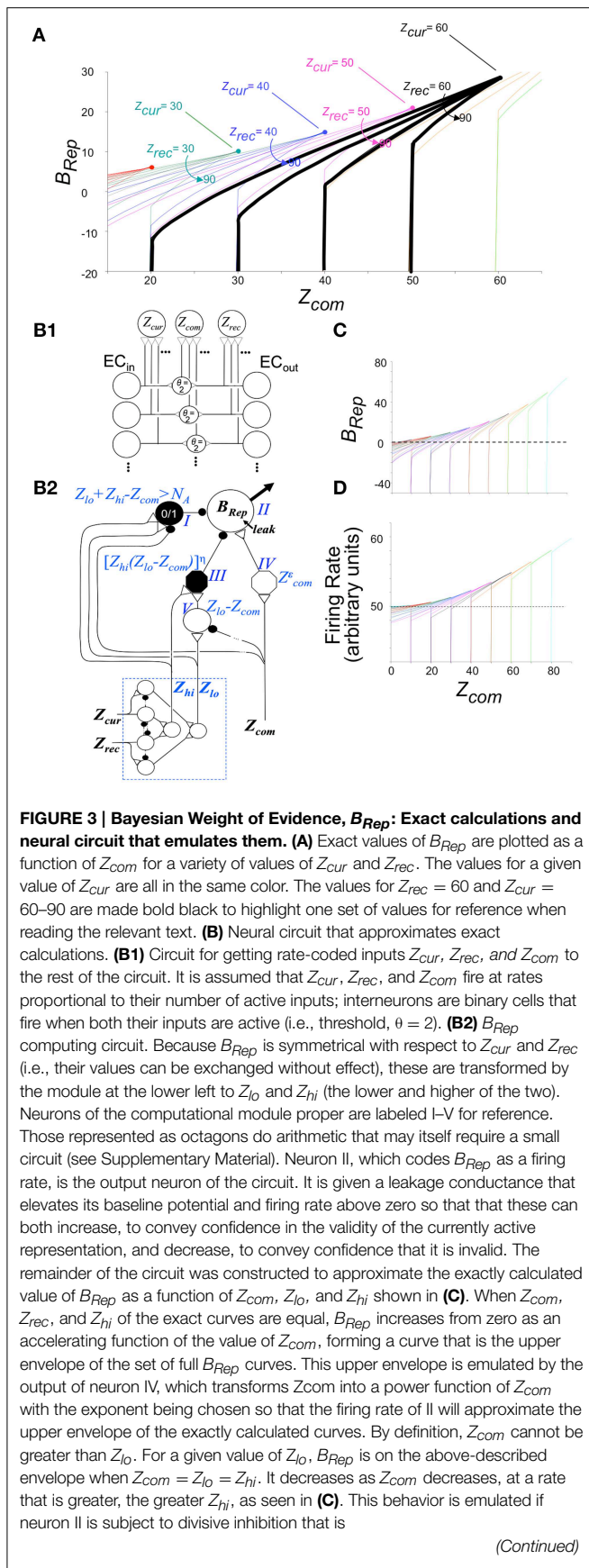
The relationships shown in **Figure 3A** conform well with logical or intuitive judgments based on the same information. Consider for example the bold black curves for  $Z_{cur} = 60$ . (1) For any given number of current and recalled attributes, the greater  $Z_{com}$ , the more likely it is that the currently active representation is in fact that of the current context, and correspondingly,  $B_{Rep}$  increases as a function of  $Z_{com}$ . (2) The more one recalls about the context one thinks one is in, the more strongly a low value of  $Z_{com}$  implies that one is actually somewhere else, and conversely. Consistent with this, for a given value of  $Z_{com}$ ,  $B_{Rep}$  decreases as  $Z_{rec}$  increases. (3) If the number of currently observed attributes that are different from ones remembered ( $Z_{cur}-Z_{com}$ ) plus the number remembered ( $Z_{rec}$ ) is greater than the number of attributes there are per context ( $N_A$ ), then the current context cannot be valid. Consistent with this,  $B_{Rep}$  goes precipitously negative if  $Z_{cur} - Z_{com} + Z_{rec} > N_A$ .

We will later discuss experiments in which animals are pre-familiarized to contexts in which they will subsequently be conditioned. Intuitively, it would be expected that the more attributes of the context were previously learned, the sooner *on average* the current context could be identified with assurance. **Figure 4A** plots  $B_{Rep}$  values for the theoretically predicted average (referred to as “expected”)  $Z_{com}$  for a test in the conditioning context; we refer to  $B_{Rep}$  so calculated as *Expected*  $B_{Rep}$ . The greater  $Z_{rec}$  due to experience with the context in prior sessions, the faster *Expected*  $B_{Rep}$  rises as more attributes are sampled in the test session.

### Control of Representation Creation and Updating

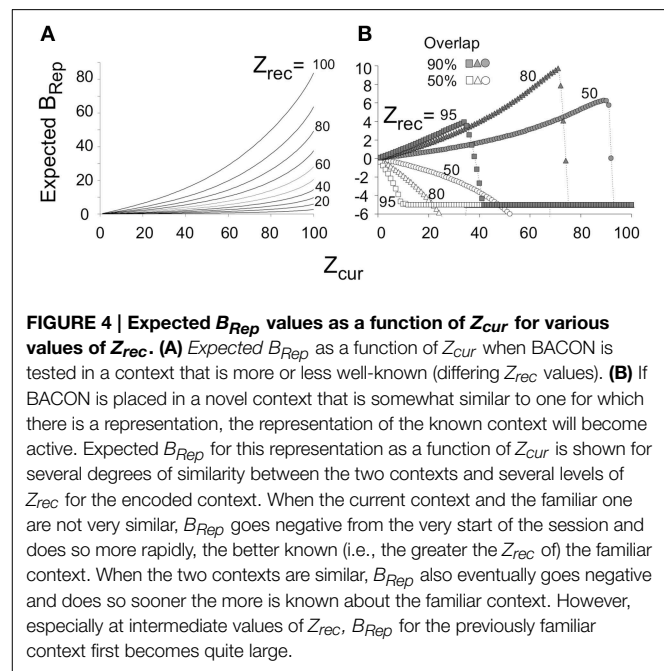
This section describes the logical rules used for controlling the state of the basic circuit; neural circuitry for implementing these rules is discussed separately.

We define certain levels of  $B_{Rep}$  ( $B_{old}$ ,  $B_{new}$ ,  $B_{add}$ , and  $B_{pv}$ ) that determine the mode of operation (**Figure 2**) of the basic circuit, sometimes along with other conditions.  $B_{old}$  is taken as the lowest level of  $B_{Rep}$  that provides reasonable but not foolproof evidence that the currently active hippocampal representation



**FIGURE 3 | Continued**

proportional to a power function of  $Z_{hi}(Z_{lo} - Z_{com})$ . Neuron III provides this inhibitory input; it receives excitatory input from both  $Z_{hi}$  and neuron V, whose output is proportional to  $Z_{lo} - Z_{com}$ . The output of neuron III is proportional to a power function of the product of its inputs. If an active hippocampal representation is valid, then  $Z_{lo} + Z_{rec} - Z_{com}$  must be less than or equal to  $N_A$ . If this condition is violated, the current context cannot be that currently represented by the hippocampus, and  $B_{Rep}$  will be infinitely negative, as seen in the exact calculations of **(A,C)**; in the circuit of **(B)** neuron I fires when  $Z_{lo} + Z_{hi} = Z_{com} > N_A$  and massively inhibits neuron II. **(C)** Exact values of  $B_{Rep}$  calculated from combinatorial formulas. Color coding as in **(A)**. **(D)** Firing rate of neuron II. Note that it emulates **(C)** quite well.



is in fact the one that was created for the current context (i.e., is “valid”). As explained in the next section, this is the level at which conditioning begins to become possible. If  $B_{Rep}$  exceeds this value, we speak of the current context as probably being “known” or “familiar.”  $B_{add}$  is taken as the level of  $B_{Rep}$  at which the automaton is virtually certain that the current representation is valid, and at which attributes are therefore allowed to be added to active representations (i.e., updating mode is entered). We note that it is crucial that a representation not be updated unless it is almost certain that it is valid, because adding invalid attributes to an existing representation would render it valid nowhere, and any conditioning that had occurred to it would be lost to effective usage. Therefore,  $B_{add}$  is made much higher than  $B_{old}$  (Table 2).  $B_{new}$  is the negative value of  $B_{Rep}$  below which it is deemed highly unlikely that the current representation is valid (i.e., BACON is probably in an unknown context for which it has no representation).

When  $B_{Rep}$  falls below  $B_{new}$ , creation mode (Figure 2A) may be entered and a new representation established, but two additional conditions must be satisfied:



- (1) New representations should be created only when enough of the current context's attributes have been sampled so that, once learned, they will suffice to allow recognition of the context on future occasions; we call this number  $Z_0$ . The minimum useful value for  $Z_0$  would be that for which  $B_{Rep}$  reached  $B_{old}$  once all  $N_A$  attributes of the context were sampled. With the parameters used here this would be  $Z_0 = 10$ . However, if  $Z_0$  were so low, recognition of the context on future occasions would require that every attribute of a context have been sampled before reasonably confident recognition could occur, which would presumably take a maladaptively long time to be completed. In fact, we will argue below that for our parameters,  $Z_0$  should be at least 45, which is the value that we adopt for the simulations of this paper.
- (2) If placed in one of two quite similar familiar contexts, vicissitudes of attribute sampling early in the session are likely to cause the representation of the incorrect context to be activated intermittently until enough information about the actual context has been acquired so that the correct representation becomes continuously active. Such invalid representations can sometimes cause  $B_{Rep}$  values sufficiently negative to trigger creation of a new representation, which would then be a second (and unwanted) representation for the same context. To reduce the likelihood of this happening, representation creation is suppressed if during the session, some representation has generated a  $B_{Rep}$  sufficiently positive that it is likely to be the actual representation of the context; the degree of positivity required is  $B_{pv}$  ("pv" for "probably valid"), and for the simulations of this paper this was set equal to  $B_{old}$ . This rule helps prevent unwanted representation creation; however, if the automaton is placed in a *new* context that is quite similar to a familiar one, the representation of the familiar context is likely to be activated and may generate a quite high  $B_{Rep}$  value (see **Figure 4B** and associated discussion). When this happens, the above rule, if unmodified, would prevent a representation of this new context from being created. However, if the context is novel, currently observed attributes and recalled attributes associated with existing representations will eventually disagree sufficiently to generate  $B_{Rep}$  values below  $B_{new}$ , and once this has occurred for all representations that formerly dictated suppression of creation, the suppression is lifted and creation of the new representation is allowed.

Given the above rules, creation of a new representation will usually occur when an active representation generates a  $B_{Rep}$  value below  $B_{new}$ . However, if there is no existing representation for some context at least somewhat similar to the current one, the number of CA3' cells initially activated (denoted  $X_{P_0}$ ) may remain below  $K_0$ ,  $EC'_{out}$  remain silent, and  $B_{Rep}$  remain zero, even after sampling considerable numbers of attributes. In such cases a new representation is created if the number of cells active in the pre-recurrent input activity pattern of CA3' ( $X_{P_0}$ ) is still less than  $K_0$  by the time  $Z_{cur}$  reaches  $Z_0$ .

Finally, if a new representation has just been created, it is by definition valid for the remainder of the session; therefore updating occurs at every subsequent sample of the session.

## Control of Context Fear Conditioning and Expression Conditionability

Context fear conditioning in BACON is due to LTP of CA3'-Amygdala' synapses, which occurs at synapses between active CA3' neurons and amygdala' cells when the amygdala' cells are depolarized by a US (Maren and Fanselow, 1995; Johansen et al., 2014). Obviously this cannot occur until a representation of the context has been created, but one would also not want it to occur unless there was reasonable evidence that the representation is valid.

We thus make conditionability increase linearly as a function of the degree to which  $B_{Rep}$  exceeds its value at  $B_{old}$ , reaching a maximum of unity when  $B_{Rep} = B_{add}$  (**Figure 5**—recall that  $B_{add}$  corresponds to virtual certainty):

$$Cnd = \text{linsig}(B_{Rep}|B_{old}, B_{add}), \quad (4)$$

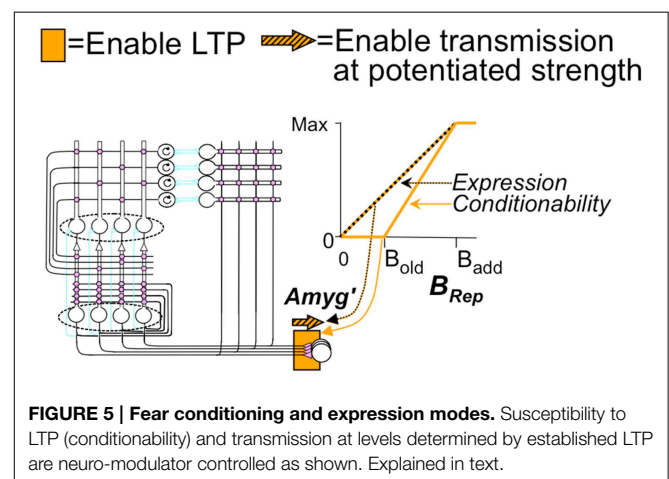
(see Methods for definition of  $\text{linsig}$ ). We then calculate the increment in strength due to a US as:

$$\Delta \text{Strength} = \alpha \cdot Cnd \quad (5)$$

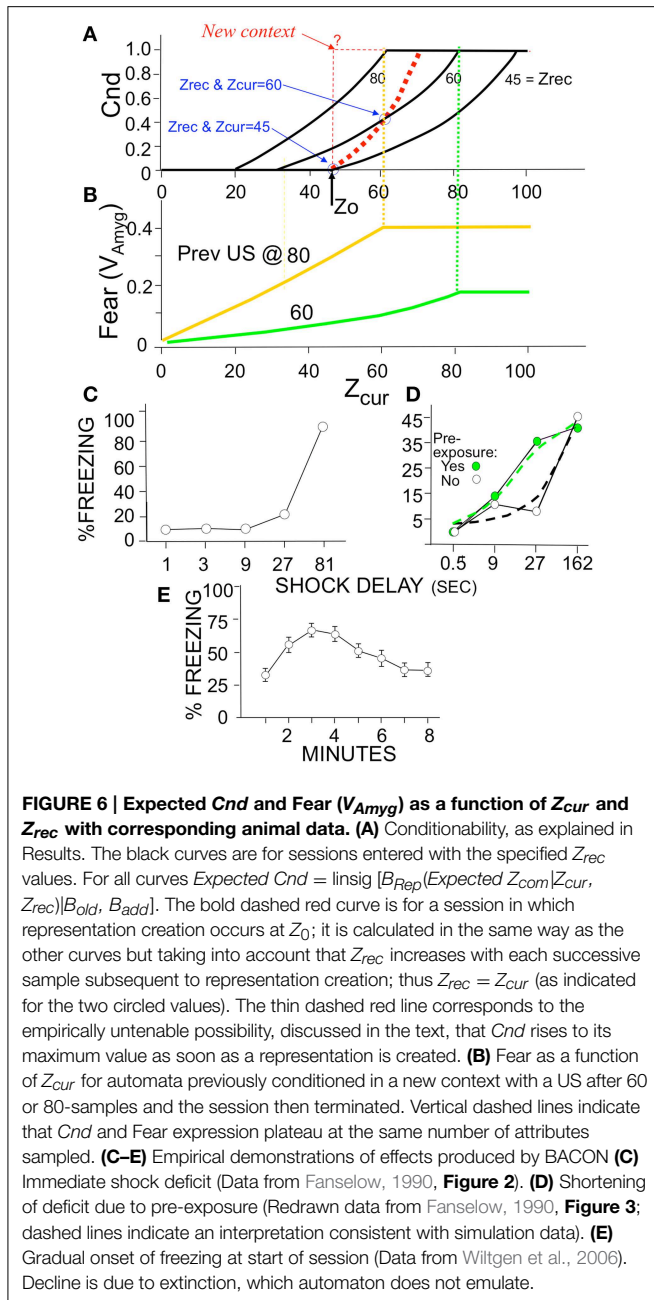
where  $\alpha$  is the amygdala learning rate parameter (**Table 2**). The black curves in **Figure 6A** show the growth of conditionability as sampling of contextual attributes progress throughout a session in a context for which a representation was previously created. The curves start rising from  $Cnd = 0$  when expected  $B_{Rep} = B_{old}$ .  $B_{Rep}$  increases both earlier and faster with increasing  $Z_{rec}$  because the more that is known about the context, the faster  $B_{Rep}$  increases as information about the current context is acquired.

This much is straight-forward and we believe plausible. However, special issues arise when USs occur *during* sessions where representations were created.

One of these concerns the value of  $Z_0$  (the minimum  $Z_{cur}$  at which a representation can be created). We argued above that given BACON's parameters,  $Z_0$  should be at least 10. However, if  $Z_0$  were in fact 10 and we introduced BACON into a new and distinctive context, a representation for the context would then be created at sample 10. Since, as pointed out above, a representation is by definition absolutely valid for the remainder



**FIGURE 5 | Fear conditioning and expression modes.** Susceptibility to LTP (conditionability) and transmission at levels determined by established LTP are neuro-modulator controlled as shown. Explained in text.



of a session in which it is created, conditionability should at least start to rise from zero at that point. However, suppose we were to remove BACON from the context right after the representation was created and then later re-introduce it. Since  $Z_{rec}$  would now be only 10, all  $N_A (=100)$  attributes would have to be sampled before  $B_{Rep}$  rose to  $B_{old}$  and conditioning became possible. That would leave us with the peculiarity that conditioning becomes possible after 10 samples in a fully novel context but only after 100 in one with which the automaton is already familiar. In order to prevent such an anomaly,  $Z_0$  would have to be such that  $Expected\ B_{Rep}$  on a session following an initial one that terminated just after  $Z_0$  samples, would reach  $B_{add}$  once this same number of samples

were made in the latter session (i.e.,  $Expected\ B_{Rep} = B_{add}$  when  $Z_{cur} = Z_{rec} = Z_0$ ). Reference to **Figure 6A** shows this to occur when  $Z_0 = 45$ , which value we have used for the simulations of this paper.

The other issue concerns the *rate* at which  $Cnd$  should increase if the automaton is left in the context after creation of the representation. Basing  $Cnd$  on  $B_{Rep}$  calculated from  $Z_{cur}$ ,  $Z_{rec}$ , and  $Z_{com}$  seems to make sense when BACON is in-essence making a statistical inference as to its whereabouts. However, if a representation for the current context has just been created, BACON can in-effect be *certain* that its active new representation is valid. Thus, it might plausibly be argued that once the representation is created, conditionability should jump to unity for the remainder of that session, as indicated by the dashed step function in **Figure 6A**. Alternatively, one might argue, somewhat as in the previous paragraph, that for a given amount of information available about the current context,  $Cnd$  should not go *down* in a later session with the same information. In that case, the upper limit for  $Cnd$  on the session of a representation's creation would be the value it would be expected to have given random sampling of attributes on a later session. The bold red dotted curve of **Figure 6A** plots this upper limit.

Whatever the cogency of either argument, it appears from available data that in a novel context conditionability rises *gradually from zero* following the period during which no conditioning is possible (Fanselow, 1990), exactly as shown in the bold dashed curve of **Figure 6A**. We therefore have constructed BACON so that it computes  $Cnd$  in this way on the session of encoding.

### Fear expression

Just as it would make little sense for fear to become conditionable before an animal has enough information about a context to know where it is, so too expression of previously conditioned fear should depend on an animal's confidence about its locale. We thus let the *effective* conductance ( $G_e'$ ) produced by a potentiated excitatory CA3'-amygdala' synapse be dependent on  $B_{Rep}$ , being zero when  $B_{Rep} = 0$  and becoming equal to its full potentiated value  $G_e$  when  $B_{Rep} = B_{add}$  (**Figure 5**) Specifically,

$$G_e' = \text{linsig}(B_{Rep}|0, B_{add}) \cdot G_e. \quad (6)$$

$G_e'$  is here taken to begin its increase at  $B_{Rep}=0$ , rather than at  $B_{old}$  (compare to equation 4) because we suppose that it is better to err on the side of being afraid. The actual expression of fear is proportional to the depolarization  $V_{Amyg}$  that is produced by the conductance  $G_e'$ :

$$V_{Amyg} = G_e' / (1 + G_e') \quad (7)$$

(see Methods). Thus, the onset of fear expression increases as  $B_{Rep}$  rises during a session (**Figure 6B**).

During a session where representation creation occurred,  $B_{Rep}$  in Equation (6) is replaced by  $Expected\ B_{Rep}$ . This is used in computing post-shock freezing.

## Neural Implementation of Control Rules and Bayesian Calculations

Although the above logical description of the BACON algorithm fully characterizes its operation (and provides a sufficient basis for reading later sections), it is important to make clear that the operations specified above could in fact be carried out by neurons. There would be many ways of doing this; we sketch one here.

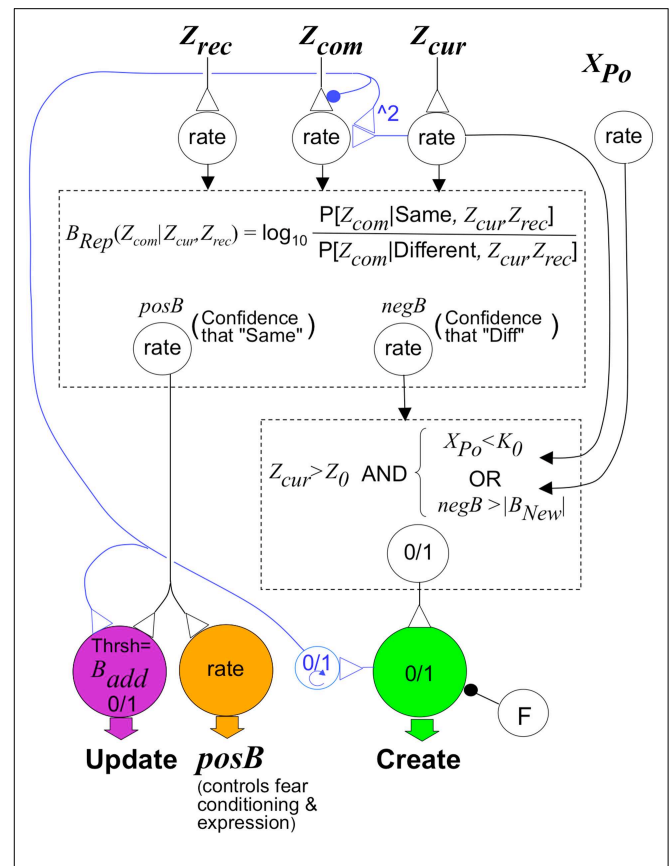
Central to this model is the control of plasticity and other neural properties. Since such control applies to entire classes of neurons and synapses, we think of it as being exercised via widely distributing neuromodulators. We assume that Recall mode is the default and imagine separate modulators for casting the network into Create and Update modes. We also employ a third modulator to signal the extent of positivity of  $B_{Rep}$  ( $posB$ ) for controlling  $Cnd$  and fear expression. Control of plasticity is assumed to be via direct effects on synaptic properties, though one could imagine the possibility of indirect control via modulation of firing rates. We note that the work of Hasselmo (2006) provides a precedent for supposing hippocampus to be thrown into different, neuromodulator-controlled modes of operation during encoding and recall and that a variety of neuromodulators are known to affect LTP (e.g., Frey et al., 1989; Markram and Segal, 1990; Huerta and Lisman, 1995; Thomas et al., 1996. Reviews: Kenney and Gould, 2008; Lovinger, 2010; Lisman et al., 2011; Hawkins, 2013).

Figures 3B<sub>1</sub>, B<sub>2</sub> sketch circuitry that can compute approximate Bayesian weights of evidence; as seen in Figure 3C, it does this fairly accurately. Figure 7 sketches circuitry that could implement the control rules described in previous sections. We think of prefrontal cortex as a plausible locus for this speculative control circuitry.

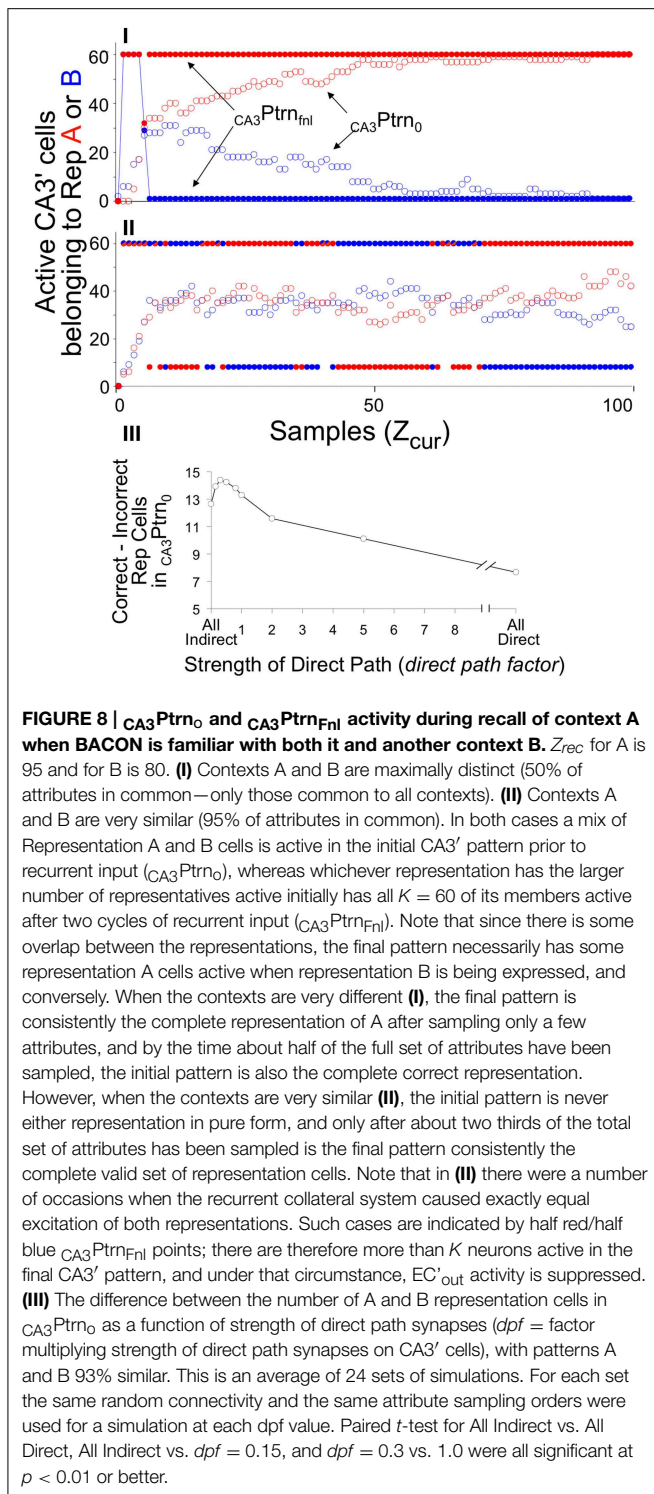
## Properties and Performance of BACON

Conditionability and fear expression will vary over time, dependent on what attributes are randomly sampled. In order to avoid proposing a theory of sampling over time, which is a complex topic of its own, we use sample number ( $Z_{cur}$ ), rather than time, as an independent variable. However, a few comments about time *per se* are needed: The earliest that a representation can be created and conditioning occur is at sample  $Z_0$ , and the earliest time that conditioning begins in rats is about 30 s (Fanselow, 1990). We therefore imagine that at the start of a session about  $Z_0$  (=45) attributes are sampled per 30 s. However, we presume that any plausible theory of sampling would have sampling slow greatly from this rate as the proportion of total attributes sampled approached the total available. Indeed, we imagine that it would be rare for full knowledge of a context's attributes to ever be achieved.

It should be noted that BACON has intentionally not been designed to extinguish. Since extinction of context fear is time-dependent, extinction could not have been included without a model of sampling, which we wished to avoid. Moreover, the behavioral consequences of the model's Bayesian inference mechanisms are seen far more clearly in the absence of extinction.



**FIGURE 7 | Sketch of circuit that would implement BACON control rules.** The upper dashed box represents a circuit that, like that in Figure 3B, can compute  $B_{Rep}$  in response to input from neurons whose rate of firing signals  $Z_{cur}$ ,  $Z_{rec}$ , and  $Z_{com}$ . It has as output, neurons that separately convey positive and negative values of  $B_{Rep}$ . These, along with activity of a neuron whose activity level  $X_{Po}$  is proportional to the number of neurons active in the pre-recurrent input activity pattern of CA3' ( $CA3'P_{trn}$ ), determine the activity of the neuromodulator-releasing neurons shown at the bottom of the figure. The neuromodulator that configures the cortical-hippocampal circuit for representation creation is released when the logical conditions in the lower dashed box are met and the creation neuron is not inhibited by neuron F. We have written the logical conditions themselves rather than a neural circuit that would implement them because this makes the logic clearer, and there are many straight-forward ways to implement these rules neurally. Inhibitory input to the Create neuron from neuron F suppresses its activity if a representation (different from the current one) for which  $B_{Rep} > B_{pv}$  was active earlier in the session (as explained in the section on creation and updating). Circuitry for controlling F is provided in the Supplementary Material (Topic D). The neuromodulator that configures the cortical-hippocampal circuit for updating is released when  $posB > B_{add}$ . However, if a new representation has been created, the neuromodulator must continue to be released until the automaton leaves the context. This is mediated by a binary ("0/1") working memory neuron that becomes active when representation creation occurs and remains so until BACON leaves the context. The binary neuron also reconfigures input to a neuron innervated by  $Z_{com}$  so that (as discussed in the text section on Conditionability) it will reflect *Expected*  $Z_{com}$  instead of  $Z_{com}$  itself. This is done by inhibiting transmitter release from the  $Z_{com}$  neuron and activating an input from  $Z_{cur}$  that causes the target cell to respond as the square of  $Z_{cur}$ ; this in-effect computes the expected value of  $Z_{com}$  given that  $Z_{rec}$  equals  $Z_{cur}$  when the automaton is updating representations. Control of fear conditioning and expression is mediated by a neuromodulator that is released in proportion to the  $posB$  signal. Its effect on conditioning and fear expression are as described in the text.



The approach proposed for adding extinction is sketched in the Discussion.

### Illustrative Simulations

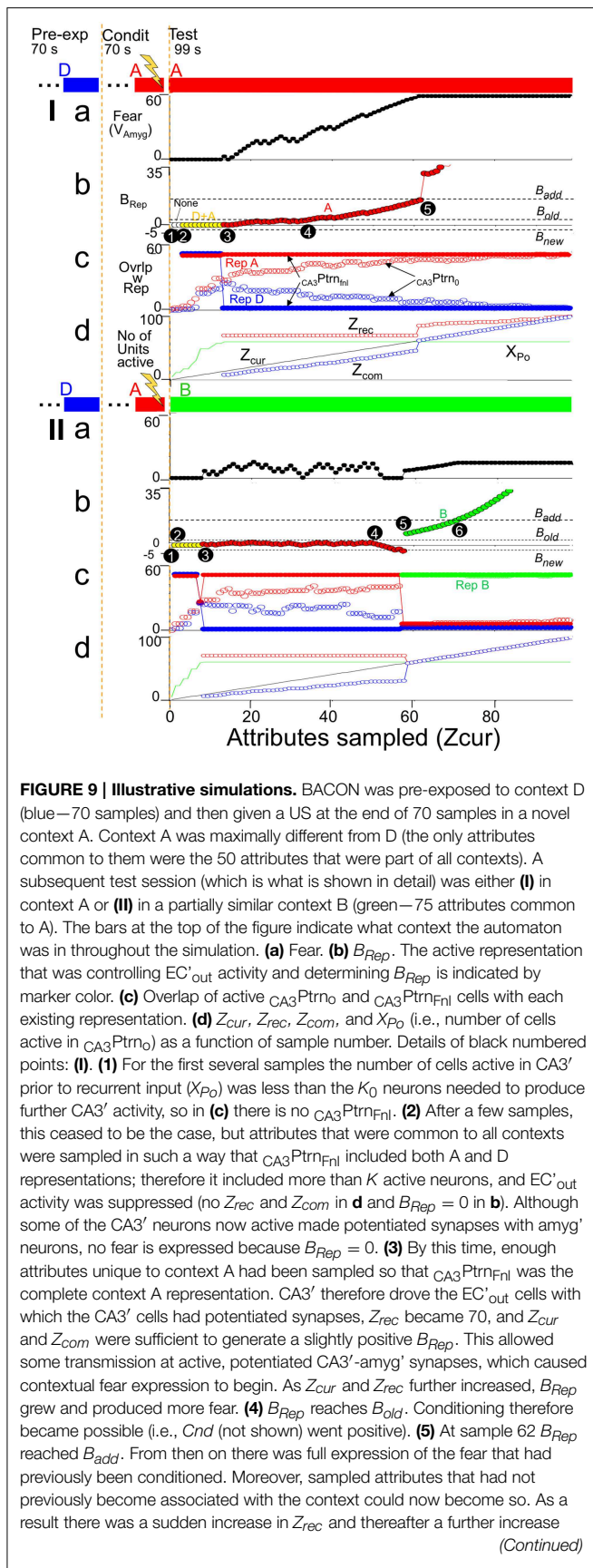
**Figure 8** illustrates the operation of BACON's hippocampus in recall mode. For this simulation the automaton was first

given sufficient exposure to two contexts (A and B) to cause creation of representations for each. **Figures 8I,II** show activity during recall tests in context A. In **Figure 8I** context A and B were fairly distinct, whereas in **Figure 8II** they were very similar. It should be borne in mind that the automaton's knowledge of those contexts for which it has representations will usually, as here, be incomplete unless it has spent a great deal of time in those contexts. Moreover, because of the vicissitudes of random attribute sampling, attributes common to two different contexts will often have become associated with one representation but not the other. Thus, if tested in either context, where it will again sample attributes at random, it may well sample common attributes that by chance became associated with the other context, and early in a session these might by chance predominate. As a result,  $EC'_{in}$  activity will often activate some elements of incorrect representations in  $CA_3'$ . The graphs show the number of active  $CA_3'$  units that were part of the representations of each context as sampling progressed during the test. This is shown both prior to input from the recurrent collateral system ( $CA_3Ptrn_0$ ) and after the final iteration of such input ( $CA_3Ptrn_{Fnl}$ ). In both simulations there was a mix of elements of each representation active prior to recurrent collateral input, but in the post-recurrent input representations, whichever pattern had the most active elements prior to recurrent input prevailed. Moreover, as sampling progressed, the valid representation increasingly predominated. When A and B were quite different, the final output of  $CA_3'$  was the valid representation starting fairly early in the session; when they were quite similar, it took many samples before the final representation was consistently the valid one.

This sort of simulation provides a good opportunity to evaluate the consequences of using the direct vs. indirect pathways during recall. As will be explained in Discussion, it was expected that use of the indirect path alone during recall should lead to more reliable activation of the actual current context's representation than use of only the direct one. However, both operating together should produce somewhat better performance than either operating alone. The ratio of strength of direct path to indirect path synapses in recall mode is the parameter  $dpf$ . In simulations I and II only the indirect pathway was operational ( $dpf = 0$ ). **Figure 8III** shows the preponderance of Representation A over B neurons in  $CA_3'$  prior to recurrent collateral input, as a function of  $dpf$  in simulations like those in I and II. As expected, the indirect pathway produces better separation than the indirect one, but both together are best. However, the advantage of incorporating input from the direct pathway is modest and is seen only if input from the direct pathway is relatively weak. For simplicity's sake we chose to use only the indirect pathway (i.e.,  $dpf = 0$ ) in the simulations below.

Having considered the activity of the hippocampus alone, **Figure 9** brings in the rest of the circuit. It shows two simulations. In both, BACON was fear-conditioned in a context A after pre-exposure to a very different context D. It was then tested either in A (**Figure 9I**) or in a different but somewhat similar new context B (**Figure 9II**). The colored bars at the top of **Figures 9I,II** show what contexts the automaton was in and when shock was given.



**FIGURE 9 | Continued**

whenever another not already known attribute was sampled. As  $Z_{cur}$ ,  $Z_{rec}$ , and  $Z_{com}$  increased, so too did  $B_{Rep}$ . **(II)** **(1)** Initially  $X_{Po} < K_0$ , so there was no  $CA3_{Ptrn_{Fnl}}$  activity. **(2)** Both Representation A and B cells were active in  $CA3_{Ptrn_{Fnl}}$ ; thus  $Z_{rec}$ ,  $Z_{com}$ , and  $B_{Rep}$  are zero. **(3)** As more attributes were sampled, the representation of context A, the conditioning context, which was more similar to the current context than to context D became expressed in  $CA3_{Ptrn_{Fnl}}$ . During some of this time, there was a small positive weight of evidence ( $B_{Rep}$ ) for the validity of this representation and a little fear was expressed. **(4)** As the sampling continued, attributes that were not among those associated with the currently active conditioning context were sampled and provided evidence that the current context was not B; this caused a negative  $B_{Rep}$  value to emerge. When  $B_{Rep}$  was negative, no fear was expressed even though the hippocampal representation of the conditioning context was active. **(5)** At sample 57  $B_{Rep}$  fell below  $B_{new}$ , and since  $Z_{cur}$  was greater than  $Z_0 (=45)$ , a representation of the current context (context B) was created. Because the representations of A and B had a few cells in common, there is some fear expressed, which grew as the  $B_{Rep}$  value for context B's representation increased. **(6)**  $B_{Rep}$  reached  $B_{add}$ . Increases in  $B_{Rep}$  beyond  $B_{add}$  do not cause any increased expression of whatever fear has been conditioned (Equation 6), so fear levels off. Fear leveled off when  $B_{Rep}$  reached  $B_{add}$ , because could not increase further once because  $B_{Rep}$  reached  $B_{add}$ .

Note first that in **Figure 9Ic**, as in **Figure 8I**, when BACON was tested in a familiar context after having been exposed to it and another fairly dissimilar one,  $CA3'$  activity separates well even prior to recurrent collateral input, and  $CA3_{Ptrn_{Fnl}}$  becomes the complete representation of context A early in the session. On the other hand, when placed in a different unfamiliar context B that was moderately similar to A, the hippocampus initially activated the representation of A, but eventually (after having acquired enough information about the current context to show that A is invalid) created a new representation for B.

**Figure 9Ib** plots  $B_{Rep}$  for the active representation throughout the simulation, with which representation was active indicated by point color. When in context A, the representation of A became the sole representation active at sample 15 (black circle point 3). This allowed  $EC'_{out}$  neurons to be activated ( $Z_{rec} \sim 70$ , **Figure 9Id**). Since most of the currently sampled attributes overlapped those then active in  $EC'_{out}$  ( $Z_{com}$  slightly less than  $Z_{cur}$  in **Figure 9Id**),  $B_{Rep}$  took on a positive value, and since the context A  $CA3'$  representation cells now active were those to which fear was conditioned, fear expression began to rise (**Figure 9Ia**). As more attributes of A were sampled,  $Z_{cur}$  and  $Z_{com}$  rose (**Figure 9Id**), and so  $B_{Rep}$  and fear increased correspondingly (**Figure 9Ia**). At black circle point 5,  $B_{Rep}$  reached  $B_{add}$ . This allows maximal expression of conditioned fear and allows updating of the representation of A so all sampled attributes that were not already associated with the representation of A became so causing sudden increments in  $Z_{rec}$  and  $Z_{com}$  (**Figure 9Id**).

When placed in context B (unfamiliar, but moderately similar to A), the representation of A was activated early on (**Figure 9IIc**). The attributes of A and B were sufficiently similar so that a positive  $B_{Rep}$  value developed at about sample 8 (**Figure 9IIb**), with a corresponding modest expression of conditioned fear (**Figure 9IIa**), but by sample 60 (black circled point 5) differences between current attributes and the recalled

attributes of representation A became so great (**Figure 9IId**) that  $B_{Rep}$  fell below  $B_{new}$  (point 5 in **Figure 9Ib**), and a representation for context B was created. Once that happened fear expression occurred only because the representation of context B had some neurons in common with that of context A.

### Representation Characteristics

Representations in BACON, even of quite similar contexts, have only limited overlap (i.e., the encoding (creation) process produces considerable “pattern separation”). Two factors contribute to the slightness of overlap. First, as often discussed, the large number of DG’ neurons, their innervation by random groups of EC<sub>in</sub> cells, and the firing of only the  $K$  most excited, promote pattern separation; these factors cause contexts with 90% overlap at the level of EC<sub>in</sub> to have about 30% overlap at DG’ during representation creation. Second, contexts are generally created when only a fraction of their total attributes have been (randomly) sampled. With this factor operating along with the first, overlap of the representations for 90% similar contexts falls to about 10%. As a result, even when two contexts are extremely similar, relatively distinct sets of attributes will control the form of their representations and very different representations result.

### Conditionability—The Immediate Shock Deficit and Pre-familiarization

When placed in a novel context that is fairly distinct from any familiar one, a representation of the context will be created at sample  $Z_0$ . Conditionability begins to increase at this point and continues to do so as  $Z_{cur}$  increases, until it reaches its maximum value of unity (**Figure 6B**, bold red dashed curve). This delay, with subsequent gradual increase in conditionability, corresponds to the immediate shock deficit of real context fear conditioning (**Figure 6C**).

If placed in a *familiar* context (i.e., one for which a representation was previously established) conditionability begins to rise when  $B_{Rep}$  reaches  $B_{old}$ . The more that is known about the context’s attributes (i.e., the greater  $Z_{rec}$ ) at the start of the session, the sooner  $B_{old}$  will be reached (**Figures 4A, 6A**). **Figure 6D** illustrates such an earlier rise in conditionability following pre-familiarization with the to-be-conditioned context in a real experiment on the immediate shock deficit. In the theoretical case the rate of rise is generally greatest in the session of creation, because expected conditionability is proportional to *Expected*  $B_{Rep}$  which is an increasing function of both  $Z_{cur}$  and  $Z_{rec}$ , and whereas  $Z_{rec}$  increases at each successive sample during the session of encoding, on later sessions it is constant throughout a session unless  $B_{Rep}$  comes to exceed  $B_{add}$ .

It is a basic feature of BACON that the attributes of a novel context cannot begin to be learned until a hippocampal representation is formed, and this cannot occur until the number of samples made exceeds  $Z_0$ . Therefore, even many repeated exposures to the context, each of which is too short for  $Z_0$  samples to be made, will not reduce the immediate shock deficit. Moreover, attributes cannot be added to an existing representation unless  $B_{Rep}$  exceeds  $B_{add}$ . Therefore, once BACON has been familiarized

sufficiently with a novel context to allow some measure of conditioning, further later exposures will only be effective in shortening the immediate shock deficit if some of the exposures are long enough to allow  $B_{Rep}$  to exceed  $B_{add}$ . To our knowledge no experiments that would evaluate either of these features have been published.

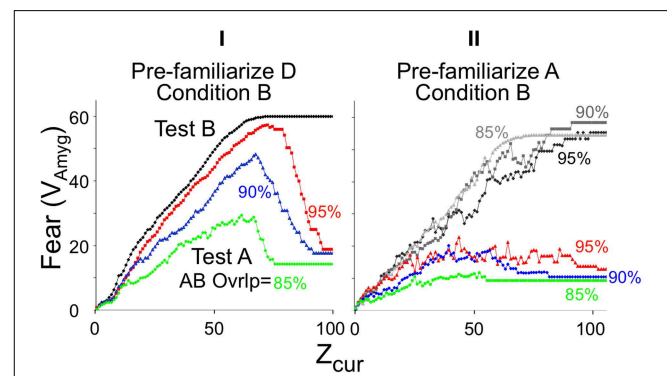
### Fear Expression—Delayed Onset of Fear and Effects of Context Familiarity

Fear expression in BACON develops gradually when the automaton is introduced into to a feared context (**Figure 6B**). As illustrated in **Figure 6E**, fear-induced freezing behavior does in fact develop gradually, though because extinction sets in, the rise is followed by a decline that is not emulated by BACON because we chose not to incorporate extinction mechanisms.

Since both conditionability and fear expression become maximal when  $B_{Rep}$  reaches  $B_{add}$ , both become maximal at about the same time (**Figures 6A,B**—note dashed lines between frames **Figures 6A,B**).

### Generalization

In BACON, fear conditioned in one context may generalize to another similar context for either of two reasons: (1) The hippocampal representations of the two contexts may have neurons in common. (2) The vicissitudes of attribute sampling may sometimes cause periods of activation of the hippocampal representation of a familiar context *similar* to the actual context rather than of the actual context itself. The latter is almost certain to occur if the context does not yet have a representation of its own. We chose the automaton’s parameters so that for contexts with 90% attribute overlap, hippocampal representations would only overlap by about 10%. Therefore, generalization in BACON



**FIGURE 10 | Generalization of context fear.** BACON was conditioned by a shock at the end of a session in novel context B and then tested either in the conditioning context (black and gray curves) or in a generalization context A (colored curves). Context A’s similarity to B was either 85% (green), 90% (blue), or 95% (red). The automaton was pre-exposed to either the generalization context (A) (**II**) or to a very different context (D) as a control for per-exposure *per se* (**I**). Because degree of conditioning is somewhat affected by pre-exposure to a context similar to the conditioning context, separate conditioning-context tests had to be run for each pre-exposed generalization context (shades of gray in **II**). Each curve is the average of 30 simulations. The pre-familiarization session was 95 samples, the conditioning session 80, and the test session 99.

is usually attributable more to misidentification of context than to representation overlap.

Degree of context fear generalization is very dependent on the extent to which the automaton is already familiar with the context in which generalization is tested. **Figure 10I** illustrates generalization when BACON is not familiar with any contexts that are at all similar to that in which it has been conditioned. During generalization tests in such cases there is initially substantial generalization of fear because the automaton at first misidentifies a similar but unfamiliar test context as the conditioning context itself. But once it is recognized that the test context is a new one, a representation for it is created, and fear then falls off; the only fear that then remains is that caused by those representation neurons that are common to the test and conditioned context. However, if BACON was pre-familiarized with the test context (**Figure 10II**), only a slight initial surge of fear occurs, because the pre-established representation of the test context itself quite rapidly becomes active.

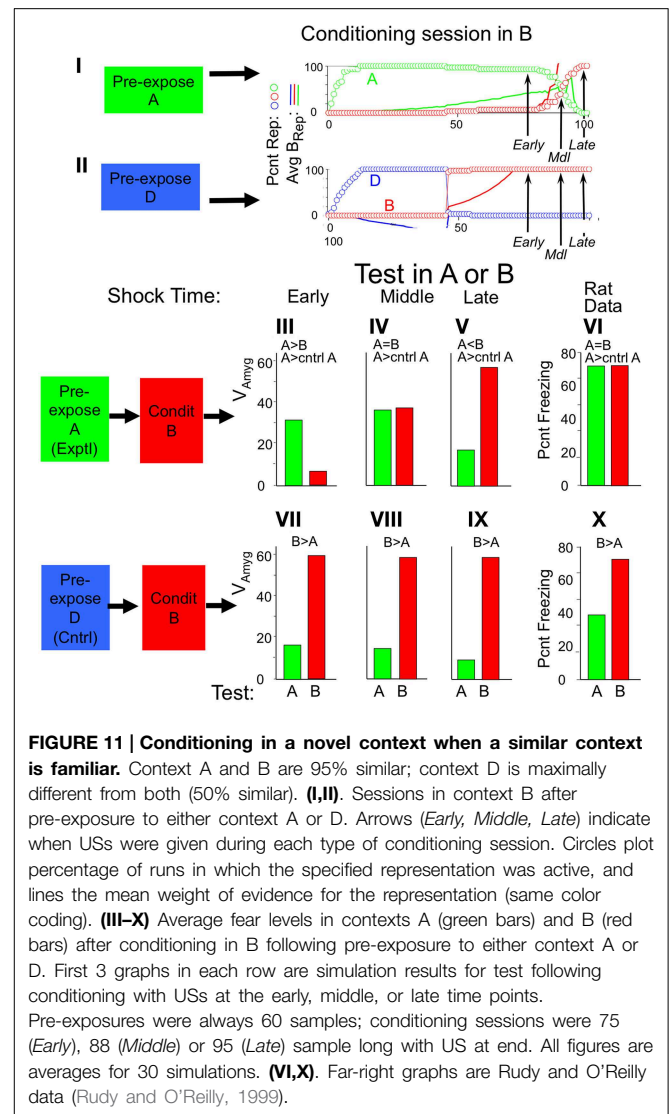
Pre-exposure to the conditioning, as opposed to the test context also diminishes generalization (not illustrated). This happens because the better-known the conditioning context (the greater its  $Z_{rec}$ ), the more rapidly a similar context will be recognized as different. Decreased generalization of stimuli that are better known has been observed both for context fear and other forms of conditioning (Kiernan and Westbrook, 1993; Rudy and O'Reilly, 1999).

### Effects on Conditioning of Familiarity with Similar Contexts

Rudy and O'Reilly (1999; see also Rudy et al., 2002) have published intriguing experiments in which, when rats were shocked in an unfamiliar context similar to one they knew, fear became conditioned to the familiar but absent context. BACON does something similar. The reason can be seen in **Figure 11I**, which shows what representations are active when BACON is placed in an unfamiliar context B similar to a context A to which it was pre-exposed. Initially the existing representation of A is activated because B has many of the same attributes as A. As sampling progresses, the automaton becomes more and more "certain" that it really is in A, as reflected by the gradual increase of A's  $B_{Rep}$  value (and as discussed above with respect to **Figure 4B**). But eventually sufficient discrepancies between currently observed attributes and those associated with A's representation mount to a point where a representation for B itself is created (and its  $B_{Rep}$  value then increases rapidly). If a US were given at the point marked "Early" fear would be conditioned to context A, if given at the "Middle" point, each representation would be active in about half the runs, and if given "Late," conditioning would be to context B.

Had BACON been pre-familiarized with a context, D, very different from B, conditioning at the corresponding points in the experiment would always have been to the actual context, B (**Figure 11II**).

Fear expressed in each context after giving BACON USs at one of these three points are shown in **Figures 11III–V, VII–IX**. When BACON had not been pre-exposed to A, it always showed



**FIGURE 11 | Conditioning in a novel context when a similar context is familiar.** Context A and B are 95% similar; context D is maximally different from both (50% similar). **(I,II)** Sessions in context B after pre-exposure to either context A or D. Arrows (*Early, Middle, Late*) indicate when USs were given during each type of conditioning session. Circles plot percentage of runs in which the specified representation was active, and lines plot the mean weight of evidence for the representation (same color coding). **(III–X)** Average fear levels in contexts A (green bars) and B (red bars) after conditioning in B following pre-exposure to either context A or D. First 3 graphs in each row are simulation results for test following conditioning with USs at the early, middle, or late time points. Pre-exposures were always 60 samples; conditioning sessions were 75 (*Early*), 88 (*Middle*) or 95 (*Late*) sample long with US at end. All figures are averages for 30 simulations. **(VI,X)** Far-right graphs are Rudy and O'Reilly data (Rudy and O'Reilly, 1999).

considerable fear of B but little of A, though some generalization did occur. However, when it had been pre-exposed to A, the outcome depended on when the US was given (**Figures 11III–V**). When given Early, fear of A was greater than B (**Figure 11III**) (and also greater than fear of A in the absence of pre-exposure to it). When given Late, fear during testing was greater for B (the actual conditioning context) than for A (**Figure 11V**). When the US had been at the Middle time point, fear was expressed mostly to A in some runs and to B in others (not shown), with the average across simulations being about equal for both (**Figure 11IV**). In the actual experiments (**Figures 11VI,X**) the pattern of results is similar to those of the Middle time point simulations. In the experiments of panels **Figures 11VI,X** two shocks were given, one at 120 and another at 240 s. In that case one could imagine the first being at our "Early" and the second at our "Late" time point. However, in other experiments, which had similar results, a single shock was given, as in our simulations.



## Discussion

### Multiple Modes of Hippocampal Function—Necessity and Benefits

The algorithm described here was an attempt to construct a neurally plausible mechanism for context fear conditioning. The role of the hippocampus in this is the creation and subsequent recall of permanent, relatively non-overlapping representations of contexts to which associations can be made. We assume that at each entry to a context its attributes are discovered serially in random order. Since attributes of a context will be sampled in different orders on different occasions, and since information about a context will usually be incomplete when its representation is created, determining whether a context is novel (in which case a representation for it should be created) or familiar (in which case pattern completion mechanisms should be allowed to operate) is a challenging problem. This led us to construct for BACON extra-hippocampal machinery designed to make optimal decisions as to the novelty or familiarity of whatever context it was in, and then to allow this machinery to configure the hippocampus specifically for representation creation, updating, or recall, as required.

This strategy solves two problems that arise if the hippocampus is assumed to have only one state:

- (1) In conventional theories of the hippocampus the occurrence of a novel stimulus configuration that is similar to one previously encoded may activate the complete code of the familiar stimulus with the result that a new representation will not be created but its attributes will, improperly, become associated with the existing representation. Or conversely, in the presence of a known stimulus, incomplete sampling of its attributes or sampling of attributes not originally encoded will cause a new representation for it to be created so that one stimulus will now have two different (pattern-separated) representations. This is the so-called “pattern separation-completion tradeoff” that has been extensively discussed in the literature (O’Reilly and McClelland, 1994; and see O’Reilly and Rudy, 2001). The sort of control of hippocampal mode that we have utilized in BACON essentially abolishes this problem, because decisions as to novelty vs. familiarity are made optimally, and hippocampal plasticity is shut off when it would be inappropriate.
- (2) Although, as will be explained below, pattern recall should work best if information about a stimulus reaches CA3’ via the indirect (EC’<sub>in</sub>-DG’-CA3’) pathway, it has been argued, and is often accepted, that properties of the indirect pathway that are needed for optimal representation encoding make it impossible for the pathway to also be used effectively during recall (Treves and Rolls, 1992; O’Reilly and McClelland, 1994; Treves et al., 2008). However, this sort of argument loses all force if, as is done in BACON, the properties of the hippocampus can be configured for its different modes of operation.

### DG-CA3 Projection

In BACON each DG’ neuron has a single dedicated CA3’ partner with which it makes a non-plastic synapse. In the

actual hippocampus there are less CA3 than dentate cells, and each dentate cell innervates a small number of CA3 cells, via synapses that show some sort of plasticity, but not the type of associative NMDA-dependent plasticity seen at EC-dentate and recurrent CA3 network synapses (Bortolotto et al., 2003; Reid et al., 2004). In so far as a KWTA competition within DG’ is used to choose the cells for a representation, it seemed to us to make good design sense to let those cells have specific partners in CA3’. Thus, representational elements are in-effect DG’-CA3’ dyads that, during recall, are excited by potentiated inputs from EC’<sub>in</sub> and that are bound to other dyads of the same contextual representation by the potentiated synapses of the recurrent collateral network.

CA3’ cells could have been paired with DG’ cells by making the synapse between a DG’ cell and one of a small number of CA3’ cells it innervates functional only if and when the DG’ neuron won a KWTA competition during representation creation. Something like this seems to happen during real nervous system development where activity-dependent mechanisms cause vertebrate muscle fibers, cerebellar Purkinje cell, and thalamic neurons to become innervated by single motor neurons, climbing fibers, or topographically close thalamic afferents (e.g., Favero et al., 2009; Kano and Hashimoto, 2009). However, for us it was simplest to just take the existence of DG’-CA3’ pairs as given. Nevertheless, it should be noted that a given DG’-CA3’ pair will never get used if the sets of attributes whose EC’ representations innervate the DG’ cell never occur together in the “life” of the automaton. Therefore, pairing up a CA3’ cell with an active DG’ cell at the time of representation creation would have allowed a smaller CA3’ layer and would have reduced the size of the huge recurrent collateral network. One can imagine that this could be happening in the biological hippocampus and could account for the smaller number of CA3 than dentate neurons, the limited fan-out of dentate neurons, and the presence of atypical plasticity at dentate-CA3 synapses.

### Direct/Indirect Pathway Balance in Recall

Since, as discussed above, we can ignore computational arguments against a role for the indirect pathway during recall, we consider the relative merits of using it vs. the direct one: Each DG’ and CA3’ cell will be innervated by  $F = 60$  of the  $N_{Ctx} = 1000$  EC’ attribute cells. The innervation of each DG’ and CA3’ cell is chosen *randomly and independently*; thus the two members of a DG’-CA3’ pair do not receive input from the same set of EC’ attribute cells. When the representation for a context is created, there will usually be on the order of  $Z_0 = 45$  EC’<sub>in</sub> attribute cells active, and on average each DG’ and each CA3’ cell will be innervated by (a different) 2–3 of them. The synapses of the active attribute cells on the  $K$  DG’ and CA3’ cells of the forming representation will become potentiated. In CA3’ it will be the approximately 2–3 active EC’<sub>in</sub>-CA3’ synapses that innervate each cell that become potentiated. However, in DG’ the cells that are active will be *selected* as the  $K$  that are most excited and therefore most heavily innervated by the active EC’ attribute cells. The DG’ cells that are active will therefore have a richer active EC’ attribute cell innervation than will the population as a whole, and it will be some 7–8 synapses on them that will become potentiated.



During recall, DG' and CA3' cells receive information from EC<sub>in</sub> only via potentiated synapses. In order for a representation cell for a context *A* to be selectively activated by context *A*, as opposed to a different but similar context *B*, a representation cell for *A* must receive effective input from attribute-representing neurons that are specific to *A*, and if *A* and *B* are very similar, there will not be many *A*-specific attribute neurons. So when very similar contexts must be differentiated, the DG' cells, which get effective input from 7 to 8 EC<sub>in</sub> cells, are more likely to get input from a given attribute (which will be passed to a cells CA3' partner) than will the CA3' cells, which get effective input from only 2 to 3 EC<sub>in</sub> cells. Thus, if only the direct or only the indirect pathway were used for recall, the indirect would be better because it would be more likely to provide effective input from a context *A*-specific neuron. However, if both the direct *and* the indirect pathway provide information that determines the response of CA3' cells prior to recurrent input, differentiation should be somewhat better than with either pathway alone. As shown in **Figure 8III**, the effective separation of the representations of two similar contexts at CA3' prior to recurrent input (i.e., CA3Ptrn<sub>0</sub>) as a function of the relative strength of the direct pathway is in fact just as expected. Moreover, it seems to be optimal for the direct pathway to be weak relative to the indirect one. This may be seen as consistent with the biological situation, where both pathways are present, but the direct one makes relatively distal synapses on CA3 neurons.

### From CA3' to EC<sub>out</sub>

As explained at the start of Results, we have allowed CA3' to innervate EC<sub>out</sub> directly, rather than via a CA1-like processing stage because complete innervation in the recurrent collateral network and in EC<sub>out</sub> made a CA1' redundant. Since it is the Shaffer collateral synapses of CA3 cells on CA1 neurons that have told us much of what we know about Hebbian LTP, it may seem ironic that we did not explicitly include a CA1 in the algorithm. However, in the current model, the Hebbian nature of *EC synapses on dentate and CA3 cells* are sufficient to produce effective encoding and recall of contextual representations. In a model of this type CA1 plasticity would serve mainly to help mitigate any errors or omissions that remained in representations after their processing by the recurrent collateral network.

### Functioning of DG' and CA3' during Representation Creation and Recall

There have been a number of experiments in which activity in dentate or CA3 in novel or familiar contexts has been assessed or manipulated. We consider here how BACON would behave in some of these.

#### Activity in Novel Contexts

BACON will start its time in a novel context in Recall mode, but fairly soon  $B_{Rep}$  will fall below  $B_{new}$  and a new representation will be created. For the remainder of the session, as each additional attribute gets sampled, excitation will be passed to CA3' over the direct and indirect pathways, CA3Ptrn<sub>Fnl</sub> determined, and since the system is in Update mode, CA3Ptrn<sub>Fnl</sub> will be reflected back to DG' so that the synapses between the added EC<sub>in</sub>

attribute cells and DG' cells can potentiate. From then on in each computational cycle hippocampal activity will consist of the *K* DG'-CA3' pairs of cells of the current representation fully active, with the rest of the cells of each layer silent. Moreover, these active cells will overlap relatively little with those of other representations. Consistent with this, it does appear that, whereas there is a lot of overlap between the DG' cells active in different *familiar* contexts, during exposure to *novel* contexts DG' activity is quite context-specific, even when the contexts are similar (Deng et al., 2013) (**Table 3, A**).

#### Activity in Familiar Contexts

In a familiar context BACON will operate in Recall mode until such time as it "decides" it almost certainly "knows" where it is (i.e.,  $B_{Rep} > B_{add}$ ) at which time Update mode will begin to operate; we will assume for discussion that it is mainly Recall mode that will be observed during most experiments. During recall, it is DG's job to pass to CA3' information as to how much excitation each DG' cell has received from EC<sub>in</sub>; therefore, during recall the KWTA rule does not apply in DG'; and cells there fire in proportion to their excitation. Thus, in DG' any cell that represents an encoded context having attributes in common with the current one will fire to some degree. However, in CA3' the KWTA rule is operative. CA3Ptrn<sub>0</sub> will typically include cells (possibly many but usually a minority) that do not represent the current context, because some representation cells of other previously encoded contexts may happen to have dense innervation from common attributes that happen to be in the current attribute sample. However, CA3Ptrn<sub>Fnl</sub>, once it emerges, will (if enough attributes have been sampled) be the complete representation of the current context and will overlap little with the representations of other, even similar, contexts. We assume for the sake of discussion that it would be mostly this final pattern of activity that would be detected in experiments. In fact, recordings from real animals do seem to show that in familiar contexts DG' uses a rate code with cells active in different contexts overlapping considerably, whereas in CA3' there is little overlap between those active in different but similar contexts. Leutgeb et al. (2007), Neunuebel and Knierim (2014), **Table 3, B**; see Supplementary Material, Topic F for further discussion and a simulation of the Leutgeb experiment.

The proportion of sampled cells firing during tests in one of two familiar contexts in rat dentate has been reported to be about a third that in CA3 (Leutgeb et al., 2007). However, since rat dentate has about 8 times more cells than CA3 (O'Reilly and McClelland, 1994), this means that in absolute terms there are perhaps some 2.5 times more cells active in rat DG than CA3 during recall. In BACON each DG' cell innervates a single CA3' cell, but whereas in Recall mode the number of cells active in CA3' is limited to *K*, there is no limit in DG', and each cell there fires in proportion to its degree of excitation by sampled attributes that have potentiated synapses with it. In any given context, a given attribute would excite any cell it innervated that had previously become one of the representation cells of some context. All *K* DG' cells that had been active during the formation of the current representation would be excited plus presumably quite a number more. Thus, the observation that 2.5

**TABLE 3 | Some behavioral properties of BACON.**

Prediction	Evidence (for, against, nil)?
A. In CA3' only a small number ( $K$ ) of cells will fire. During the first exposure to a novel context, there should be a fairly sudden shift in the population of active cells as a representation for the context is created, and the population of active cells should thereafter be constant. For familiar contexts, the same cells should always be active during different exposures to the same context, and substantially different cells active in different contexts. Even quite similar familiar contexts should activate fairly different populations of cells	For <sup>a,b</sup>
B. In DG' patterns of activity will be very different for the remainder of sessions in which new representations are created (i.e., in novel contexts) than at other times. Following representation creation activity patterns will have the same properties as those in CA3' (few cells active; little overlap in cells active in different contexts). At other times cells will be active at graded rates, and firing rate profiles across cells will differ more, the more different the contexts	For <sup>a,b,c</sup>
C. Upon introduction to a distinctive <i>unfamiliar</i> context, context fear conditioning will be impossible until after a delay (we refer to the delay for an unfamiliar context as " $D_0$ ," which is about 30 s.), after which conditionability gradually increases (the immediate shock deficit)	For <sup>d,e,f</sup>
D. Upon introduction to a <i>familiar</i> context, there will also be a delay before context conditioning can occur, but it will be a time $D$ that is shorter than $D_0$	For <sup>e,g</sup>
E. Repeated exposures of duration less than the $D_0$ to an unfamiliar context will not shorten $D$	Nil
F. No amount of pre-familiarization will completely abolish the immediate shock deficit (i.e., reduce $D$ to zero)	For <sup>e</sup> Against <sup>g,h</sup>
G. Upon introduction to a feared context, expression of fear will increase gradually	For <sup>i,j,k</sup>
H. Fear expression will rise gradually from very near the start of a testing session (in contrast to conditionability, which will begin to rise when $B_{Rep}$ reaches $B_{Old}$ ). However, conditionability and expression will reach their maxima at the same time (when $B_{Rep}$ reaches $B_{Add}$ ) and thereafter be constant	Nil
I. Generalization of fear to familiar contexts will be less than to unfamiliar ones	Nil
J. Generalization of fear from more familiar contexts will be less than from less familiar ones	For <sup>h,l</sup>
K. Conditioning to a novel context that is very similar to a familiar one may be to the familiar rather than the actual context	For <sup>l,o</sup>
L. The above effect will be less, the greater the delay of the US	Nil
M. Generalization tests will often show initial fear that later falls off much too fast to be explained by known rates of extinction	Nil
N. Loss of indirect path input to CA3' during both encoding and recall does not seriously disrupt encoding or recall, but does reduce ability to distinguish similar contexts. However, inactivation of DG' during encoding alone has substantial effects on recall of contextual fear	For <sup>m,p</sup>
O. Inactivation of DG' plasticity has no effect on selection of cells during encoding, but reduces ability to distinguish similar contexts during recall, because without EC'-DG' plasticity recall must occur entirely via direct path	For <sup>n</sup>

<sup>a</sup>Deng et al. (2013). <sup>b</sup>Leutgeb et al. (2007). <sup>c</sup>Neunuebel and Knierim (2014). <sup>d</sup>Fanselow (1986). <sup>e</sup>Fanselow (1990). Also Fanselow, unpublished. <sup>f</sup>Blanchard et al. (1976). <sup>g</sup>Matus-Amat et al. (2004). During pre-exposures and conditioning, rats were consistently carried to the test chamber in the same covered black bucket; the black bucket could therefore be considered a part of the context, and "immediate" shocks should be seen as in-effect occurring at the transport-time after the beginning of contextual exposure. <sup>h</sup>Kiernan and Westbrook (1993). Whether, as in the previous experiment, a consistent environment immediately preceding nominal context entry may have made the "immediate" shock in-effect somewhat delayed is not known. <sup>i</sup>Wiltgen et al. (2006). <sup>j</sup>Fanselow (1982). <sup>k</sup>Lester and Fanselow (1986). <sup>l</sup>Rudy and O'Reilly (1999). <sup>m</sup>Nakashiba et al. (2012). <sup>n</sup>McHugh et al. (2007). <sup>o</sup>Rudy et al. (2002). <sup>p</sup>Kheirbek et al. (2013).

times more cells are active in DG relative to CA3 in a familiar environment seems consistent with what BACON would predict (see also Supplementary Material Topic F).

### Inactivation of Dentate

Several studies have examined the effect of inactivating DG'. In one study, dentate-CA3' transmission was suppressed with tetanus toxin during encoding, conditioning, and recall with only very modest effects. Fear expressed during a test in the conditioning context was roughly normal, as was degree of generalization to a quite different context, but there were some deficits in discrimination learning (Nakashiba

et al., 2012). Apparently consistent with this, a more recent study found apparently normal fear during a test with DG' inactivated following encoding and conditioning with DG intact (Kheirbek et al., 2013). However, if DG was suppressed during encoding and conditioning but *not* during recall, fear of the conditioned context was diminished (Kheirbek et al., 2013).

BACON would emulate these observations under similar conditions: If DG' were inoperative during both encoding and recall, the indirect pathway would be used in both cases, and recall would be fairly normal, though discrimination of similar contexts would be somewhat reduced because one would lose

the cooperative advantages of using both pathways, as discussed above. However, if DG' were suppressed only during encoding, there would be a great deficit in conditioned fear. The CA3' cells selected to comprise the representation would be those most richly innervated by active EC<sub>in</sub> neurons via the *direct* path, and it would be only they that would become potentiated as part of the representation creation process. However, during recall with both pathways operative, input via the *indirect* path dominates. But because synapses to DG' did not become potentiated during encoding, the input to CA3' that determines CA3Ptrn<sub>o</sub> will be via DG' input synapses that had become potentiated during *previous* encodings when DG' was intact and will not promote the firing of representation neurons for the *current* context. If, on the other hand, DG' were suppressed only during recall, there would be only a modest effect, as seen in **Figure 8III** and as discussed in "Direct/Indirect pathway balance in recall," above (**Table 3, N**). There would also be only a modest deficit if DG' plasticity were permanently eliminated: During encoding DG' would still select DG'-CA3' dyads, but since potentiation would be impossible in DG', it would be EC<sub>in</sub>-CA3' synapses that would become potentiated and determine what CA3' cells would be activated during recall. Consistent with findings of McHugh et al. (2007), there would be some deficits in discrimination ability but no major dysfunction.

### Direct Optogenetic Activation of Representation Cells

Liu et al. (2012) recently studied mice in which dentate cells active during encoding of a context A were labeled with channelrhodopsin so that they could be selectively activated by stimulating dentate with light. The mice were then fear conditioned in A but tested in a different context B, with and without the dentate light stimulus on. As expected, the light stimulus did evoke fear reactions. In BACON direct activation of the DG' cells that were active during encoding of A would evoke a CA3Ptrn<sub>Fnl</sub> that was a complete representation of context A even if only a portion of the relevant DG' cells were successfully activated. However, there would be a definite mismatch between the attributes of context B observed during the test and the recalled attributes of context A activated by the representation. This mismatch would degrade  $B_{Rep}$  and decrease fear expression. Thus, contextual fear would be expressed when the context A representation was activated in context B, but consistent with the behavioral findings, the fear would be of sub-normal magnitude.

### Predictions, Tested, and Untested

BACON emulates many known aspects of real contextual fear conditioning (Results and **Table 3**), suggesting the possibility that biological context fear circuitry may work in a somewhat BACON-like way. It should be appreciated that the algorithm's ability to do this was not explicitly built in; it is a natural consequence of designing an algorithm that deals in an adaptive way with the problems posed by context fear learning. The algorithm also predicts a number of behavioral characteristics that are presently unknown (**Table 3**) and that seem worth experimental test. However, it should be emphasized that predictions that depend on the algorithm's Bayesian evaluations of context identity could well be falsified without it ruling out the

more general idea that hippocampus operates in several distinctly different modes and that control of these is determined in some principled (but not necessarily Bayesian) way by comparisons between attributes of the current context and those associated with contexts from the animal's past.

### Unfinished Business

The present work grew out of an earlier attempt to construct a fairly comprehensive model of fear conditioning that included cue as well as context conditioning (Krasne et al., 2011). That model accounted for much of the phenomenology of fear conditioning; however, hippocampal representations of context were assumed to develop gradually during contextual exposure, without any theory as to mechanism. The current algorithm supplies a mechanism, and ultimately the two models should be merged.

A number of simplifying assumptions were made during the design of BACON, and ways of moving beyond them should be explored in future work. Some matters especially needing consideration are these: (1) A theory of sampling. How does attribute sampling vary over time; does it slow as the set of attributes so-far observed during a session becomes more complete? (2). We assumed that all contexts have the same number of attributes. Ways of dealing with cases where this is an implausible assumption are needed. Weight of Evidence calculations depend on the total number of attributes available for sampling in a situation (our  $N_A$ ). How might this number be estimated on the fly without extensive prior sampling of a context? (3) It seems not unlikely that sampling slows or stops during freezing. If so, it would greatly affect BACON's behavior and needs to be considered as part of a theory of sampling. (4) Random sampling is a useful simplifying assumption, but some attributes are surely more salient than others. This ultimately should be incorporated into a model such as this. (5) BACON assumes that existing representations are activated serially and evaluated for validity as they become active; a more sophisticated algorithm that can compare validities of multiple attributes in parallel would produce better performance. (6) Extinction, which we chose to ignore, must be taken into account. It will be introduced into the algorithm at a later time by assuming, as done in the above-mentioned model of the amygdala, that extinction occurs because unreinforced fear responses cause potentiation at synapses between active CA3' neurons and amygdala interneurons that inhibit BACON's fear-producing neurons.

### Conclusion

We have constructed an algorithm, BACON, which randomly and serially samples attributes of its current context, uses its hippocampus to activate representations of whatever known context best matches the current one, and computes the Bayesian Weight of Evidence that the hippocampal representation is or is not that of the actual context. If this value indicates that the representation is valid, context fear conditioning and expression are enabled. If not, the hippocampus is configured to create a representation for the apparently novel context.

This machinery works well, accounts for several well-known aspects of context fear learning, such as the immediate shock deficit, and makes new predictions. BACON's use of external control to enable hippocampal representation creation does away with several problems that have plagued thinking about hippocampal function (especially the conflict between pattern separation and completion and problems that arise if the "indirect" pathway from entorhinal cortex to CA3 is used both during representation creation and recall). Overall, we believe that the approach taken provides insight into both context fear learning mechanisms and hippocampal function in general.

## References

- Bast, T., Zhang, W. N., and Feldon, J. (2003). Dorsal hippocampus and classical fear conditioning to tone and context in rats: effects of local NMDA-receptor blockade and stimulation. *Hippocampus* 13, 657–675. doi: 10.1002/hipo.10115
- Blair, H. T., Schafe, G. E., Bauer, E. P., Rodrigues, S. M., and LeDoux, J. E. (2001). Synaptic plasticity in the lateral amygdala: a cellular hypothesis of fear conditioning. *Learn. Mem.* 8, 229–242. doi: 10.1101/lm.30901
- Blanchard, R. J., Fukunaga, K. K., and Blanchard, D. C. (1976). Environmental control of defensive reactions to footshock. *Bull. Psychon. Soc.* 8, 129–130. doi: 10.3758/BF03335103
- Boatman, J. A., and Kim, J. J. (2006). A thalamo-cortico-amygdala pathway mediates auditory fear conditioning in the intact brain. *Eur. J. Neurosci.* 24, 894–900. doi: 10.1111/j.1460-9568.2006.04965.x
- Bortolotto, Z. A., Lauri, S., Issaac, J. T., and Collingridge, G. L. (2003). Kainate receptors and induction of mossy fibre long-term potentiation. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* 358, 657–666. doi: 10.1098/rstb.2002.1216
- Calandrea, L., Desmedt, A., Decorte, L., and Jaffard, R. (2005). A different recruitment of the lateral and basolateral amygdala promotes contextual or elemental conditioned association in Pavlovian fear conditioning. *Learn. Mem.* 12, 383–388. doi: 10.1101/lm.92305
- Contractor, A., Rogers, C., Marion, C., Henkeneyer, M., Swanson, G. T., and Heinemann, S. F. (2002). Trans-synaptic Eph receptor signaling in hippocampal mossy fiber LTP. *Science* 296, 1864–1869. doi: 10.1126/science.1069081
- Cowansage, K. K., Shuman, T., Dillingham, B. C., Chang, A., Golshani, P., and Mayford, M. (2014). Direct reactivation of a coherent neocortical memory of context. *Neuron* 84, 432–441. doi: 10.1016/j.neuron.2014.09.022
- Deng, W., Mayford, M., and Gage, F. H. (2013). Selection of distinct populations of dentate granule cells in response to inputs as a mechanism for pattern separation in mice. *Elife* 2:e00312. doi: 10.7554/eLife.00312
- Fanselow, M. S. (1982). The post-shock activity burst. *Anim. Learn. Behav.* 10, 448–454. doi: 10.3758/BF03212284
- Fanselow, M. S. (1986). Associative vs. topographical accounts of the immediate shock freezing deficit in rats: implications for the response selection rules governing species specific defensive reactions. *Learn. Motiv.* 17, 16–39. doi: 10.1016/0023-9690(86)90018-4
- Fanselow, M. S. (1990). Factors governing one-trial contextual conditioning. *Anim. Learn. Behav.* 18, 264–270. doi: 10.3758/BF03205285
- Fanselow, M. S. (2000). Contextual fear, gestalt memories, and the hippocampus. *Behav. Brain Res.* 110, 73–81. doi: 10.1016/S0166-4328(99)00186-2
- Fanselow, M. S., and Dong, H. W. (2010). Are the dorsal and ventral hippocampus functionally distinct structures? *Neuron* 65, 7–19. doi: 10.1016/j.neuron.2009.11.031
- Fanselow, M. S., and LeDoux, J. E. (1999). Why we think plasticity underlying Pavlovian fear conditioning occurs in the basolateral amygdala. *Neuron* 23, 229–232. doi: 10.1016/S0896-6273(00)80775-8
- Favero, M., Massella, O., Cangiano, A., and Buffelli, M. (2009). On the mechanism of action of muscle fibre activity in synapse competition and elimination at the mammalian neuromuscular junction. *Eur. J. Neurosci.* 29, 2327–2334. doi: 10.1111/j.1460-9568.2009.06779.x

## Acknowledgments

This work was supported in part by NIH Grant Number RO1MH062122 to MF and UCLA Office of Research Administration Grant Number 444040-XH444042-XH-07427 to JC and FK.

## Supplementary Material

The Supplementary Material for this article can be found online at: <http://journal.frontiersin.org/article/10.3389/fnbeh.2015.00112/abstract>

- Frey, U., Hartmann, S., and Matthies, H. (1989). Domperidone, an inhibitor of the D2-receptor, blocks a late phase of an electrically induced long-term potentiation in the CA1-region in rats. *Biomed. Biochim. Acta* 48, 473–476.
- Gallistel, C. R. (2009). The importance of proving the null. *Psychol. Rev.* 116, 439–453. doi: 10.1037/a0015251
- Gardner-Medwin, A. R. (1976). The recall of events through the learning of associations between their parts. *Proc. R. Soc. Lond. B.* 194, 375–402. doi: 10.1098/rspb.1976.0084
- Hasselmo, M. E. (2006). The role of acetylcholine in learning and memory. *Curr. Opin. Neurobiol.* 16, 710–715. doi: 10.1016/j.conb.2006.09.002
- Hasselmo, M. E., and Sarter, M. (2010). Modes and models of forebrain cholinergic neuromodulation of cognition. *Neuropsychopharmacology* 36, 52–73. doi: 10.1038/npp.2010.104
- Hasselmo, M. E., Schnell, E., and Barkai, E. (1995). Dynamics of learning and recall at excitatory recurrent synapses and cholinergic modulation in rat hippocampal region CA3. *J. Neurosci.* 15, 5249–5262.
- Hawkins, R. D. (2013). Possible contributions of a novel form of synaptic plasticity in Aplysia to reward, memory, and their dysfunctions in mammalian brain. *Learn. Mem.* 20, 580–591. doi: 10.1101/lm.031237.113
- Huerta, P. T., and Lisman, J. E. (1995). Bidirectional synaptic plasticity induced by a single burst during cholinergic theta oscillation in CA1 *in vitro*. *Neuron* 15, 1053–1063. doi: 10.1016/0896-6273(95)90094-2
- Jin, J., and Maren, S. (2015). Fear renewal preferentially activates ventral hippocampal neurons projecting to both amygdala and prefrontal cortex in rats. *Sci. Rep.* 5:8388. doi: 10.1038/srep08388
- Johansen, J. P., Diaz-Mataix, L., Hamanaka, H., Ozawa, T., Ycu, E., Koivumaa, J., et al. (2014). Hebbian and neuromodulatory mechanisms interact to trigger associative memory formation. *Proc. Natl. Acad. Sci. U.S.A.* 111, E5584–E5592. doi: 10.1073/pnas.1421304111
- Kano, M., and Hashimoto, K. (2009). Synapse elimination in the central nervous system. *Curr. Opin. Neurobiol.* 19, 154–161. doi: 10.1016/j.conb.2009.05.002
- Kass, R. E., and Raftar, A. E. (1995). Bayes factors. *J. Am. Stat. Assoc.* 90, 773–795. doi: 10.1080/01621459.1995.10476572
- Kenney, J. W., and Gould, T. J. (2008). Modulation of hippocampus-dependent learning and synaptic plasticity by nicotine. *Mol. Neurobiol.* 38, 101–121. doi: 10.1007/s12035-008-8037-9
- Kesner, R. P., and Rolls, E. T. (2015). A computational theory of hippocampal function, and tests of the theory: new developments. *Neurosci. Biobehav. Rev.* 48, 92–147. doi: 10.1016/j.neubiorev.2014.11.009
- Ketz, N., Morkonda, S. G., and O'Reilly, R. C. (2013). Theta coordinated error-driven learning in the hippocampus. *PLOS Comput. Biol.* 9:e1003067. doi: 10.1371/journal.pcbi.1003067
- Kheirbek, M. A., Drew, L. J., Burghardt, N. S., Costantini, D. O., Tannenholz, L., Ahmari, S. E., et al. (2013). Differential control of learning and anxiety along the dorsoventral axis of the dentate gyrus. *Neuron* 77, 955–968. doi: 10.1016/j.neuron.2012.12.038
- Kiernan, M., and Cranney, J. (1992). Immediate-startle stimulus presentation fails to condition freezing responses to contextual cues. *Behav. Neurosci.* 106, 121–124. doi: 10.1037/0735-7044.106.1.121



- Kiernan, M. J., and Westbrook, R. F. (1993). Effects of exposure to a to-be-shocked environment upon the rat's freezing response: evidence for facilitation, latent inhibition, and perceptual learning. *Q. J. Exp. Psychol.* 46, 271–288. doi: 10.1080/14640749308401089
- Krasne, F. B., Fanselow, M. S., and Zelikowsky, M. (2011). Design of a neurally plausible model of fear learning. *Front. Behav. Neurosci.* 5:41. doi: 10.3389/fnbeh.2011.00041
- Ladeira-Fernandez, J., Decola, J. P., Kim, J. J., and Fanselow, M. S. (2006). Immediate shock deficit in fear conditioning: effects of shock manipulations. *Behav. Neurosci.* 120, 873–879. doi: 10.1037/0735-7044.120.4.873
- LeDoux, J. E., Cicchetti, P., Xagoraris, A., and Romanski, L. M. (1990). The lateral amygdaloid nucleus: sensory interface of the amygdala in fear conditioning. *J. Neurosci.* 10, 1062–1069.
- LeDoux, J. E., and Clugnet, M. C. (1990). Synaptic plasticity in fear conditioning circuits: induction of LTP in the lateral nucleus of the amygdala by stimulation of the medial geniculate body. *J. Neurosci.* 10, 2818–2824.
- Lester, L. S. and Fanselow, M. S. (1986). Naloxone's enhancement of freezing: Modulation of perceived intensity or memory processes? *Physiol. Psychol.* 14, 5–10.
- Leutgeb, J. K., Leutgeb, S., Moser, M. B., and Moser, E. I. (2007). Pattern separation in the dentate gyrus and CA3 of the hippocampus. *Science* 315, 961–966. doi: 10.1126/science.1135801
- Lisman, J., Grace, A. A., and Duzel, E. A. (2011). neoHebbian framework for episodic memory; role of dopamine-dependent late LTP. *Trends Neurosci.* 34, 536–547. doi: 10.1016/j.tins.2011.07.006
- Liu, X., Ramirez, S., Pang, P. T., Puryear, C. B., Govindarajan, A., and Deisseroth, K. S. (2012). Optogenetic stimulation of a hippocampal engram activates fear memory recall. *Nature* 484, 381–385. doi: 10.1038/nature11028
- Lovinger, D. M. (2010). Neurotransmitter roles in synaptic modulation, plasticity and learning in the dorsal striatum. *Neuropharmacology* 58, 951–961. doi: 10.1016/j.neuropharm.2010.01.008
- Maren, S., and Fanselow, M. S. (1995). Synaptic plasticity in the basolateral amygdala induced by hippocampal formation stimulation *in vivo*. *J. Neurosci.* 15, 7548–7564.
- Maren, S., and Holt, W. G. (2004). Hippocampus and Pavlovian fear conditioning in rats: muscimol infusions into the ventral, but not dorsal, hippocampus impair the acquisition of conditional freezing to an auditory conditional stimulus. *Behav. Neurosci.* 118, 97–110. doi: 10.1037/0735-7044.118.1.97
- Markram, H., and Segal, M. (1990). Acetylcholine potentiates responses to N-methyl-D-aspartate in the rat hippocampus. *Neurosci. Lett.* 113, 62–65. doi: 10.1016/0304-3940(90)90495-U
- Marr, D. (1971). Simple memory: a theory for archicortex. Philosophical transactions of the royal society of london. *Ser. B Biol. Sci.* 262, 23–81. doi: 10.1098/rstb.1971.0078
- Matus-Amat, P., Higgins, E. A., Barrientos, R. M., and Rudy, J. W. (2004). The role of the dorsal hippocampus in the acquisition and retrieval of context memory representations. *J. Neurosci.* 24, 2431–2439. doi: 10.1523/JNEUROSCI.1598-03.2004
- McClelland, J. L., McNaughton, B. L., and O'Reilly, R. C. (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. doi: 10.1037/0033-295X.102.3.419
- McHugh, T. J., Jones, M. W., Quinn, J. J., Balthasar, N., Coppari, R., Elmquist, J. K., et al. (2007). Dentate gyrus NMDA receptors mediate rapid pattern separation in the hippocampal network. *Science* 207, 94–99. doi: 10.1126/science.1140263
- McNaughton, B. L., and Morris, R. G. M. (1987). Hippocampal synaptic enhancement and information storage within a distributed memory system. *Trends Neurosci.* 10, 408–415. doi: 10.1016/0166-2236(87)90011-7
- Myers, C. E., and Scharfman, H. E. (2011). Pattern separation in the dentate gyrus: a role for the CA3 backprojection. *Hippocampus* 21, 1190–1215. doi: 10.1002/hipo.20828
- Nakashiba, T., Cushman, J. D., Pelkey, K. A., Renaudineau, S., Buhl, D. L., McHugh, T. J., et al. (2012). Young dentate granule cells mediate pattern separation, whereas old granule cells facilitate pattern completion. *Cell* 149, 188–201. doi: 10.1016/j.cell.2012.01.046
- Neunuebel, J. P., and Knierim, J. J. (2014). CA3 retrieves coherent representations from degraded input: direct evidence for CA3 pattern completion and dentate gyrus pattern separation. *Neuron* 81, 416–427. doi: 10.1016/j.neuron.2013.11.017
- Onishi, B. K., and Xavier, G. F. (2010). Contextual, but not auditory, fear conditioning is disrupted by neurotoxic selective lesion of the basal nucleus of amygdala in rats. *Neurobiol. Learn. Mem.* 93, 165–174. doi: 10.1016/j.nlm.2009.09.007
- O'Reilly, R. C., and McClelland, J. L. (1994). Hippocampal conjunctive encoding, storage, and recall: avoiding a trade-off. *Hippocampus* 4, 661–682. doi: 10.1002/hipo.450040605
- O'Reilly, R. C., and Rudy, J. W. (2001). Conjunctive representations in learning and memory: principles of cortical and hippocampal function. *Psychol. Rev.* 108, 311–345. doi: 10.1037/0033-295X.108.2.311
- Parsons, T. C., and Otto, T. (2008). Temporary inactivation of dorsal hippocampus attenuates explicitly nonspatial, unimodal, contextual fear conditioning. *Neurobiol. Learn. Mem.* 90, 261–268. doi: 10.1016/j.nlm.2008.03.007
- Reid, C. A., Dixon, D. B., Takahashi, M., Bliss, T. V. P., and Fine, A. (2004). Optical quantal analysis indicates that long-term potentiation at single hippocampal mossy fiber synapses is expressed through increased release probability, recruitment of new release sites, and activation of silent synapses. *J. Neurosci.* 24, 3618–3626. doi: 10.1523/JNEUROSCI.3567-03.2004
- Romanski, L. M., and LeDoux, J. E. (1992). Equipotentiality of thalamo-amygdala and thalamo-cortico-amygdala circuits in auditory fear conditioning. *J. Neurosci.* 12, 4501–4509.
- Rudy, J. W. (2009). Context representations, context functions, and the parahippocampal-hippocampal system. *Learn. Mem.* 16, 573–585. doi: 10.1101/lm.1494409
- Rudy, J. W., Barrientos, R. M., and O'Reilly, R. C. (2002). Hippocampal formation supports conditioning to memory of a context. *Behav. Neurosci.* 116, 530–538. doi: 10.1037/0735-7044.116.4.530
- Rudy, J. W., and O'Reilly, R. C. (1999). Contextual fear conditioning, conjunctive representations, pattern completion, and the hippocampus. *Behav. Neurosci.* 113, 867–880. doi: 10.1037/0735-7044.113.5.867
- Rudy, J. W., and O'Reilly, R. C. (2001). Conjunctive representations, the hippocampus, and contextual fear conditioning. *Cogn. Affect. Behav. Neurosci.* 1, 66–82. doi: 10.3758/CABN.1.1.66
- Scharfman, H. E. (2007). The CA3 “Backprojection” to the dentate gyrus. *Prog. Brain Res.* 163, 627–637. doi: 10.1016/S0079-6123(07)63034-9
- Schenberg, E. E., and Oliveira, M. G. (2008). Effects of pre or posttraining dorsal hippocampus D-AP5 injection on fear conditioning to tone, background, and foreground context. *Hippocampus* 18, 1089–1093. doi: 10.1002/hipo.20475
- Schon, K., Newmark, R. E., Ross, R. S., and Stern, C. E. (2015). A working memory buffer in parahippocampal regions: evidence from a load effect during the delay period. *Cereb. Cortex* bhv013. doi: 10.1093/cercor/bhv013. [Epub ahead of print].
- Skaggs, W. E., and McNaughton, B. L. (1992). Computational approaches to hippocampal function. *Curr. Opin. Neurobiol.* 2, 209–211. doi: 10.1016/0959-4388(92)90014-C
- Stote, D. L., and Fanselow, M. S. (2004). NMDA receptor modulation of incidental learning in Pavlovian context conditioning. *Behav. Neurosci.* 118, 253–257. doi: 10.1037/0735-7044.118.1.253
- Suzuki, W. A., Miller, E. K., and Desimone, R. (1997). Object and place memory in the macaque entorhinal cortex. *J. Neurophysiol.* 78, 1062–1081.
- Tanaka, K. Z., Pevzner, A., Hamidi, A. B., Nakazawa, Y., Graham, J., and Wiltgen, B. J. (2014). Cortical representations are reinstated by the hippocampus during memory retrieval. *Neuron* 84, 347–354. doi: 10.1016/j.neuron.2014.09.037
- Thomas, M. J., Moody, T. D., Makhinson, M., and O'Dell, T. J. (1996). Activity-dependent beta-adrenergic modulation of low frequency stimulation induced LTP in the hippocampal CA1 region. *Neuron* 17, 475–482. doi: 10.1016/S0896-6273(00)80179-8
- Treves, A. (1995). Quantitative estimate of the information relayed by the Schaffer collaterals. *J. Comput. Neurosci.* 2, 259–272. doi: 10.1007/BF00961437
- Treves, A., and Rolls, E. T. (1992). Computational constraints suggest the need for two distinct input systems to the hippocampal CA3 network. *Hippocampus* 2, 189–200. doi: 10.1002/hipo.450020209

- Treves, A., and Rolls, E. T. (1994). Computational analysis of the role of the hippocampus in memory. *Hippocampus* 4, 374–391. doi: 10.1002/hipo.450040319
- Treves, A., Tashiro, A., Witter, M. P., and Moser, E. I. (2008). What is the mammalian dentate gyrus good for? *Neuroscience* 154, 1155–1172. doi: 10.1016/j.neuroscience.2008.04.073
- Wiltgen, B. J., Sanders, M. J., Anagnostaras, S. G., Sage, J. R., and Fanselow, M. S. (2006). Context fear learning in the absence of the hippocampus. *J. Neurosci.* 26, 5484–5491. doi: 10.1523/JNEUROSCI.2685-05.2006
- Young, B. J., Otto, T., Fox, G. D., and Eichenbaum, H. (1997). Memory representation within the parahippocampal region. *J. Neurosci.* 17, 5183–5195.
- Young, S. L., Bohenek, D. L., and Fanselow, M. S. (1994). NMDA processes mediate anterograde amnesia of contextual fear conditioning induced by hippocampal damage: immunization against amnesia by context preexposure. *Behav. Neurosci.* 108, 19–29. doi: 10.1037/0735-7044.108.1.19
- Conflict of Interest Statement:** The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

*Copyright © 2015 Krasne, Cushman and Fanselow. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.*