



Published in final edited form as:

*Psychon Bull Rev.* 2011 June ; 18(3): 455–468. doi:10.3758/s13423-011-0086-9.

## Human memory reconsolidation can be explained using the Temporal Context Model

**Per B. Sederberg,**

Department of Psychology, The Ohio State University, Columbus, OH 43210, USA

**Samuel J. Gershman,**

Department of Psychology and Princeton Neuroscience Institute, Princeton University, Princeton, NJ 08540, USA

**Sean M. Polyn,** and

Department of Psychology, Vanderbilt University, Nashville, TN 37240, USA

**Kenneth A. Norman**

Department of Psychology and Princeton Neuroscience Institute, Princeton University, Princeton, NJ 08540, USA

### Abstract

Recent work by Hupbach et al. (2007, 2009) suggests that episodic memory for a previously studied list can be updated to include new items, if participants are reminded of the earlier list just prior to learning a new list. The key finding from the Hupbach studies was an asymmetric pattern of intrusions, whereby participants intruded numerous items from the second list when trying to recall the first list, but not vice-versa. Hupbach et al. explain this pattern in terms of a cellular reconsolidation process, whereby the first-list memory is rendered labile by the reminder, and the labile memory is then updated to include items from the second list. Here we show that the Temporal Context Model (TCM) of memory, which lacks a cellular reconsolidation process, can account for the asymmetric intrusion effect using well-established principles of contextual reinstatement and item-context binding.

### Introduction

One of the most provocative and exciting ideas to emerge from the animal learning and memory literature in recent years is the idea of *reconsolidation*. According to this idea, retrieving a memory makes its molecular substrate malleable; when the memory is in this malleable state, it can be changed or even erased (for reviews, see Dudai & Eisenberg, 2004; Riccio, Millin, & Bogart, 2006; Dudai, 2009; Lee, 2009; Wang & Morris, 2010). The recent excitement over reconsolidation can be traced back to a fear conditioning study conducted by Nader, Schafe, and Le Doux (2000). In this study, Nader et al. demonstrated that a well-learned tone-shock fear conditioning memory could be erased, in an apparently permanent fashion, if rats were first reminded of the original association (by presenting the tone by itself) and then were injected with a protein synthesis inhibitor; both factors (the reminder and the injection) were necessary in order to get this effect. Nader et al. argued that the reminder initiated a window of vulnerability (during which the memory had to be reconsolidated at a molecular level) and the protein synthesis inhibitor blocked this reconsolidation process (thereby causing permanent erasure of the memory). Subsequent to the Nader study, this basic finding (forgetting after a reminder and a protein synthesis

blocker) has been extended to other animals (e.g., crabs: Pedreira, Pérez-Cuesta, & Maldonado, 2004; medaka fish: Eisenberg, Kobil, Berman, & Dudai, 2003) and other learning paradigms (e.g., spatial learning in the Morris water maze: Morris et al., 2006).

More recently, researchers have started to explore the relevance of reconsolidation to human learning and memory (for a review, see Hardt, Einarsson, & Nader, 2010). The protein synthesis inhibitors used in the animal studies reviewed above are too toxic to use in human studies. However, researchers have started to devise other ways to explore reconsolidation in humans using purely behavioral means. For example, a recent series of studies by Hupbach and colleagues (Hupbach, Gomez, Hardt, & Nadel, 2007; Hupbach, Gomez, & Nadel, 2009) used a simple list-learning paradigm and found that list memories can be altered by first reminding participants of a previously-studied list (to make the memory malleable) and then giving participants new material to learn; the focus of these studies was not on erasing the original list memory, but rather on showing that the original list memory could be updated with new information when it was in a malleable state (induced by the reminder). These Hupbach et al. studies are described in more detail below. In addition to the Hupbach et al. studies, several other human memory studies using different types of paradigms (procedural learning in sleep: Walker, Brakefield, Hobson, & Stickgold, 2003; fear conditioning: Schiller et al., 2009) have tested the basic idea that reminding participants of a previously formed memory can make that memory malleable.

The goal of the present study is to explore the implications of this “human reconsolidation” work for theories of human learning and memory (in particular, theories of declarative memory: recalling and recognizing lists of stimuli). Over the past half-century, researchers working with human memory data have built up a rich vocabulary of mechanisms, instantiated in mathematical and computational models, to account for behavioral recall and recognition results (for reviews, see Norman, Detre, & Polyn, 2008; Raaijmakers & Shiffrin, 2002; Raaijmakers, 2005). Do these theories need to be fundamentally revised to account for the findings mentioned above? Answering this question is challenging because reconsolidation theory and mathematical memory models have traditionally been described at different levels of analysis (reconsolidation in terms of molecular/cellular processes, and mathematical memory models in terms of abstract, implementation-independent algorithms for storing and retrieving memory traces). One possibility is that reconsolidation theory and cognitive models describe the same process at different levels of analysis: In this case, it should be possible to fit behavioral data from reconsolidation studies using existing cognitive/mathematical models, even if these models do not explicitly discuss the molecular/cellular mechanisms that determine the malleability of the underlying memory trace. Another possibility is that reconsolidation entails new cognitive mechanisms that are absent from existing models; in this case, it will not be possible to account for behavioral data from reconsolidation studies using existing models.

To address these questions, we set out to explore whether an existing computational model of memory retrieval (the Temporal Context Model, TCM; Howard & Kahana, 2002; Sederberg, Howard, & Kahana, 2008) could explain findings from Hupbach and colleagues that have been cited as evidence for reconsolidation in humans. We selected the Hupbach findings (out of all of the “human reconsolidation” studies mentioned above) because these findings fall squarely within the data domain (list-learning in humans) that extant computational models were designed to address. Regarding our choice of TCM: As we discuss later, the properties that allow TCM to account for the Hupbach results are shared with several other models. As such, TCM should be viewed as a representative of a larger class of models; however, TCM also gives us the ability to make predictions about recall dynamics that can be tested in subsequent experiments – these predictions are described in the Discussion section.

In the next section, we discuss what Hupbach found and why these findings have been described as evidence for reconsolidation; then, we discuss the Temporal Context Model and how it can be used to address these findings.

### The Hupbach et al. (2007) list-learning paradigm

The paradigm used by Hupbach et al. (2007) involved participants studying two separate lists. On Monday, participants studied a set of objects that were pulled (one by one) out of a blue basket (call this *List A*). After a day off, participants returned to the lab on Wednesday and studied a new set of objects (this time, instead of pulling the objects from a blue basket, the objects were all spread out over a table; call this *List B*). The key manipulation was that, prior to studying list B, some participants were given a reminder of the list A learning experience. In the reminder condition, the same experimenter that was present during the list A study phase took the participant back to the same room that was used during list A study phase; the experimenter showed the blue basket to the participant and asked whether they remembered studying the items in the basket (participants were stopped if they started to recall any of those items out loud); finally, the participant was allowed to learn the list B items (in the same room where they learned the list A items). In the no-reminder condition, a new experimenter took the participant to a new room to study the list B items. Hupbach et al. also varied whether participants were immediately tested after learning list B or whether they had another day off before their memory test. In the delayed-test condition, participants in both conditions (reminder and no-reminder) returned on Friday and were given a free recall test where they were either asked to recall list A items or they were asked to recall list B items. In the immediate-test condition, participants in both the reminder and no-reminder groups were asked to recall list A items immediately after they learned the B items to criterion on day 2 (Wednesday).

The key finding was an *asymmetric intrusion effect* in the reminder condition with the delayed test: Participants in this condition intruded a large number of list B items when they were asked to recall list A items, but they did not intrude a substantial number of list A items when they were asked to recall list B items. Participants in the no-reminder (delayed-test) condition did not intrude a substantial number of items on either test. Also, participants in the immediate-test condition (both with and without reminders) did not intrude a substantial number of list B items when asked to recall list A items.

Hupbach et al. explain the asymmetric intrusion effect on the delayed test in the following manner: Presenting the list A reminder (prior to list B study) makes the list A memory malleable; when list B items are presented (subsequent to the reminder) these list B items are used to update the newly-malleable list A memory. Later, when participants are asked to recall list A items, they recall both the actual list A items and also the list B items that were part of the “update” to the list A memory. To explain the lack of intrusions on the immediate test, Hupbach et al. refer to rodent data from Nader et al. (2000) showing that (in Nader’s fear conditioning paradigm) the effects of the reminder on memory for the original event are not apparent if memory is tested immediately after the reminder – rather, these effects only emerge after some time has passed. Nader et al. (2000) explain this by positing that the molecular changes underlying reconsolidation unfold slowly over time, and Hupbach et al. argue that similar principles account for the effects of immediate vs. delayed testing in their “human reconsolidation” study.

### A contextual reinstatement account of the Hupbach et al. data

We set out to determine whether these results can be explained with TCM, which has previously been applied to a wide range of list-learning phenomena. In recent years, TCM and similar models have been used to explain many of the fundamental properties of

episodic memory, including recency and contiguity effects, encoding task and semantic effects, and transitive inference (Sederberg et al., 2008; Polyn, Norman, & Kahana, 2009; Howard, Jing, Rao, Probyn, & Datey, 2009). In TCM, temporal context is modeled as a recency-weighted average of past experience. Encoding an item in TCM involves binding that item to the temporal context in which it is experienced. Retrieval involves cuing with the current state of temporal context. Finally, the successful retrieval of an item also retrieves the temporal context in which the item was encoded, which is used to update the current context for subsequent recalls.

We hypothesized that contextual reinstatement and item-context binding in TCM would be sufficient to explain the asymmetric intrusion effect described above. According to TCM, when items from list A are presented, they will be linked to the currently active set of contextual elements (call this the *list A context*). Note that the list A context will drift over time, but it will also contain features that are common to all list A items, owing to these items being presented in the same fashion, by the same experimenter, in the same room. When the reminder is presented (prior to studying list B), it will trigger reinstatement of the list A context. Items from list B will be linked with these (reinstated) list A context features, as well as contextual features uniquely associated with list B (call this the *list B context*). Note the crucial asymmetry: In the reminder condition, the list A context is linked to items from both list A and list B, whereas the list B context is only linked to items from list B. Because of this asymmetrical linkage, cuing with the list A context at test will trigger recall of both list A items (correct recalls) and list B items (intrusions), whereas cuing with the list B context will only trigger recall of list B items (correct recalls).

Our explanation for the lack of intrusions in the immediate-test condition centers on participants' use of a recall-to-reject strategy (Hintzman, Caulton, & Levitin, 1998). Specifically, we hypothesize that — on the immediate test — list B items come to mind, but participants can reject these list B items via a process where 1) they compare retrieved contextual information to the current test context, and 2) they reject items where there is a match. Intuitively, the idea is that — for each item that is recalled on the immediate test — participants ask themselves “did I study this today” (by comparing retrieved context for that item to the current context) and they reject items where the answer is “yes”. Crucially, this recall-to-reject strategy is not viable on the delayed test; on Friday, none of the retrieved items (from Monday or Wednesday) will match the “today” context (on Friday), so the presence/absence of contextual match can not be used to discriminate between list A and list B items.

While this is a plausible story, there is no guarantee that TCM will be able to reproduce the critical findings from Hupbach et al. (2007). Below, we present the simulations showing that TCM can successfully account for the aforementioned results from Hupbach et al., thereby providing a mechanistic explanation for these human reconsolidation findings.

## The Temporal Context Model

For the simulations in this manuscript we used the same implementation of TCM that was used in Sederberg et al. (2008). Called TCM-A, this version has a recall rule based on the Usher and McClelland (2001) leaky-accumulator decision model that allows it to capture a wide array of recall behaviors that the original TCM was not designed to capture. This section provides a high-level overview of the model; for additional details please refer to Sederberg et al. (2008).

The core of TCM is composed of a two-layer neural network, with an item layer  $\mathbf{f}$  and a context layer  $\mathbf{t}$ . These layers are fully connected via association matrices. The context-to-

item associations are stored in a matrix,  $\mathbf{M}^{TF}$ , that allows contextual states to cue items. The item-to-context matrix,  $\mathbf{M}^{FT}$ , allows items to recover previous states of context. Our use of the terms “item” and “context” to describe the two layers fits with the terminology used by previous TCM papers. However, we should emphasize that the functions of the two layers diverge in important ways from common-sense interpretations of the terms “item” and “context”. In the model, the difference between the item and context layers relates to the time scale over which things are represented, rather than the types of content that are represented. The item layer represents the identity of the current object, and it also includes information that is considered to be more contextual in nature, such as the room, the experimenter, and the task that is currently being performed. The job of the context layer is to represent *temporal context*, which is operationally defined as a recency-weighted running average of features that were active on the item layer.

In Sederberg et al. (2008), the item layer contained only units that represented (in a localist fashion) the identities of individual items. A key feature of the present simulations is that the item layer also contains additional units that represent the features of the experimental environment: a “list A” unit that represents the features of the list A environment (the room, the experimenter, and the study task) and a “list B” unit that represents the unique features of the list B environment (i.e., the features that differentiate it from the list A environment). The item layer also contains “day” units (one per day) that represent slowly-changing features unique to each day of the experiment. For example, you might be hungry on day 3 of the experiment but not day 1 or day 2 – this is an example of the kind of “day-specific” feature that would be represented by activation of the day 3 unit. The context layer contains the same number of units as the item layer. These units represent the same things as the corresponding item-layer units (i.e., object identity, features of the list environments, and features of the current day) but in a time-averaged fashion, as described below.

An important feature of the present model is that multiple item-layer units can be active at the same time. For example, if an object is presented in the list A environment on day 1, then the item-layer unit corresponding to that object, the “list A” item-layer unit, and the unit for day 1 will all be active to some extent. These activity levels (which we also refer to as *strength* levels) can vary from zero to one; when multiple item-layer units are active, the activity values of these units determine the relative influence of these units on the context layer. Consequently, both the item and context layers of the present model contain units that represent the objects being studied and the features of the environment in which those objects are studied. Whereas the specific items being processed may change rapidly during the course of the experiment, other environmental features change more slowly and stay activated throughout the study session (e.g., the list unit representing the room, task, and experimenter will stay active throughout the entire session, as will the unit representing the day-specific contextual features).

Patterns of item-layer activity are propagated to the context layer via the item-to-context association matrices. This input causes the pattern of activity in the context layer to evolve according to the contextual drift equation:

$$\mathbf{t}_i = \rho_i \mathbf{t}_{i-1} + \beta \mathbf{t}_i^{IN}, \quad (1)$$

where  $\beta$  is a parameter that determines the rate of contextual drift during encoding,  $\rho_i$  is a scaling parameter chosen at each time step such that  $\mathbf{t}_i$  is always of unit length, and  $\mathbf{t}_i^{IN}$  is the input at time step  $i$ . The input pattern  $\mathbf{t}_i^{IN}$  that drives the evolution of context in Eq. 1 is calculated from the item-to-context association matrix  $\mathbf{M}^{FT}$  and the item  $\mathbf{f}_i$  as:

$$\mathbf{t}_i^{IN} \propto \mathbf{M}^{FT} \mathbf{f}_i, \quad (2)$$

where the proportionality symbol reflects the fact that  $\mathbf{t}_i^{IN}$  is normalized to be of unit length before contributing to Eq. 1.

As in Sederberg et al. (2008), the item-to-context and context-to-item associative matrices have both pre-experimental and newly-learned (experimental) components, which are combined via the following equations:

$$\mathbf{M}^{FT} = (1 - \gamma_{FT}) \mathbf{M}_{pre}^{FT} + \gamma_{FT} \mathbf{M}_{exp}^{FT} \quad (3)$$

$$\mathbf{M}^{TF} = (1 - \gamma_{TF}) \mathbf{M}_{pre}^{TF} + \gamma_{TF} \mathbf{M}_{exp}^{TF}. \quad (4)$$

The pre-experimental item-to-context and context-to-item matrices encode pre-existing semantic relationships between items. In the simulations presented here, we were not interested in modeling the effects of semantic similarity, so we set the pre-experimental matrices equal to the identity matrix for all of our simulations (this corresponds to an assumption of orthogonal semantic representations for the different objects). The experimental item-to-context and context-to-item matrices encode newly-learned episodic associations; these matrices were initialized to zero at the start of each simulation. Note that  $\gamma_{TF}$  was fixed to 1.0 in the simulations presented here; this parameter setting means that recall of items (given contextual cues) was entirely driven by newly-formed episodic associations, as opposed to pre-existing semantic associations.

During encoding, items are bound via standard Hebbian association to the pattern of activation in the context layer that was present when each item is experienced. In other words, the weights in the newly-learned experimental matrices between a unit in the item layer and a unit in the context layer are increased by an amount proportional to the product of their activations. Consequently, unlike other memory models that posit direct connections between items, associations between items in TCM are mediated by their shared context. In order to make predictions regarding the effect of variable learning rates on the number of intrusions we also scaled item-to-context weights by an additional learning rate parameter  $\alpha$ . This parameter was fixed at 1.0 for all simulations reported in the Results section; in the *Novel predictions* section we describe the effects of selectively setting  $\alpha$  to 0.5 during list A learning or list B learning.

At retrieval, the current state of context provides the cue for each recall. In our simulations, recall is initiated for a particular study list (e.g., list A) by activating the item-layer unit representing that list, which then updates the context via Equation 1. This assumption (that context activity at the start of the test reflects the context that the participant is trying to retrieve, as opposed to the environmental context that the participant is actually in) is important, and we talk about it in more detail in the Discussion. Once recall has been initiated, recall progresses as in Sederberg et al. (2008): Context activates items via episodically-formed context-to-item associations and the items then compete for recall via a set of Usher and McClelland (2001) accumulators. After an item is recalled, it updates the context layer with a combination of itself and the context that was present when it was originally experienced (the relative proportions of these two types of updates are determined by the model parameter  $\gamma_{FT}$ ).

In addition to the standard recall mechanism in TCM-A, we have implemented a recall-to-reject mechanism based on the recency that items were learned. Consider a situation where the model is asked (on day 2) to recall list A items from day 1; if – when an item is retrieved – the retrieved context overlaps with the current day unit (i.e., the retrieved activation of the current day unit is greater than zero), the model will reject the response because it is too recent. Given that this recall-to-reject mechanism relies on the time-of-test context overlapping with the day unit in the retrieved context, it has no effect on the delayed (day 3) test because the day 3 context does not overlap with either the day 1 context (when list A was studied) or the day 2 context (when list B was studied).

The new state of context, whether or not the recalled item was rejected, serves as the cue for the next recall and this cycle continues until the end of the recall period, which, for these simulations, was set at 45 seconds (with the accumulator time step parameter  $\delta t$  set to 10 ms, “45 seconds” of retrieval time translates into 4500 accumulator time steps).

Finally, to simulate experiences that are orthogonal to the memory task, we presented a distractor item that was orthogonal to all of the other item-layer patterns, and we used this distractor item to drift context. We fixed  $\beta_{dist}$  at 0.99 to simulate the large contextual drift resulting from the passing of a day between list A study and list B study; we also used this method (with a different distractor pattern, orthogonal to the first distractor pattern) to simulate the passing of a day between list B study and the final test.

## Simulation and Results

Table 1 outlines the procedure followed by Hupbach et al. (2007) and the corresponding steps in our TCM simulation. As outlined above, the key manipulation in the study was whether the participants were reminded of learning list A prior to learning list B. In the reminder condition, the experimenter and the room were the same for lists A and B; also, the experimenter asked participants a reminder question (participants were asked to think back to the blue basket). In the no-reminder condition, a new experimenter took the participant to a new room where they were asked to learn list B items. Note that, for both the reminder and no-reminder conditions, the study task was different for list B than for list A (for list B, objects were set out on a table instead of being pulled from a blue basket).

We simulated the reminder manipulation by activating the list A unit (in the item layer) prior to list B study and allowing this list A unit to drift context. We also allowed the list A unit to stay active to some degree during list B study (this reflects the fact that key aspects of the list A study environment – the room and the experimenter – were persistently present during list B study in the reminder condition). In the no-reminder condition, the list A unit was also allowed to be active during list B study, but at a much reduced level (relative to the reminder condition). This low (but non-zero) residual level of list-A activation in the no-reminder condition reflects the fact that — even in the no-reminder condition — there were still some similarities between list B study and list A study (e.g., both study phases took place in the same building; they both involved studying miscellaneous objects).

In addition to the reminder vs. no-reminder manipulation, we also simulated the immediate vs. delayed test manipulation from Hupbach et al. (2007). To simulate the immediate-test condition, we allowed the context layer to drift slightly in the direction of an orthogonal distractor following the presentation of list B, while maintaining the same day 2 unit activation. This drift is meant to capture the fading of the context representation that takes place in the (relatively short) interval between the end of list B study and the start of the list A test. For the delayed-test condition, we cleared context with a large distractor presentation (to simulate the full day off between learning list B and being tested on list A) and we activated the new day 3 unit prior to recalling the list A items.

We fit the model to the overall recall levels reported in the reminder/no-reminder, list A/list B, and immediate/delayed test conditions from Hubbach et al. (2007) by simulating 2000 lists for each condition (there were only 6 conditions, not 8, because the Hubbach et al. immediate-test condition only tested memory for list A items). Note that these were all between-participant manipulations and reported as three separate experiments by Hubbach et al., yet we fit them simultaneously. First, we ran a differential evolution genetic algorithm multiple times with different random starting populations in the free parameter ranges listed in Table 2 (Storn & Price, 1997). These all converged to the same approximate parameter range, indicating that we were in a stable parameter space and not stuck in a local minima. We then made minor adjustments to the parameters by hand to achieve better fits. The final parameters used to generate the results are included in Table 2.

Figure 1 shows the behavioral and simulation results for the Hubbach et al. (2007) data. We achieved a close quantitative fit with an RMSD of 1.71 in units of percent recalled across the 12 behavioral recall values. Importantly, the model was able to capture the asymmetric pattern of intrusions, whereby list B items were frequently intruded into list A recall in the reminder condition (left panel), but not vice-versa (right panel). In addition to the intrusion asymmetry, the model was able to capture the lack of intrusions in the immediate test condition, as well as the reduced rate of correct list A recall in the immediate test condition (center panel).

As discussed in the Introduction, TCM's ability to capture this complex pattern of intrusions is a direct result of contextual reinstatement and item-context binding: In the reminder condition, the list A context unit is associated with both list A and list B items, whereas the list B context unit is only ever bound to the list B items. Consequently, cuing with the list A context unit triggers recall of both list A and list B items, whereas cuing with the list B context unit primarily triggers recall of list B items. When list A is tested on the same day that list B is learned, however, the model is able to avoid list B intrusions using the recall-to-reject rule described earlier (whereby items are rejected if retrieved context matches the current test context). The overall decrease in correct list A recall on the immediate test occurs because list B items keep getting in the way — since list B items were studied more recently (relative to list A items) in the immediate-test vs. the delayed-test condition, list B items are retrieved more frequently in the immediate-test condition; the recall-to-reject rule prevents the model from intruding these items, but they take a toll in terms of reduced opportunities to recall list A items (every rejected list B item is a missed opportunity to recall a list A item).

We should note that the actual retrieval dynamics in the model are more nuanced than the simple account given above. During list B recall, when the model retrieves list B items, it also retrieves (via the item-to-context association matrix) the context that was present when the list B items were studied. In the reminder condition, this retrieved contextual information includes features of both list A and list B, which are then used to cue for more items. Naively, one might expect that including list A features (along with list B features) in the cue might lead to elevated intrusions of list A items, but this does not occur — there is an increase in list A intrusions in the reminder condition, both in the model and the data, but it is very slight. The reason that more list A intrusions do not occur is because recall in TCM is a competitive process: When list A and list B features are conjointly used as a retrieval cue, this cue has some degree of contextual match with list A items (since these items were bound to the list A context at study), but it has a higher degree of match with list B items (since these items were bound to both the list A and list B contexts at study). Because the level of contextual match is higher for list B items, these items are recalled by the accumulator instead of the list A items. Figure 2 provides a schematic summary of why the model predicts an asymmetric pattern of intrusions in the reminder condition.

In addition to simulating the results from Hupbach et al. (2007), we also simulated the results from a follow-up study conducted by Hupbach et al. (2009). The procedure used in the 2009 study was identical to the procedure used in the delayed-test condition of the 2007 study, except the final memory test was different: Instead of doing free recall of list A or list B items, participants were shown items one at a time. For each item, participants had to judge whether or not the item had been studied, and (if studied) whether it had been studied on list A or list B. Hupbach et al. found that source errors were rare overall in the no-reminder condition. In the reminder condition, Hupbach et al. observed an asymmetric pattern of source errors, where list B items were frequently mislabeled as list A items (but list A items were rarely mislabeled as list B items). Importantly, this asymmetric pattern of results was observed regardless of whether participants used the blue basket to learn list A and the table to learn list B (as in the 2007 paper) or whether they used the table to learn list A and the blue basket to learn list B – this rules out explanations of the intrusion asymmetry based on encoding task differences. Overall, these results fit with our hypothesis that, in the reminder condition, list B items are getting bound to the list A context, but not vice-versa.

To simulate the Hupbach et al. (2009) results, we used the exact same learning procedure that we used in the Hupbach et al. (2007) simulation. At test, however, instead of cuing with a list unit (list A or list B) and then freely recalling items, we cued with each item and tested the model's ability to retrieve the corresponding list unit (A or B) with a time-limit of 2 seconds (200 accumulator cycles) per item. The only parameters that we changed from the previous simulations were  $\beta_{cue}$ , which determines the degree of contextual drift triggered by retrieval cues at test, and also the parameters controlling the accumulator decision rule ( $\eta$ ,  $\kappa$ , and  $\lambda$ ). These changes were needed to accommodate the very different types of retrieval cues used in the 2009 study (specific item cues vs. a generalized list cue) and also the very different nature of the decision being made (here, the memory test is a two-choice decision between list A and list B, whereas – in our previous simulations – recall in the accumulator involved a competition between all 40 studied items). The final parameters for the source memory simulation were:  $\beta_{cue} = 0.8$ ,  $\eta = 15.0$ ,  $\kappa = .03$ , and  $\lambda = .05$ ; all other parameters were identical to the previous simulation of recall performance.

Again, the simulations achieve accurate quantitative fits to the behavioral data with an RMSD of 3.49. As shown in Figure 3, the model successfully captures the asymmetry whereby the reminder causes a significant increase in source errors for list B items but not list A items.

## Discussion

In the simulations presented here, we demonstrated that data from Hupbach et al. that have (previously) been explained in terms of reconsolidation can be explained using an existing computational model of memory (TCM). In our simulations, we achieved accurate quantitative fits to four separate experiments (List A delayed free-recall data from Hupbach et al., 2007, Experiment 1; List A immediate free-recall data from Hupbach et al., 2007, Experiment 2; List B delayed free-recall data from Hupbach et al., 2007, Experiment 3; and source recognition memory data from Hupbach et al., 2009, Experiment 1) with the same set of parameters governing the learning process in the model. The parameters governing retrieval were also the same across our simulations, except for some minor changes to accommodate the use of free recall in the 2007 study vs. source-memory cued-recall in the 2009 study. The key finding from our simulations was that the asymmetric pattern of intrusions observed by Hupbach et al. arises as a natural consequence of contextual reinstatement and item-context binding in the temporal context model (TCM): When the list A reminder is presented, this leads to reinstatement of contextual information related to list A; any list B items that are subsequently presented are linked to this reinstated context,

thereby increasing the odds that they will be retrieved when participants cue with the list A context at test. We also demonstrated one possible way that participants can avoid list B intrusions during the immediate-test condition: They can reject list B items if the context retrieved by the item matches the current test context.<sup>1</sup>

At the beginning of the paper, we posed the question of whether reconsolidation theory and models like TCM are compatible or incompatible at the cognitive level. Our simulation results support the former view: While these two frameworks describe memory in very different ways – reconsolidation papers talk about updating the molecular substrate of the memory whereas TCM focuses on contextual reinstatement and formation of new item-context associations – the basic behavioral predictions that come out of the two frameworks are quite similar. Based on our simulation results, we think that it is quite plausible (at least in the list-learning domain) to think of reconsolidation and TCM-style contextual reinstatement as descriptions of the same learning phenomenon at two different levels of analysis.

We should note that TCM is not the only computational model of episodic memory that could, in principle, account for the Hupbach et al. results shown in Figure 1 and Figure 3. The main prerequisites for explaining asymmetric intrusions are a representation of context and item-context binding, and several other models possess these properties. For example, the Context Maintenance and Retrieval model (CMR, Polyn et al., 2009) is based on TCM and thus it would probably be able to generate similar quantitative fits to the Hupbach et al. (2007, 2009) data. The Search of Associative Memory model (SAM, Raaijmakers & Shiffrin, 1981; Mensink & Raaijmakers, 1988) handles contextual drift differently from TCM and it uses a different retrieval rule (i.e., not an accumulator); however, it incorporates item-context binding and list-specific contextual information and – as such – it may be possible to use SAM to explain the Hupbach et al. (2007, 2009) findings.

### Sensitivity to parameters

In this section, we highlight how key model predictions depend (or do not depend) on model parameters.

**Intrusion asymmetry**—The asymmetric pattern of intrusions shown in Figure 1 (whereby, in the reminder condition, list B intrusions during list A are much more prevalent than list A intrusions during list B) depends on a combination of two factors. The first factor is the asymmetric nature of contextual reinstatement in the model (whereby the list A context is active during list A and list B study, but the list B context is only active during list B study); this is a parameter-independent property of the model. The second factor is our assumption that, on the final test, participants can contextually target list A or list B by selectively activating the context unit corresponding to that list, regardless of the contextual information that is actually present in the environment. In our simulations, we modeled list B recall in the reminder condition by saying that the list B unit (in the item layer) was 100% active, whereas the list A unit in the item layer was 0% active (even though participants were tested in the room where they had learned list A). Effectively, we are assuming that — when asked to recall list B — participants can focus on information that they think is diagnostic (e.g., the table-based encoding task, which was used in list B but not list A) and they can ignore non-diagnostic information (e.g., the room, which was the site of both list A and list B study). If we relax this contextual focusing assumption and activate the list A

<sup>1</sup>Another strategy that would have almost the same effect would be to assess the familiarity of recalled items and reject highly familiar items. If we assume that familiarity decays quickly, then list B items will be much more familiar than list A items on the immediate test, making it possible to discriminate between the two lists based on familiarity.

item-layer unit to a similar extent during list A and list B recall, this leads to an increase in the level of list A intrusions during list B recall, but crucially it does not cause a reversal of the pattern of intrusions (list A intrusions during list B recall are still less-than-or-equal-to list B intrusions during list A recall). The only way to generate a reversal in the pattern of intrusions is to massively reduce the learning rate  $\alpha$  during the study of list B items – this leads to a reduction in list B intrusions during list A recall, but it also has the side-effect of reducing correct recall of list B items during list B recall; for additional discussion of the effects of learning-rate manipulations, see the *Novel predictions* section below. The bottom line is that there is no way for the model in its present form to generate results where levels of correct recall are similar to those observed in the actual data, but intrusions are flipped (such that the rate of list A intrusions during list B recall is greater than the rate of list B intrusions during list A recall).

**More intrusions in the reminder condition**—The model’s prediction of greater intrusions in the reminder condition (vs. the no-reminder condition) is dependent on  $\nu_{A(Rem)} > \nu_{A(NoRem)}$ . This parameter configuration instantiates the common-sense idea that the list B study context is more similar to the list A study context in the reminder condition vs. the no-reminder condition.

**Contextual reinstatement**—The  $\gamma_{FT}$  parameter controls the extent to which experimental context (i.e., associated contextual information from earlier in the experiment) is reinstated in the context layer when an item is retrieved. The simulations shown here were run with  $\gamma_{FT}$  set to 0.25. Strictly speaking, we could fit the results shown in Figure 1 and Figure 3 with  $\gamma_{FT} = 0$  (i.e., no automatic reinstatement of contextual information from earlier in the experiment). We decided to use a positive  $\gamma_{FT}$  value because of other TCM modeling work showing that a positive  $\gamma_{FT}$  value is needed to explain contiguity effects in free recall (Howard & Kahana, 2002; Sederberg et al., 2008). We discuss some predictions deriving from our use of positive  $\gamma_{FT}$  in the *Novel predictions* section below.

### Relation to other findings

A number of related memory distortion findings, including the misinformation and hindsight bias effects, have been cited as support for reconsolidation (Hardt et al., 2010). For example, in studies of the misinformation effect (E. Loftus, Miller, & Burns, 1978; E. F. Loftus, 2005), participants are given misleading information during the course of answering questions about a recently-viewed event. On a later memory test, participants incorrectly intrude the false details when recounting the original event. According to reconsolidation accounts, these intrusions occur because the post-event questions serve as a reminder, causing the original event to be retrieved and become labile; the relabialized memory is then modified based on the misleading information in the questions (Hardt et al., 2010). Our TCM simulations suggest a reframing of this account in terms of contextual reinstatement and item-context binding: According to this view, the post-event questions reinstate the context of the original event, and participants bind that context to the new (misleading) information; on the final test, when participants cue with the original study context, this triggers retrieval of the misleading details. The same general principles can also be used to explain the hindsight bias effect (Hawkins & Hastie, 1990), whereby participants’ memory for their initial response to a question is biased by subsequently presented information.

The TCM account of retroactive interference phenomena like the misinformation effect and the hindsight bias effect is highly compatible with the account provided by the Johnson, Hashtroudi, and Lindsay (1993) source monitoring framework (SMF) – contrary to claims made by E. Loftus and Loftus (1980) that the original memory is modified or possibly overwritten, both TCM and SMF posit that the original memory is left intact, and that errors

are caused by participants retrieving post-event information and attributing it to the wrong source (Lindsay & Johnson, 1989). A key prediction generated by the TCM/SMF interpretation is that, since the original memory is never modified, it should be possible to access the original details of the memory given a sufficiently specific cue. In keeping with this view, numerous studies of the misinformation effect have found that it is possible to retrieve the original details when participants are given specific cues (e.g., on a forced-choice recognition test; McCloskey & Zaragoza, 1985). We should emphasize that, although TCM and SMF provide highly compatible accounts of these effects, TCM and SMF are very different kinds of theories: SMF is a verbally defined theory whereas TCM is a computational model that can be used to make concrete predictions about contextual reinstatement and its consequences for subsequent memory performance; this ability to generate concrete predictions about contextual reinstatement allows TCM to predict phenomena like asymmetric intrusions that do not naturally fall out of verbally-defined theories like SMF.

For completeness, we should also note that we have focused (in our simulations, and in our discussion of misinformation and hindsight effects) on source memory errors that are driven by item-context associations that were formed at study – list B items are linked to the list A context, so cuing with the list A context at test results in retrieval of some list B items. Importantly, this is not the only possible mechanism that can result in source memory errors. For example, on a forced-choice source test (of the sort used by Hupbach et al., 2009), participants might fail to retrieve any diagnostic source information for a given item and then make a random guess about the source. This kind of guessing mechanism could result in bidirectional source memory errors (the model will sometimes attribute list A items to list B, and they will sometimes attribute list B items to list A). While it was not necessary to incorporate this kind of guessing mechanism into the model to explain the Hupbach results, it might be necessary to incorporate this kind of random guessing into the model to account for bidirectional patterns of source memory errors that have been observed in the literature (e.g., Hintzman et al., 1998).

### Novel predictions

A key benefit of applying computational models to the Hupbach et al. data is that we can use these models to generate novel predictions. For example, we can use TCM to explore how the intrusion effect that Hupbach et al. observed in the reminder condition would interact with serial position effects. The Hupbach et al. (2007) data are not suitable for looking at serial position effects because the list B items were presented all at once. Nonetheless, we can use the model to explore what would have happened if the list B items had been presented one at a time. The model's predictions are shown in Figure 4: According to TCM, reinstated contextual information should be most strongly active right after the reminder and then fade over time; as such, the probability of intruding a list B item (when participants are asked to recall list A items) should be highest for items at the beginning of list B (when reinstated list A information is most strongly active) and lower for items at the end of the list (when reinstated information is less active). Note that, while the reminder effect is predicted to be strongest for items at the beginning of the list, the model predicts an elevated level of intrusions for all of the items in list B. The model makes this prediction because – in the reminder condition – information relating to list A (in particular, the experimenter and the room) is present in the environment throughout the list B study period. This persistently-available list A information gets bound to all of the list B items, thereby increasing the odds that they will be intruded later.

Another prediction from TCM is that, when participants are trying to retrieve memories, presenting a previously studied item should automatically trigger reinstatement of associated contextual information from earlier in the experiment (as noted above, this is a consequence

of  $\gamma_{FT}$  being set to a positive value). Consider a variant of the Hupbach paradigm where there is no overlap in environmental context between list A and list B (i.e., the rooms are different; the experimenters are different) but, instead, some of the items from list A are also included in list B. Furthermore, assume that (during list B study) participants are given the extra demand of pressing a button whenever they recognize an item from list A. In this situation, TCM predicts that repetition of list A items (during list B study) will have the same effect as the explicit “reminder” in the Hupbach et al. (2007, 2009) studies. Specifically, the repeated items will trigger reinstatement of the list A context (assuming that participants successfully notice that the item was presented during list A). This reinstated “list A” contextual information will then be linked to subsequently presented list B items, making it more likely that these list B items will be intruded during list A recall. As with the situation depicted in Figure 4, the reinstated “list A” contextual information should be strongest right after a list A item is presented, and then gradually dissipate due to contextual drift. As such, the intrusion effect should be strongest for list B items studied just after the re-presented list A items.

We can also make predictions regarding the effects of learning rate manipulations. In the simulations described in the *Results* section, the learning rate  $\alpha$  was fixed at 1.0. Here, we discuss the effects of changing the learning rate on correct recall and intrusions. Making global adjustments in the learning rate affects the total number of recalls (higher learning rate = more items recalled) but – up to the point where we start to see edge effects (recalling very few items or recalling all items) – global adjustments to the learning rate do not affect the ratio of correct to intruded items. However, as shown in Figure 5, varying the learning rates for list A and list B independently of one other can have a striking effect on the ratio of correct recalls to intrusions. If the learning rate for list A is lowered to 0.5 and the list B learning rate is held at 1.0 ( $A < B$ ), TCM predicts more intrusions of list B items (during list A recall) due to better binding of the list B items to the list A context. The opposite pattern of learning rates ( $A > B$ ) gives rise to fewer intrusions of list B items; in this situation, list A items win the competition against the list B items, which were only weakly bound to the list A context (as noted in the *Sensitivity to parameters* section, lowering the list B learning rate also has the added cost of reducing correct recalls for list B items – there is no way to eliminate list B intrusions while also maintaining the level of correct list B recall that was observed in the actual data). These predictions about changes in the intrusion rate can be tested with focused vs. divided attention or rote vs. elaborative encoding manipulations, where list A is encoded using a relatively effective encoding task and list B is encoded using a relatively ineffective encoding task (or vice-versa).

Lastly, given that the retrieval process in TCM-A is based on the Usher and Mc-Clelland (2001) leaky-accumulator model, we can make predictions about response times in addition to recall performance. As stated earlier, our explanation of the immediate-test data from Hupbach et al. (2007) is that list B items come to mind more strongly in the immediate-test condition (because of recency) but participants are able to reject them using a recall-to-reject decision rule based on contextual matching. This leads to the prediction that, at least at the start of recall, participants will be slower (on average) to recall list A items in the immediate-test condition (vs. the delayed-test condition) because of the extra time spent rejecting list B items. In the model, the mean number of time-steps to the first response in immediate test ( $M = 672.99$ ,  $SEM = 19.44$ ) was significantly greater than in the delayed test ( $M = 465.84$ ,  $SEM = 8.08$ ) by an independent samples t-test ( $t(3843) = 10.12$ ,  $p < .0001$ ).<sup>2</sup>

<sup>2</sup>We assume that time steps in the model map linearly onto msec, but we have not attempted to fit a scaling factor here.

## Limitations of the model

**TCM does not explain the differential efficacy of spatial reminders**—In another study (not simulated above), Hupbach, Hardt, Gomez, and Nadel (2008) explored what makes an effective reminder. The reminder in the Hupbach et al. (2007) study was composed of three parts: using the same experimenter, using the same room, and asking participants to think back to the original study event. The goal of the 2008 study was to explore the relative efficacy of these three components in triggering the asymmetric intrusion effect. The results from Hupbach et al. (2008) clearly showed that the asymmetric intrusion effect was being driven by the use of the same room as a reminder; this component of the reminder was effective when used on its own, and the other two components were ineffective on their own (and in combination with one another). While these results do not directly contradict any of the principles in TCM, it is equally clear that this result is not predicted by TCM — that is, there is nothing in the basic form of TCM that would predict that reinstatement of spatial context would be especially effective in triggering mental context reinstatement (compared, e.g., to reinstating the experimenter who was present during the original event). Understanding why some types of reminders are more effective than others is an important direction for future research.

**TCM does not model strategic memory targeting processes**—In the *Sensitivity to Parameters* section, we discussed the idea that – at test – participants can strategically weight retrieval cues at test according to the perceived relevance and diagnosticity of these cues (e.g., when trying to recall list B, participants can strategically cue with the table-based encoding task, which was used in list B but not list A). This principle has strong face validity – if you are trying to recall a trip to the beach while sitting in your office, you can ignore your immediate spatial environment and focus on contextual features that are unique to beaches. In the present version of our model, we simply assume that participants can do strategic cue weighting. In future work, we plan to develop a version of TCM that directly addresses how participants assign weights to retrieval cues.

**Other types of reconsolidation**—In this paper, we have focused on addressing a particular kind of reconsolidation data (reconsolidation in list-learning). As reviewed in the Introduction, reconsolidation can occur in numerous domains of learning and memory outside of list-learning (e.g., fear conditioning). More work is needed to assess whether other types of reconsolidation can be explained using our TCM-based model.

## Conclusions

In this paper, we applied a well-established computational model of recall (TCM) to data from Hupbach et al. (2007, 2009) that have been described as evidence for reconsolidation in humans. In doing this, we hoped to determine whether the Hupbach results were consistent with extant cognitive models of recall, or whether these results require a fundamental re-thinking of how we model recall and forgetting. Our simulation results support the former alternative: The distinctive asymmetric pattern of intrusions obtained by Hupbach et al. (2007, 2009) can be explained in a principled way by TCM, using widely-accepted ideas about item-context binding and contextual reinstatement.

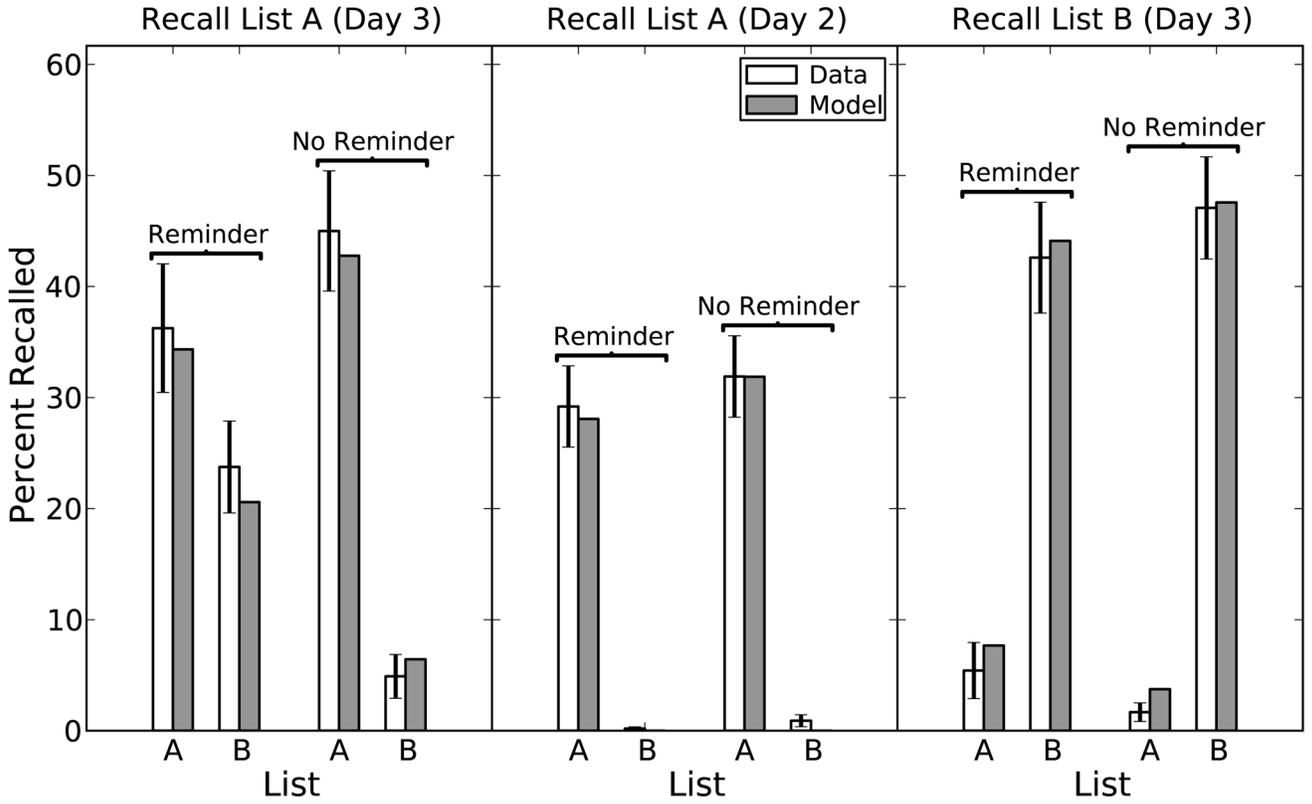
## Acknowledgments

The authors acknowledge support from the National Institutes of Health research grants MH080526 awarded to PBS and MH069456 awarded to KAN.

## References

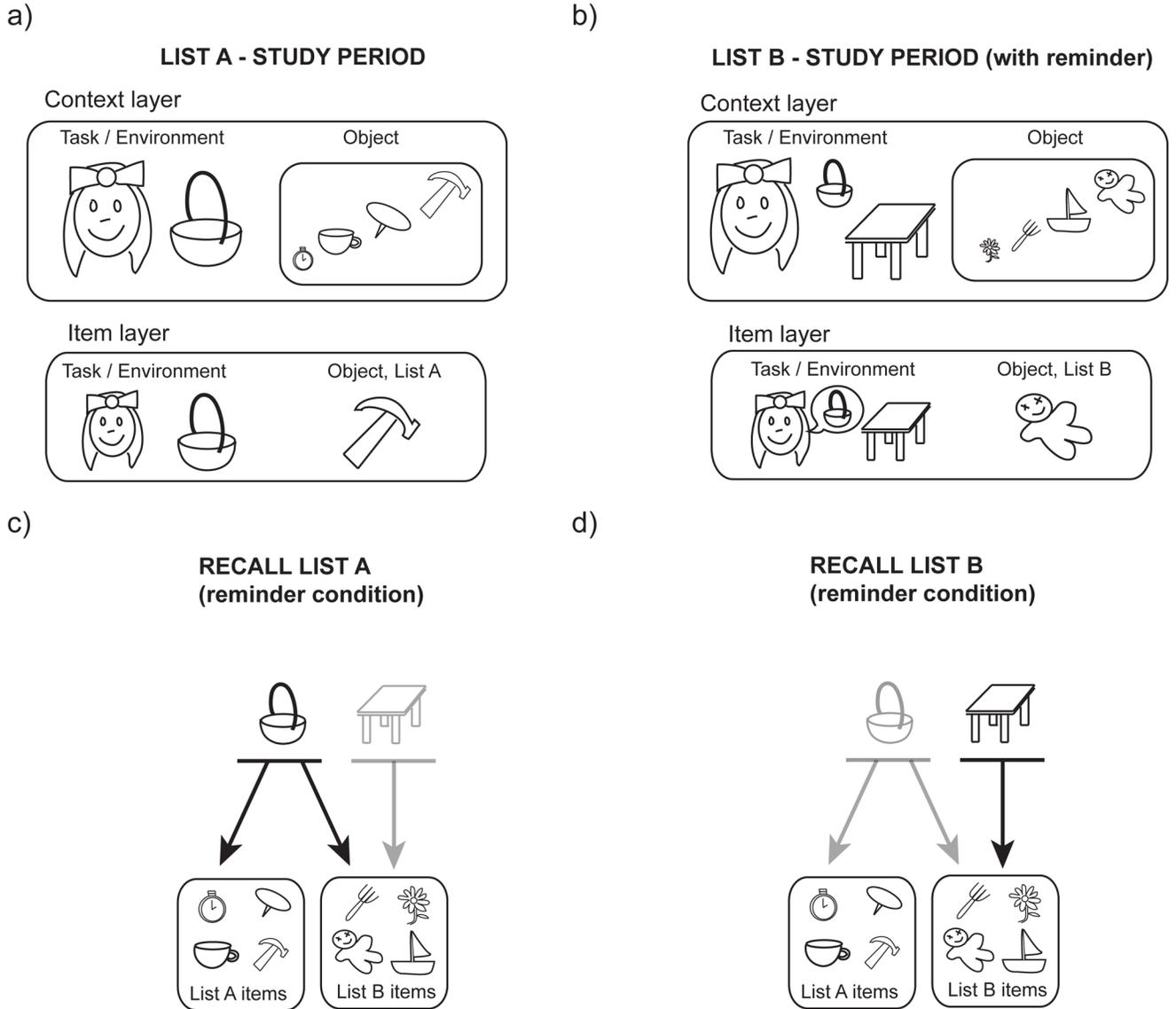
- Dudai Y. Predicting not to predict too much: how the cellular machinery of memory anticipates the uncertain future. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*. 2009; 364(1521):1255–62.
- Dudai Y, Eisenberg M. Rites of passage of the engram: reconsolidation and the lingering consolidation hypothesis. *Neuron*. 2004; 44(1):93–100. [PubMed: 15450162]
- Eisenberg M, Kobilo T, Berman DE, Dudai Y. Stability of retrieved memory: inverse correlation with trace dominance. *Science (New York, NY)*. 2003; 301(5636):1102–4.
- Hardt O, Einarsson EO, Nader K. A bridge over troubled water: reconsolidation as a link between cognitive and neuroscientific memory research traditions. *Annual Review of Psychology*. 2010; 61:141–67.
- Hawkins S, Hastie R. Hindsight: Biased judgments of past events after the outcomes are known. *Psychological Bulletin*. 1990; 107(3):311–327.
- Hintzman DL, Caulton DA, Levitin DJ. Retrieval dynamics in recognition and list discrimination: Further evidence of separate processes of familiarity and recall. *Memory and Cognition*. 1998; 26:449–462.
- Howard MW, Jing B, Rao V, Probyn J, Datey A. Bridging the gap: Transitive associations between items presented in similar temporal contexts. *Journal of Experimental Psychology: Learning, Memory, and Cognition*. 2009; 35(2):391–407.
- Howard MW, Kahana MJ. A Distributed Representation of Temporal Context. *Journal of Mathematical Psychology*. 2002; 46:269–299.
- Hupbach A, Gomez R, Hardt O, Nadel L. Reconsolidation of episodic memories: A subtle reminder triggers integration of new information. *Learning & Memory*. 2007; 14(1–2):47. [PubMed: 17202429]
- Hupbach A, Gomez R, Nadel L. Episodic memory reconsolidation: updating or source confusion? *Memory*. 2009; 17(5):502–510. [PubMed: 19468955]
- Hupbach A, Hardt O, Gomez R, Nadel L. The dynamics of memory: Context-dependent updating. *Learning & Memory*. 2008; 15(8):574. [PubMed: 18685148]
- Johnson MK, Hashtroudi S, Lindsay DS. Source monitoring. *Psychological Bulletin*. 1993; 114:3–28. [PubMed: 8346328]
- Lee JLC. Reconsolidation: maintaining memory relevance. *Trends in Neurosciences*. 2009; 32(8):413–20. [PubMed: 19640595]
- Lindsay DS, Johnson MK. The eyewitness suggestibility effect and memory for source. *Memory and Cognition*. 1989; 17:349–358.
- Loftus E, Loftus GR. On the permanence of stored information in the human brain. *American Psychologist*. 1980; 35(5):409–420. [PubMed: 7386971]
- Loftus E, Miller D, Burns H. Semantic integration of verbal information into a visual memory. *Journal of Experimental Psychology*. 1978; 4(1):19–31.
- Loftus EF. Planting misinformation in the human mind: a 30-year investigation of the malleability of memory. *Learning & Memory (Cold Spring Harbor, NY)*. 2005; 12(4):361–6.
- McCloskey M, Zaragoza M. Misleading Postevent Information and Memory for Events: Arguments and Evidence Against Memory Impairment Hypotheses. *Journal of Experimental Psychology: General*. 1985; 114(1):1–16. [PubMed: 3156942]
- Mensink G, Raaijmakers J. A model for interference and forgetting. *Psychological Review*. 1988; 95:434–455.
- Morris RGM, Inglis J, Ainge JA, Olverman HJ, Tulloch J, Dudai Y, et al. Memory reconsolidation: sensitivity of spatial memory to inhibition of protein synthesis in dorsal hippocampus during encoding and retrieval. *Neuron*. 2006; 50(3):479–89. [PubMed: 16675401]
- Nader K, Schafe G, Le Doux J. Fear memories require protein synthesis in the amygdala for reconsolidation after retrieval. *Nature*. 2000; 406(6797):722–726. [PubMed: 10963596]
- Norman, KA.; Detre, GJ.; Polyn, SM. Computational models of episodic memory. In: Sun, R., editor. *The Cambridge handbook of computational psychology*. New York: Cambridge University Press; 2008.

- Pedreira ME, Pérez-Cuesta LM, Maldonado H. Mismatch between what is expected and what actually occurs triggers memory reconsolidation or extinction. *Learning & Memory (Cold Spring Harbor, NY)*. 2004; 11(5):579–85.
- Polyn SM, Norman KA, Kahana MJ. A context maintenance and retrieval model of organizational processes in free recall. *Psychological Review*. 2009; 116:129–156. [PubMed: 19159151]
- Raaijmakers, J. Modeling implicit and explicit memory. In: Izawa, C.; Ohta, N., editors. *Human learning and memory: Advances in theory and application*. Mahwah, NJ: Erlbaum; 2005. p. 85-105.
- Raaijmakers J, Shiffrin R. Search of associative memory. *Psychological Review*. 1981; 88(2):93–134.
- Raaijmakers, J.; Shiffrin, RM. Models of memory. In: Pashler, H.; Medin, D., editors. *Stevens' handbook of experimental psychology*. 3. Vol. 2. New York: John Wiley and Sons; 2002. p. 43-76. Memory and cognitive processes
- Riccio DC, Millin PM, Bogart AR. Reconsolidation: a brief history, a retrieval view, and some recent issues. *Learning & Memory (Cold Spring Harbor, NY)*. 2006; 13(5):536–44.
- Schiller D, Monfils M-H, Raio CM, Johnson DC, LeDoux JE, Phelps EA. Preventing the return of fear in humans using reconsolidation update mechanisms. *Nature*. 2009; 463(7277):49–53. [PubMed: 20010606]
- Sederberg PB, Howard MW, Kahana MJ. A context-based theory of recency and contiguity in free recall. *Psychological Review*. 2008; 115(4):893–912. [PubMed: 18954208]
- Storn R, Price K. Differential evolution—a simple and efficient heuristic for global optimization over continuous spaces. *Journal of Global Optimization*. 1997; 11(4):341–359.
- Usher M, McClelland JL. The time course of perceptual choice: The leaky, competing accumulator model. *Psychological Review*. 2001; 108(3):550–592. [PubMed: 11488378]
- Walker MP, Brakefield T, Hobson J, Stickgold R. Dissociable stages of human memory consolidation and reconsolidation. *Nature*. 2003; 425(6958):616–620. [PubMed: 14534587]
- Wang SH, Morris RGM. Hippocampal-neocortical interactions in memory formation, consolidation, and reconsolidation. *Annual Review of Psychology*. 2010; 61:49–79.



**Figure 1. Hupbach et al. (2007) simulation**

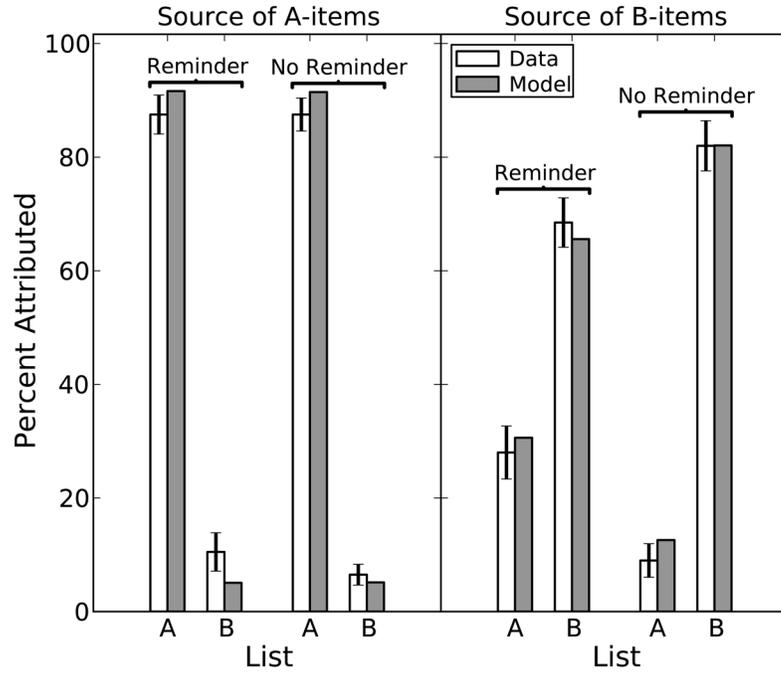
The left panel shows the percent recalled from each list when participants were asked to recall list A items on day 3 (the delayed test condition) in both the reminder and the no-reminder conditions. In this panel, any item recalled from list B is an intrusion. The middle panel shows the percent recalled from each list when participants were asked to recall list A immediately after learning list B on day 2. The right panel plots the percent recalled from each list when participants were asked to recall list B items (in both the reminder and the no-reminder conditions) on day 3. In this panel, any item recalled from list A is an intrusion. For all panels, the behavioral data are shown in white and the model fits are shown in gray. Error bars are standard errors of the mean.



**Figure 2. Schematic explanation of why the model shows an asymmetric intrusion effect in the reminder condition**

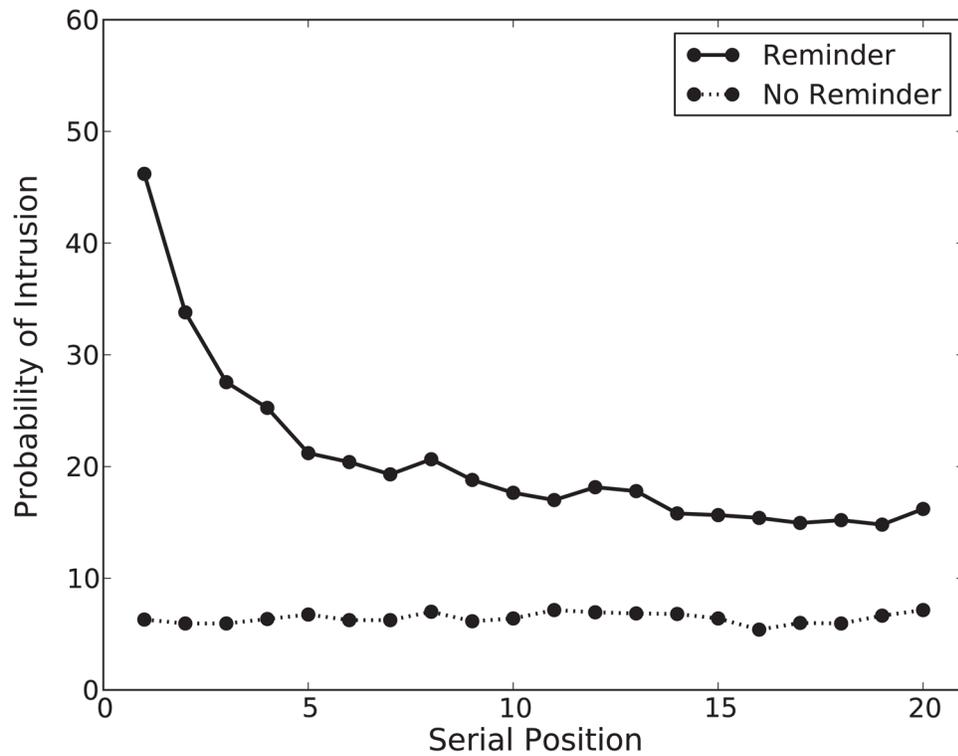
a) Snapshot of model activity during list A study. The item layer represents the object that is currently being studied (the hammer), currently-active task/environment features (e.g., the experimenter, the basket used for the encoding task, and the room; the room is not shown here), and features of the current day (not shown here). The context layer represents a recency-weighted average of item-layer activity; this includes the currently-active object, task/environment, and day-specific features as well as features relating to recently presented objects (tack, cup, stopwatch). b) Snapshot of model activity during list B study (reminder condition). The context layer contains features relating to unique aspects of list B (the table-based encoding task), as well as list A features that were triggered by the reminder (the experimenter and the basket). During list B study, these list A context features are episodically bound to list B objects. c) To cue recall of objects from list A, we activate features of the list A task/environment (e.g., the basket; the table is greyed out because we hypothesize that participants focus on features that are specifically relevant to the to-be-

recalled list and they ignore other features). Since these list A contextual features were linked to both list A and list B objects at study, both list A and list B objects are recalled. d) To cue recall of objects from list B, we activate unique features of the list B task/environment (here the table is highlighted and the basket is greyed out, because the table-based encoding task is specifically relevant to list B). Since the active set of contextual features provides more support to list B objects than list A objects, the model tends to recall list B objects more than list A objects — list A intrusions are relatively rare.



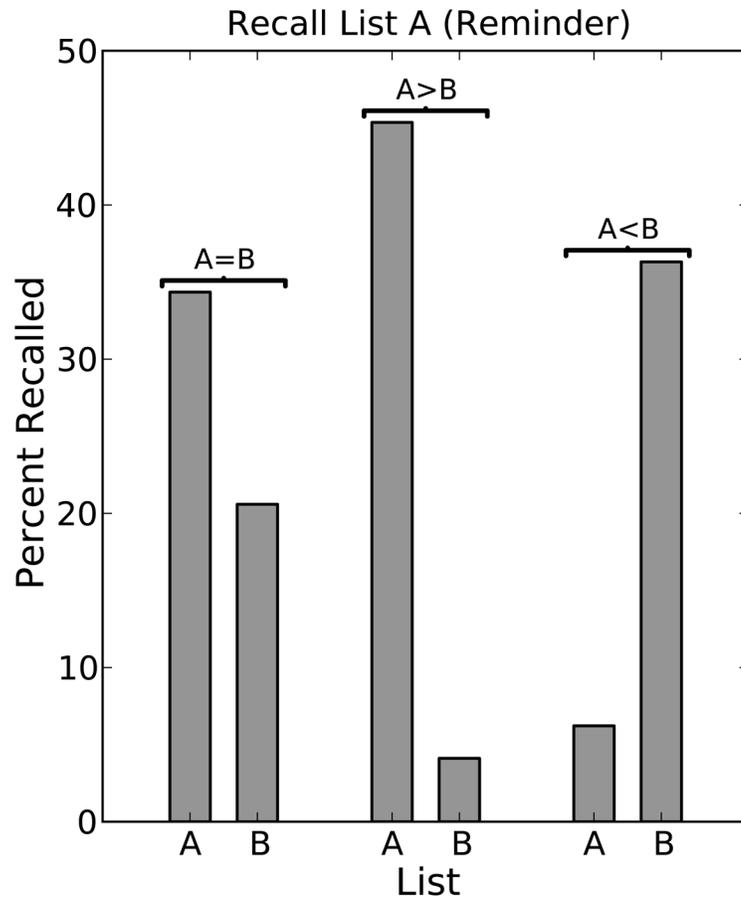
**Figure 3. Hupbach et al. (2009) simulation**

The left panel depicts the percent of items from list A attributed to list A or B in the reminder and no-reminder conditions. The right panel plots the percent attributed to each list for the B items. For both panels, the behavioral data are shown in white and the model fits are shown in gray. Error bars are standard errors of the mean.



**Figure 4. Intrusions as a function of serial position**

TCM's predictions regarding the probability of intruding a list B item into list A recall, as a function of the list B item's serial position and the presence/absence of a reminder.



**Figure 5. Intrusions as a function of list learning rate**

TCM's predictions regarding the probability of intruding a list B item into list A recall in the reminder condition on day 3 (delayed test), as a function of the relative learning rates for list A and list B. The  $A = B$  condition (provided for reference) is identical to the condition simulated in Figure 1 (left panel); in this condition, the learning rate was set to 1.0 for both lists. The  $A < B$  condition used a learning rate of 0.5 for list A and a learning rate of 1.0 for list B. The  $A > B$  condition used a learning rate of 1.0 for list A and a learning rate of 0.5 for list B. For all three conditions, list A recalls are correct responses and list B recalls are intrusions.

**Table 1**  
**Summary of Hupbach et al. (2007) simulation**

Side-by-side comparison of stages of the Hupbach et al. (2007) study and the corresponding stages in our simulation. The description of the simulation highlights the roles of the novel TCM parameters listed in Table 2.

	Hupbach et al. (2007)	Simulation
1	Unrelated experiences prior to experiment	Initialize context with an orthogonal distractor
2	Day 1 of experiment	Turn on the item-layer unit representing day 1 and drift context by $\beta_{dist}$
3	Enter room A with experimenter A	Turn on the item-layer unit representing list A and drift context by $\beta_{init}$
4	Learn 20 objects with task A	Present 20 items, one at a time, at strength $\nu_{word}$ while keeping the item-layer unit representing list A active at $\nu_A$ and the day 1 unit active at $\nu_{day}$ ; for each item, drift context by $\beta$
5	Day off filled with unrelated activity	Turn on a new distractor unit in the item layer and drift context by $\beta_{dist}$
6	Day 2 of experiment	Turn on the item-layer unit representing day 2 and drift context by $\beta_{dist}$
7a	<i>Reminder</i> : Discuss task A in room A with experimenter A, then commence study of list B	Turn on the item-layer unit representing list A with strength $\nu_{A(Rem)}$ and drift context by $\beta_{cue}$ ; next, turn on the item-layer unit representing list B at strength $\nu_{B(Rem)}$ (keeping the list A unit active at strength $\nu_{A(Rem)}$ ) and drift context by $\beta_{init}$
7b	<i>No Reminder</i> : Enter room B with experimenter B, then commence study of list B	Turn on the item-layer unit representing list A at strength $\nu_{A(No Rem)}$ and drift context by $\beta_{cue}$ ; next, turn on the item-layer unit representing list B at strength $\nu_{B(No Rem)}$ (keeping the list A unit active at strength $\nu_{A(No Rem)}$ ) and drift context by $\beta_{init}$
8	Learn 20 objects with task B	Present 20 items, one at a time; during this phase, keep the item-layer units representing lists A and B active at either their reminder levels ( $\nu_{A(Rem)}$ , $\nu_{B(Rem)}$ ) or no-reminder levels ( $\nu_{A(No Rem)}$ , $\nu_{B(No Rem)}$ ), depending on the condition, and keep the day 2 unit active at $\nu_{day}$ ; for each item, drift context by $\beta$
9a	<i>Immediate Test</i> : Stay on day 2 for immediate test	Turn on a new distractor unit in the item layer and drift context by $\beta_{distimmed}$
9b	<i>Delayed Test</i> : Day off filled with unrelated activity	Turn on a new distractor unit in the item layer and drift context by $\beta_{dist}$ , then turn on the item-layer unit representing day 3
10	Experimenter A asks participant to recall list A or B	Activate the item-layer unit representing list A or B as a recall cue, and drift context by $\beta_{cue}$

**Table 2**  
**Summary of free and fixed parameters in TCM-A**

Fixed parameters have only a single value listed in the Range column. The section labeled *Novel* lists new parameters added to simulate the Hupbach et al. (2007, 2009) experiments.

Category	Parameter	Description	Range	Value
TCM Base	$\beta$	Rate of contextual drift due to item encoding and retrieval.	0.0–1.0	0.65
	$\beta_{dist}$	Rate of contextual drift due to distractor presentation and the drift due to a new day.	0.99	0.99
	$\beta_{dist<sub>immed</sub>}$	Rate of contextual drift due to distractor presentation prior to the immediate recall on day 2.	0.0–1.0	0.75
	$\gamma_{FT}$	Relative weight of exp. to pre-exp. context (item-to-context).	0.0–1.0	0.25
	$\gamma_{TF}$	Relative weight of exp. to pre-exp. context (context-to-item).	1.0	1.0
	$\alpha$	Learning rate of associations between	0.0–1.0	1.0
Primacy	$\varphi_s$	Primacy scale factor.	0.0–5.0	0.5
	$\varphi_d$	Primacy decay rate.	0.0–5.0	1.0
Accumulator	$\kappa$	Strength of recurrent inhibition.	0.0–0.9	0.85
	$\lambda$	Strength of lateral inhibition	0.0–0.9	0.02
	$\eta$	Noise in the accumulator.	0.0–15.0	11.542
	$\theta$	Recall threshold.	1.0	1.0
	$\delta t$	Time step size.	10 ms	10 ms
	$\tau$	Rate of growth at each time step.	1000 ms	1000 ms
	Novel	$\beta_{mit}$	Rate of contextual drift due to entering a list context.	0.5–1.0
$\beta_{cue}$		Rate of contextual drift due to the reminder and retrieval cues.	0.0–1.0	0.5
$\nu_{word}$		Strength of studied items in the item layer.	0.5–1.0	0.75
$\nu_{day}$		Strength of the day units in the item layer.	0.0–0.5	0.01
$\nu_A$		Strength of the list A item-layer unit during list A study.	0.0–0.5	0.4
$\nu_{A(Rem)}$		Strength of the list A item-layer unit prior to and during list B study (reminder condition).	0.0–0.5	0.2
$\nu_{A(No Rem)}$		Strength of the list A item-layer unit prior to and during list B study (no reminder condition).	0.0–0.5	0.05
$\nu_{B(Rem)}$		Strength of the list B item-layer unit during list B study (reminder condition).	0.0–0.5	0.35
$\nu_{B(No Rem)}$		Strength of the list B item-layer unit during list B study (no reminder condition).	0.0–0.5	0.35