Cognitive Psychology 156 (2025) 101714

Contents lists available at ScienceDirect

Cognitive Psychology

journal homepage: www.elsevier.com/locate/cogpsych

A position coding model that accounts for the effects of event boundaries on temporal order memory

Xiaojing Peng [a], Yifei Cao [a], Jintao Sheng [a], Yu Zhou [a], Huinan Hu [a], Gui Xue [a,b,*]

[a] *State Key Laboratory of Cognitive Neuroscience and Learning & IDG/McGovern Institute for Brain Research, Beijing Normal University, China*
[b] *Chinese Institute for Brain Research, China*

ARTICLE INFO

Keywords:
Episodic memory
Event boundaries
Temporal order memory
Position coding

ABSTRACT

Episodic memories, particularly temporal order memory, are influenced by event boundaries. Although numerous theoretical and computational models have been developed to explain this phenomenon, creating a model that can explain a wide range of behavioral data and is supported by neural evidence remains a significant challenge. This study presented a new model, grounded in ample evidence of position coding, to account for the impact of event boundaries on temporal order memory. The proposed model successfully simulated various behavioral effects in previous experiments measuring temporal order memory. Our model outperformed the context-resetting model in fitting all the data and capturing the full set of effects in previous and newly conducted experiments, including the boundary effect, the distance effect, the local primacy effect, and the absence of boundary number effect. These findings underscore a novel mechanism in which event boundaries affect temporal order memory by resetting the local position coding of events.

1. Introduction

Our real-life experiences unfold continuously over time. However, our memories of events are grouped into meaningful units (Michelmann et al., 2023). Event boundaries, experienced as context shifts (e.g., spatial, goal, and perceptual changes), act as breakpoints and influence how ongoing episodes are processed and encoded into memory (Clewett et al., 2019). These influences manifest in different ways, including the enhancement of item-context binding for boundary items compared to non-boundary items (Heusser et al., 2018), the inflation of temporal distance estimates for items spanning a perceptual boundary (Ezzyat & Davachi, 2014), and the impairment of temporal order memory for cross-boundary items compared to within-boundary items (DuBrow & Davachi, 2013). For example, using a variety of experimental manipulations, such as stimulus categories (DuBrow & Davachi, 2013), videos spanning diverse topics (Zheng et al., 2022), different virtual rooms (Horner et al., 2016), and images with varying color frames (Heusser et al., 2018; Pu et al., 2022; Wen & Egner, 2022), extant studies have consistently reported the impairment effect of event boundaries on temporal order memory (for summaries, see Wen & Egner, 2022). Nevertheless, the cognitive and computational mechanisms underlying the effects of event boundaries on temporal order memory remain elusive.

Several theoretical and computational models have been developed to address the role of event boundaries in episodic memory in general and temporal order memory in particular. According to event segmentation theory, event boundaries are crucial for

* Corresponding author at: State Key Laboratory of Cognitive Neuroscience and Learning and IDG/McGovern Institute for Brain Research, Beijing Normal University, 19 Xinjiekou Wai Street, Beijing, 100875, China.

E-mail address: gxue@bnu.edu.cn (G. Xue).

https://doi.org/10.1016/j.cogpsych.2025.101714
Received 3 August 2023; Received in revised form 9 January 2025; Accepted 9 January 2025
Available online 15 January 2025

segmenting continuous experiences into discrete events, enabling better processing of sensory inputs and improved predictions about the future (Zacks et al., 2007). As the relationships are better coded for within-event items than cross-event ones (Radvansky & Zacks, 2017), the temporal order memory for items across event boundaries is poorer than for items within the same event. According to chaining mechanism, it has been suggested that sequential items are linked in a chain, and temporal order judgment depends on the reconstruction of this chain (Lewandowsky & Murdock, 1989). Event boundaries disrupt the associative chaining of items, resulting in worse performance in judging the order of items spanning event boundaries (DuBrow & Davachi, 2013).

The temporal context models posit that items are associated via temporally drifting context (Estes, 1955; Howard & Kahana, 2002). Contextual representation of items changes gradually within an event but shifts more rapidly during the transitions between events. This mechanism could explain the effect of event boundaries on temporal adjacency judgment (Horner et al., 2016). In Horner et al.'s (2016) study, participants navigated through several rooms in a virtual reality environment, where boundary experiences were manipulated as transitions between rooms. During the temporal memory task, participants were asked to choose which of the three items appeared temporally next to the cue item. Behavioral results showed better performance for items located within the same room as the cue item compared to those located in different rooms. The authors found a simple computational model could capture this effect by assuming that the item showing the highest contextual similarity to the cue item would be the one coming next.

It has recently been posited that this effect of event boundaries on temporal order memory can be explained by a temporal context resetting mechanism, which hypothesizes that the temporal context of boundary items partially resets to that of the first initial item in a sequence (Pu et al., 2022). In their study, participants were asked to remember a series of objects embedded in differently colored frames, where perceptual shifts of color created boundary experiences. They proposed that recency judgments could be based on the dissimilarity of contextual representations between probed items and the very first item in the sequence, with item exhibiting greater dissimilarity to the first item being judged as more recent. Due to temporal context reset, the context of post-boundary items would show higher similarity to that of the first item compared to pre-boundary items, resulting in worse memory performance for cross-boundary pairs than for within-boundary pairs. This model successfully accounts for several effects in temporal order judgment tasks, including the boundary effect, the distance effect (i.e., better performance for item pairs with longer temporal distances), and the

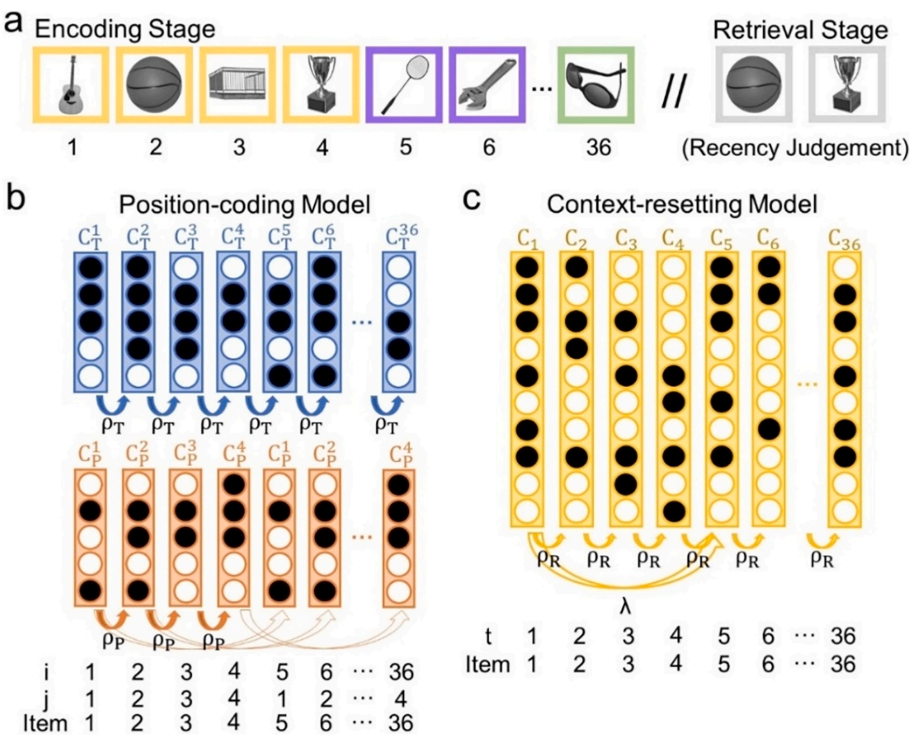


Fig. 1. Paradigm of recency judgment task and the models. **a**, The schematic plot of the recency judgment paradigm. The experiment included an encoding stage of learning 36 trial-unique objects embedded in colored frames and a retrieval stage of making recency judgments of just-encoded objects. Event boundaries were manipulated by introducing shifts in the frame colors. **b**, The schematic plot of the position-coding model. Each item (object picture) carried two kinds of contextual information: temporal and position codes. Filled and empty circles represented features 0 and 1, respectively. The parameters ρ_T (where "i" indexes the list position of items) and ρ_P (where "j" indexes the local position of items) regulated the probability that each element (i.e., circle) would remain the same. In the example shown, an event boundary occurred after every four items. **c**, The schematic plot of the context-resetting model. The context vector of each boundary item was partially reset to that of the first item in the sequence, controlled by the parameter λ. The proportion $(1 - \lambda)$ of the context for boundary items (i.e., non-resetting context), along with the context for non-boundary items, drifted gradually from the context for previous item. The probability of drift was controlled by the parameter ρ_R. The list position of items was indexed by "t" in the figure.

local primacy effect (i.e., better temporal order memory for item pairs located in early rather than late local positions).

These models have received varying degrees of support from neural and behavioral data. Consistent with the sharp random drift model (Horner et al., 2016), neural studies have provided evidence for stronger neural responses (Baldassano et al., 2017; Kurby & Zacks, 2008; Sridharan et al., 2007; Zacks et al., 2001) and sharp shifts of neural representation at event boundaries (Baldassano et al., 2017; DuBrow et al., 2017; Zheng et al., 2022). However, as suggested by a more recent study (Pu et al., 2022), the sharp random drift model failed to capture the distance effect in temporal order memory. In contrast, although the context-resetting model (Pu et al., 2022) could account for various behavioral effects, there is scarce evidence to suggest that temporal context resets to that of the very first item in a sequence at event boundaries. This limitation is critical as one of the major goals of computational neuroscience is to develop neurally feasible models (Whittington et al., 2020).

The current study aimed to develop a cognitively and neurally feasible model to account for the effects of event boundaries on temporal order memory and related behavioral phenomena. Our model is inspired by mounting evidence showing reliable position coding during sequence processing. For example, behavioral studies have revealed two lines of evidence supporting position coding. First, intrusions from prior lists tend to occur in the same position as their previous within-list position (Conrad, 1960; Fischer-Baum & McCloskey, 2015; Osth & Dennis, 2015). Second, when a list is grouped into several groups (e.g., by a longer temporal pause after presentation of every third item), items recalled from incorrect group tend to occur in the same within-group position as from their original group (Hensen, 1999; Ng & Mayberry, 2002). Later modeling work suggests that these effects could only be explained by models incorporating position-coding mechanism (Logan & Cox, 2023; Osth & Hurlstone, 2023). Using neural imaging technologies and multivoxel pattern analysis, recent studies have also observed clear neural representation of position codes in sequence memory (Hsieh et al., 2014; Liu et al., 2022; Liu et al., 2019; Xie et al., 2022).

This current study aimed to develop a new position-coding model and evaluated its performance by fitting it to previously reported various behavioral effects (Pu et al., 2022). We further conducted two new experiments to test a novel prediction from the position-coding model, i.e., the absence of boundary number effect, and to provide additional data for model comparison. The results showed that our position-coding model could account for the full set of effects and outperformed the context-resetting model.

2. Study 1: The position-coding model fits existing behavioral data

Here, we developed a position-coding model grounded in well-reported evidence of position coding and temporal context drift. We tested its performance by fitting previous empirical data (Pu et al., 2022) that characterized various behavioral effects in temporal order memory, including the boundary, distance, and local primacy effects.

2.1. Methods

2.1.1. The temporal order memory task

In a typical temporal order memory experiment (Fig. 1a), a sequence of trial-unique items embedded in different color frames was presented sequentially at the encoding phase. Participants were encouraged to imagine the items within their respective colored frame and to remember the order of their appearance. At the retrieval phase, two items were chosen from the just-learned sequence, and participants needed to judge which item appeared later in the sequence.

2.1.2. The position-coding model

Our position-coding model was constructed based on two key assumptions. First, each item i carries two types of contextual signals (Fig. 1b): the temporal code (C_T^i) (where i indexes the list position of item) and the position code (C_P^j) (where j indexes the local position of item). These two components were concatenated to form a single vector (Equation (1)). Following Pu et al. (2022), the context was coded as a set of binary elements (i.e., 0 or 1) with 100 features. The rate at which the temporal code drifted from item m – 1 to item m was regulated by the parameter ρ_T, which determined the probability that the binary elements did not change (i.e., a ρ_T value close to 1 indicated slower drift) (Equation (2)). Similarly, the position code changed sequentially as items progressed from local position 1 to n, with the rate of change regulated by the parameter ρ_P (Equation (3)). This assumption was based on previous studies, which found adjacent positions shared more similar neural representations compared to far positions (Liu et al., 2022). For simplicity, we assumed equal weighting of the temporal and position codes (50 % vs. 50 %) in the contextual representation. C_T^1 and C_P^1 were binarized vectors with 50 features, which were created by randomly selecting 50 values from an evenly distributed interval (0,1) and then rounding them up into either 0 or 1. In the actual simulation, a set of position vectors was pre-generated, with the number of vectors determined by the maximum number of items in any event within the simulated sequence. The C_P^j of each item i was chosen from this set, based on its position j within the event. The position code was reset to the first position whenever a new context began. In contrast, the temporal vector C_T^i of each item i was continuously updated with each new contextual representation.

$$C^i = C_T^i \oplus C_P^j \tag{1}$$

$$C_T^i = \begin{cases} C_T^1, i = 1 \\ \rho_T C_T^{i-1} + (1-\rho_T) C_T^{noise}, 2 \le i \le m \end{cases} \tag{2}$$

$$C_P^j = \begin{cases} C_P^1, j = 1 \\ \rho_P C_P^{j-1} + (1-\rho_P) C_P^{noise}, 2 \le j \le n \end{cases} \tag{3}$$

Second, we posited that the item with a more dissimilar contextual representation to the first item would be judged as more recent. Following Pu et al. (2022), we used the Memory Index (MI) to quantify recency judgments. Specifically, MI = $D_2 - D_1$, where D_2 refers to the dissimilarity in contextual representation between the more recent item and the first item in a sequence, and D1 refers to the dissimilarity in contextual representation between the more distant item and the first item. To be consistent with Pu et al., we did not include a "decision-making" process at retrieval and used the MI values to fit the averaged accuracy data from a group of subjects.

2.1.3. The context-resetting model

In the context-resetting model, the context of each item in a sequence was also represented as a vector of one hundred binary elements (Fig. 1c). The context fluctuated gradually at non-boundary time points, with the rate of fluctuation regulated by the parameter ρ_R (Equation (4)). At boundary time points, the context would reset to that of the first item at a rate determined by the parameter λ (Equation (5)). C_1 was created by randomly selecting 100 values from a uniform distribution between 0 and 1, which were then rounded to the nearest integer (either 0 or 1). In the actual simulation, for boundary items, some of the context features would reset to those of the first item in the sequence, while the remaining features would update according to the context drift algorithm.

$$C_t = \rho_R C_{t-1} + (1-\rho_R) C^{noise} \tag{4}$$

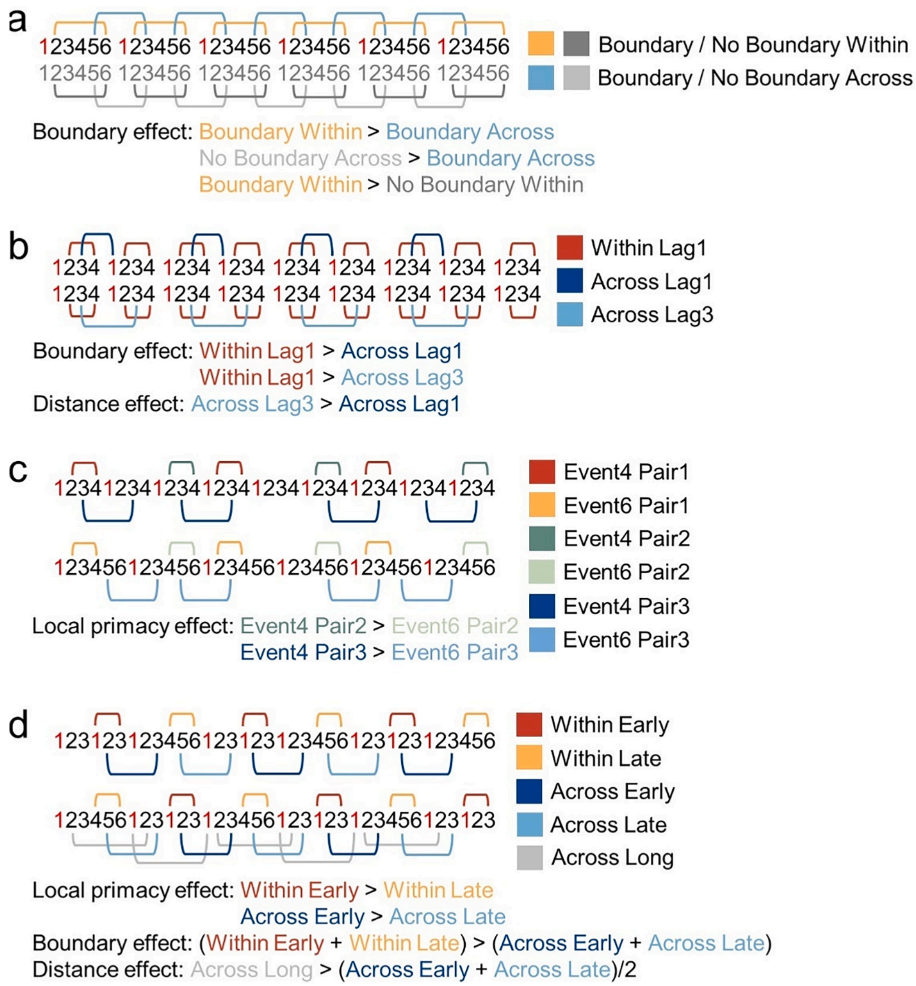


Fig. 2. Schematic plots of experimental designs. At the encoding stage, 36 trial-unique objects were presented. Event boundaries were manipulated by changing frame colors. The number 1, in red font, represents the first item of an event. Item pairs labeled with colored square brackets were tested in later recency judgment task. There were four, three, six, and five conditions in **a**, experiment 1, **b**, experiment 2, **c**, experiment 3, and **d**, experiment 4, respectively. In experiment 1, it is important to note that half of the sequences contained no change in frame colors, indicated by grey font.

$$C_t = (1-\lambda)\left(\rho_R C_{t-1} + (1-\rho_R)C^{noise}\right) + \lambda C_1 \tag{5}$$

As can be seen here, the two models involved different computational mechanisms. In our position-coding model, the context of each item consisted of two components, with their rates of change controlled by two independent parameters. In contrast, the context-resetting model assumed that the entire context vector of boundary items could be reset. The proportion ($1 - \lambda$) of the context for boundary items that is not reset, along with the context for non-boundary items, would gradually drift from that of the previous item.

2.1.4. The participants and experimental designs of Pu et al

We used the empirical data from Pu et al. (2022) to test the position-coding model. All participants in their study were recruited from Frankfurt am Main and its neighboring areas. All of them were right-handed and had normal or corrected-to-normal vision, normal color perception, and no past or present psychiatric disorders during the experiment.

The data consisted of four experiments using the typical temporal order memory paradigm (Fig. 1a). Fig. 2 shows the experimental designs, and Table 1 lists the participants' demographics and the experimental conditions. For all experiments, object images were chosen from a standardized stimulus bank (Brodeur et al., 2010; Brodeur et al., 2014) and resized to 350 by 350 pixels. Each object was presented for 2.5 s during the encoding phase, followed by a 2-second inter-trial interval during which the colored frame remained. Before each object presentation, a fixation cross appeared within the colored frame for 0.5 s. When a boundary occurred, the color of the frame was updated simultaneously with presentation of next object. The colors were easily distinguishable and did not repeat within a single learn-test round. Participants needed to remember the order of 36 trial-unique objects in each sequence and indicate whether they liked the combinations of objects and colored frames by pressing a button. Immediately after learning, they were asked to make recency judgments for pairs of just-encoded objects. There were 14 learn-test rounds in each experiment.

2.1.5. Model simulation

Thirty-six context signals were simulated in each experiment, corresponding to the number of items learned at the encoding phase. Event boundaries were set according to the specific experimental design (Fig. 2). Cosine dissimilarity coefficient (i.e., 1 – cosine similarity) quantified the dissimilarity between two contextual representations. The temporal context drift rate ρ_T and position change rate ρ_P varied from 0.01 to 0.99 with a step size of 0.01. Two hundred iterations were simulated for each experiment and each combination of parameter values. The simulations were performed in MATLAB R2017b (The MathWorks, Inc.).

2.1.6. Statistical analysis

We used the MI values to fit the empirical data (ACC) of all experimental conditions, using generalized linear mixed model (GLMM) and maximum likelihood estimation. To control potential dependency of group-averaged ACC in the same experiment, GLMM was specified to explain ACC as a function of MI and experimental group (ExpGroup) specific intercept and slope (Equation (6)). "ACC" is

Table 1
The demographics and experimental designs.

Demographic		Design	
Number of Subjects (female)	Mean Age (years)	Experimental Conditions	Number of Trials
Experiment 1 (Fig. 2a)			
N = 26 (18)	25	Boundary Within	42
		Boundary Across	35
		No Boundary Within	42
		No Boundary Across	35
Experiment 2 (Fig. 2b)			
N = 27 (18)	24	Within Lag 1	126
		Across Lag 1	56
		Across Lag 3	56
Experiment 3 (Fig. 2c)			
N = 31 (18)	26	Event4 Pair1	21
		Event6 Pair1	21
		Event4 Pair2	21
		Event6 Pair2	21
		Event4 Pair3	28
		Event6 Pair3	28
Experiment 4 (Fig. 2d)			
N = 30 (22)	24	Within Early	42
		Within Late	42
		Across Early	35
		Across Late	35
		Across Long	70

the response variable. “1” and “MI” represent the constant (i.e., intercept) and the predictor, which should be estimated, respectively. The term (1 + MI | ExpGroup) means that a random intercept and random slope of ExpGroup need to be estimated. The GLMM, specified by Equation (6), was implemented in MATLAB R2017b (The MathWorks, Inc.) using the function fitglme.

$$ACC \sim 1 + MI + (1 + MI|ExpGroup) \quad (6)$$

In addition to explained variance (i.e., R^2) of GLMM, Akaike information criterion (AIC) and Bayesian information criterion (BIC) were also computed to evaluate fitting performance. When making model comparisons, the model with lower AIC and BIC was considered to have better performance.

In addition to evaluating overall model fit, we also tested whether the model could capture individual behavioral effects. Specifically, for the boundary effect, behavioral performance was better in the “Within” than the “Across” condition. For the distance effect, pairs with a larger temporal lag showed better order memory. For the local primacy effect, accuracy was higher for pairs located at the early local positions than for those located at late local positions. To test these effects, one-tailed Student’s t-tests ($\alpha = 0.05$) were conducted on MI values from 200 iterations, treating each iteration as a subject.

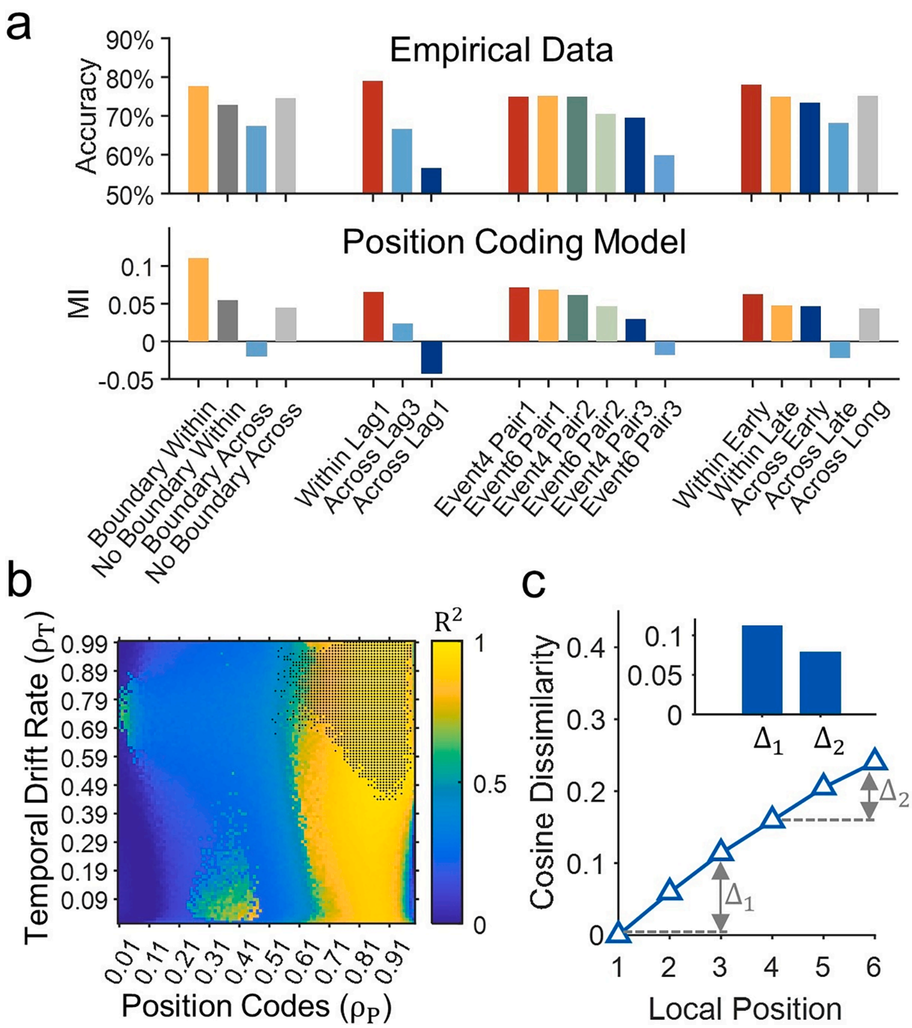


Fig. 3. Performance of the position-coding model. **a**, Accuracy of Pu et al. experiments (upper panel) and MI outputs of the position-coding model (with $\rho_T = 0.94$ and $\rho_P = 0.94$) (bottom panel). **b**, R^2 obtained by fitting empirical data with MI outputs from the position-coding model under different combinations of parameter values using GLMM. Both parameters were set from 0.01 to 0.99 with a step size of 0.01. The black dots indicate the combinations of parameter values that could capture all behavioral effects. **c**, Cosine dissimilarity of each local position to the first local position. The dissimilarity of other positions to position 1 starts steep and later tends to flatten, which could help capture the local primacy effect. For example, the difference in cosine dissimilarity to local position 1 between the early local positions (i.e., Δ_1) was larger than that between the late local positions (i.e., Δ_2).

2.2. Results

2.2.1. Model fitting performance

Overall, results showed that the position-coding model could successfully explain the data (i.e., mean accuracy, Fig. 3a). The model ($\rho_T = 0.94$ and $\rho_P = 0.94$) accounted for 98 % of the variance in the behavioral pattern (GLMM, $p < 0.001$, AIC = 88.73, BIC = 94.07).

Meanwhile, the position-coding model captured various behavioral effects (Fig. 3b, indicated by black dots). For example, with $\rho_T = 0.94$ and $\rho_P = 0.94$, the MI was higher for "Boundary Within" (MI = 0.1101) than for "Boundary Across" condition (MI = −0.0199) (one-tailed, $t(199) = 34.14$, $p < 0.001$), capturing the boundary effect. We also observed significant distance effect, with a higher MI for "Across Lag3" (MI = 0.0238) compared to "Across Lag1" trials (MI = −0.0432) (one-tailed, $t(199) = 37.32$, $p < 0.001$). For the local primacy effect, "Within Early" trials (MI = 0.0629) exhibited higher MI values compared to "Within Late" trials (MI = 0.0476) (one-tailed, $t(199) = 8.12$, $p < 0.001$). Similarly, "Across Early" trials (MI = 0.0468) showed higher MI values than "Across Late" trials (MI = −0.0223) (one-tailed, $t(199) = 32.20$, $p < 0.001$) (Fig. 3a).

To probe the model further, we plotted the dissimilarity of each positional context vector to the first (Fig. 3c) and the last local positions (Fig. S1 in Supplementary material). It was clear that the dissimilarity of each positional context vector to the first one decreased deceleratingly, which helped to capture the distance effect and the local primacy effect (the inset of Fig. 3c).

2.2.2. Other weight ratios of temporal and position codes

In the above simulation, we assumed an equal weight for temporal and positional codes (i.e., 50 % vs. 50 %). To examine whether our results were robust to different combinations of temporal and positional codes, we systematically varied the weight of the position code from 30 % to 70 % and repeated the simulation and GLMM fitting procedures as described above. Results showed that other weight ratios also worked well ($R^2 > 0.85$, all $p < 0.001$) (Supplementary material, Fig. S2). When the weight of the positional code was very low (e.g., 5 %), the model could no longer capture overall behavioral pattern (GLMM, $R^2 = 0.38$, $p > 0.05$) (Supplementary material, Fig. S3), emphasizing the critical role of position coding. In subsequent simulations, we only focused on the model with an equal weight ratio.

3. Study 2: Test new prediction from the position-coding model

Study 1 suggests that our model provides a good fit for the boundary, distance, and local primacy effects. In addition, our model could predict a new effect. That is, tested pairs sharing the same local and list positions would exhibit comparable temporal order memory (Fig. 4a), regardless of the number of boundaries between them **(i.e., the absence of boundary number effect)** (Fig. 4b). This is because MI is determined solely by list position (that affects temporal code) and local position (which affects position code). In contrast, the context-resetting model would predict that temporal order memory for short events with more boundaries (e.g., Across Two) would be worse than that for longer events (e.g., Across One) (Fig. 4c). This is because the former condition experiencing more resetting, leading to a smaller MI **(i.e., boundary number effect)**. It should be noted that under certain conditions (e.g., when the likelihood of resetting is low), the context-resetting model could also predict equivalent temporal memory for the two conditions. However, the parameters might not account for the full effects. In Study 2, we conducted two new experiments to test these hypotheses. In addition, by combining data from both Pu et al. and the new experiments, we could do more comprehensive model comparisons.

3.1. Method

3.1.1. Participants

The number of participants in the two new experiments was determined based on the existing study (Pu et al., 2022), which suggested a minimum sample size of 26. In experiment 1, 29 participants (females = 19, mean age = 21.45) were included in the final data analysis. One additional participant was recruited but excluded from further analysis due to errors in data recording.

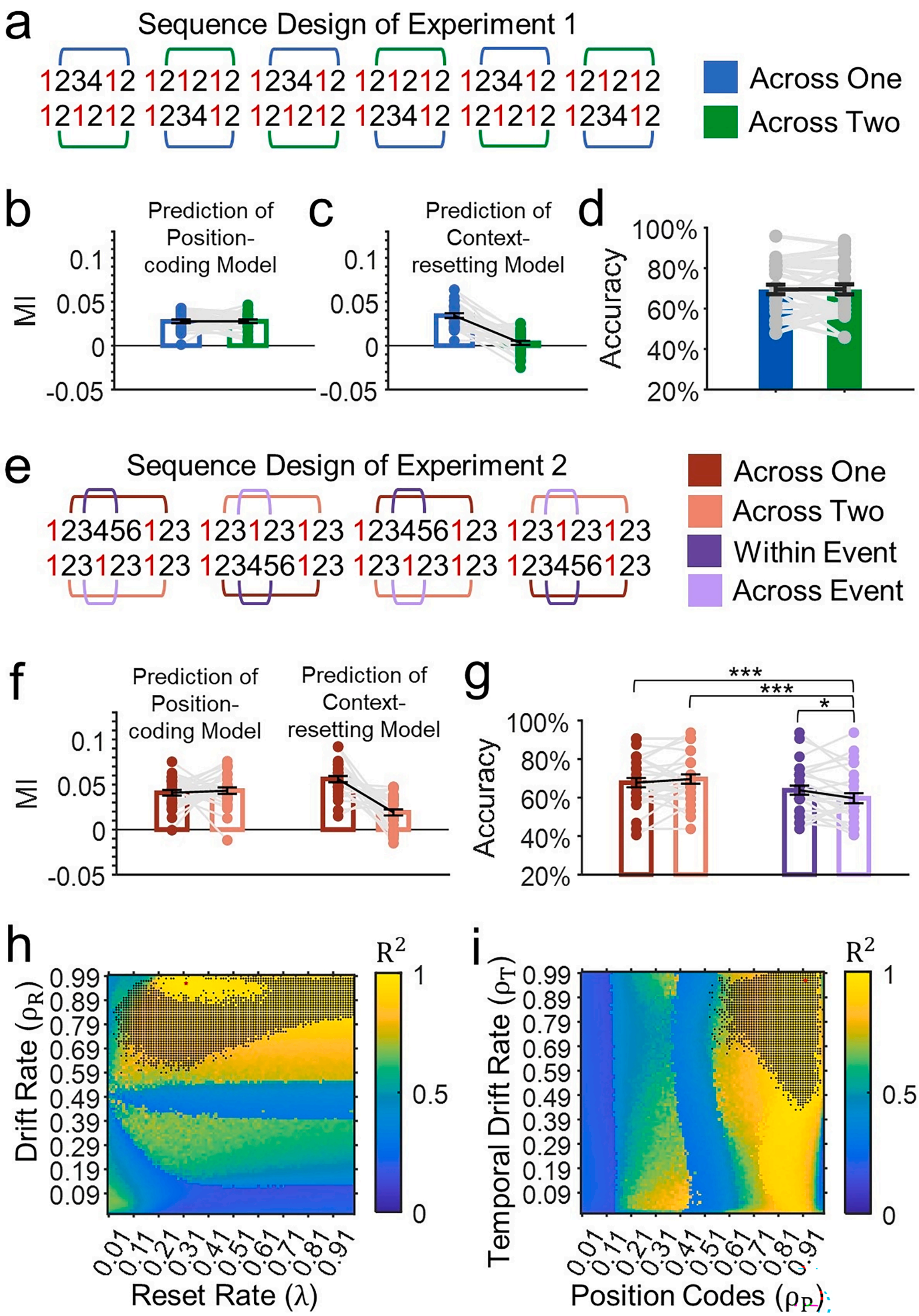
In experiment 2, 28 participants (females = 26, mean age = 20.79) were included in the final data analysis. Two additional participants were recruited but excluded due to errors in data recording. Another two participants were excluded because of overall accuracy lower than 50 % (i.e., chance level).

All participants were recruited from Beijing Normal University and reported having normal or corrected-to-normal vision and normal color perception. The Ethics Council of Beijing Normal University approved the experiments, and informed consent was obtained prior to the experiments.

3.1.2. Experimental designs and procedures

The experimental materials and procedures were adopted from previous study (Pu et al., 2022) (as described in Study 1) and were programmed via MATLAB R2020b and Psychtoolbox3.0.18. Briefly, during encoding, participants were required to remember the temporal order of 36 trial-unique object images embedded in colored frames and simultaneously rated their liking for the combinations of objects and colors. The retrieval task was administered immediately following learning, where participants needed to make recency judgments on image pairs. Items from the first half of the sequence were tested first.

In experiment 1, four groups of 15 unique RGB colors were chosen as background colors and were recycled after every four encoding-retrieval blocks. Two types of lists were created, with items grouped by background colors as follows (i.e., the numbers indicate the number of items sharing the same background color): 4–2–2–2–2–4–2–2–2–2–4–2–2–2–2 (list type 1, 8 sequences) or 2–2–2–4–2–2–2–2–4–2–2–2–2–4–2 (list type 2, 8 sequences). Sixteen encoding-retrieval blocks (i.e., 16 sequences) were divided into



(caption on next page)

Fig. 4. Experiments, model predictions, and fitting results. **a**, Design of experiment 1. The number 1 with red font represents the first item within a local event, and square brackets label the tested pairs at the retrieval phase, where participants were asked to make recency judgment of two items. Predictions of **b**, the position-coding model (with $\rho_T = 0.95$ and $\rho_p = 0.90$) and **c**, the context-resetting model (with $\rho_R = 0.97$ and $\lambda = 0.31$) based on 29 iterations, with parameter values having the best fitting performance of Pu et al. data. **d**, Behavioral results of experiment 1. Error bar represents standard error of mean. **e**, Design of experiment 2. Four conditions were tested. **f**, Predictions of the position-coding model and the context-resetting model based on 28 iterations, using parameter values showing the best fitting performance of Pu et al. data. **g**, Behavioral results of experiment 2. Error bar represents standard error of mean. *$p < 0.05$, *$p < 0.001$. R^2 was obtained by using GLMM to fit Pu et al. and our new data with the MI outputs of **h**, the context-resetting model and **i**, the position-coding model. All parameters were set from 0.01 to 0.99 with a step size of 0.01. The black dots indicate the combinations of parameter values that could capture the full set of behavioral effects in Pu et al. and two new experiments. The red pentagram indicates the combination of parameter values that best fit all the data. Specifically, when using the parameter values that provided the best fit, the context-resetting model ($\rho_R = 0.96$ and $\lambda = 0.32$) failed to account for the absence of boundary number effect, whereas the position-coding model ($\rho_T = 0.96$ and $\rho_p = 0.92$) could capture all the effects.

four mini-blocks. Within each mini-block, four sequences of the same type (either list type 1 or 2) were learned and tested. The order of mini-blocks was counterbalanced across participants (i.e., the list order was either 1111–2222–1111–2222 or 2222–1111–2222–1111). Two types of pairs were tested (i.e., “Across One” vs. “Across Two”) (Fig. 4a). In total, there were 48 pairs of “Across One” and “Across Two” trials, respectively.

In experiment 2, four groups of 10 unique RGB colors were selected and recycled after every four encoding-retrieval blocks. Two types of lists were created: 6–3–3–3–3–6–3–3–3–3 (list type 1, 8 sequences) or 3–3–3–6–3–3–3–3–6–3 (list type 2, 8 sequences). Sixteen encoding-retrieval blocks were divided into four mini-blocks within which four sequences of the same list type were learned and tested. The order of mini-blocks was counterbalanced across participants. To address the concern that the boundary manipulation in the experiment 1 might not be salient enough due to frequent background changes, we reduced the number of boundaries in experiment 2. Additionally, we aimed to replicate the boundary effect and the distance effect. If we are able to replicate the absence of boundary number effect along with these classical effects in the same paradigm, it would provide more conclusive evidence to validate our design and to support the position-coding model.

There were 32 pairs for each condition: Across One (i.e., pairs across one boundary), Across Two (i.e., pairs across two boundaries), Within Event (i.e., pairs within the same event), and Across Event (i.e., pairs across different events but with shorter list distance than that in conditions “Across One” and “Across Two”). The boundary effect is defined as “Within Event” > “Across Event”, and the distance effect is defined as “Across One” > “Across Event” and “Across Two” > “Across Event”.

3.1.3. Statistical analysis

The position-coding model predicted comparable memory performance between “Across One” and “Across Two” trials, so Bayes factor (BF_{01}) was computed by JASP 0.19.0.0 to compare relative evidence for the hypothesis that “Across One” was equal to “Across Two” (H0: $\mu 1 = \mu 2$) over the alternative hypothesis that “Across One” was different to “Across Two” (H1: $\mu 1 \neq \mu 2$). The context-resetting model predicted higher accuracy for “Across One” than for “Across Two” condition, so one-tailed Student’s *t*-test ($\alpha = 0.05$) was conducted. To further examine the context-resetting model, equivalence test (Lakens et al., 2018) was conducted to test whether the mean difference was smaller than that predicted by the context-resetting model, using the TOSTER R package. Contrary to classical null hypothesis significance test, equivalence test aims to prove that means of two conditions are close enough to be considered equivalent (i.e., H0: $\mu 1 \neq \mu 2$ and H1: $\mu 1 = \mu 2$). Equivalence bounds are set to test whether the effect size of the difference between two conditions is equal to or smaller than the smallest effect size of interest (SESOI). In our study, the lower and upper equivalence bounds were determined by the effect sizes predicted by the context-resetting model with parameters that could best fit Pu et al. data. The significance of equivalence test indicated that the observed effect size was smaller than the prediction made by the context-resetting model, thereby providing evidence against the context-resetting model.

In addition, one-tailed Student’s t-tests ($\alpha = 0.05$) were used to estimate the boundary and distance effects in the experiment 2, based on previous findings that temporal order memory was better for within than across event trials, as well as for item pairs with longer than shorter distances.

3.1.4. Model recovery simulation

To further demonstrate the differences between the two models, we performed a model recovery analysis to examine how well the data generated by the context-resetting model and the position-coding model could be recovered by each model. For each model, 17*17 = 289 datasets were generated using all combinations of 17 values for each parameter (0.01:0.06:0.99, i.e., from 0.01 to 0.99 with a step size of 0.06). The weight ratio of position code for the position-coding model was fixed at 50 %. Each dataset was separately fitted with MI outputs of the context-resetting model and the position-coding model using GLMM fitting. The AIC from the best-fitting model would be compared.

3.2. Results

In experiment 1, the means (standard deviations) of accuracy for “Across One” and “Across Two” conditions were 0.6947 (0.1275) and 0.6954 (0.1331), respectively (Fig. 4d). The one-tailed Student’s *t*-test was not significant, $t(28) = -0.0404$, $p = 0.516$, Cohen’s d = 0.0075. $BF_{01} = 5.062$, which showed moderate evidence for the hypothesis that “Across One” was equal to “Across Two”. This is consistent with the prediction by the position-coding model using parameters that could best fit Pu et al. data ($\rho_T = 0.95$ and $\rho_p = 0.90$)

(Fig. 4b; one-tailed Student's t-test: $t(28) = 0.23$, $p = 0 .41 > 0.05$; equivalence test: $t(28) = 8.047$, $p = < 0.001$).

Equivalence test was further conducted to test whether the effect size of the difference between two conditions is equal to or smaller than the effect size predicted by the context-resetting model with parameters that could best fit Pu et al. data ($\rho_R = 0.97$ and $\lambda = 0.31$). Given the predicted effect size (i.e., Cohen's d = 1.4972 for the experiment 1), the equivalence test was significant ($t(28) = 8.022$, $p < 0.001$; two one-sided tests, α = 0.05, 90 % CI: [-0.031, 0.03]), which challenged the context-resetting model. For a more stable model prediction, we repeated the procedure on MI values from 200 iterations, which yielded consistent results (Supplementary material, Fig. S4).

In experiment 2 (Fig. 4e-g), we replicated the classical boundary effect ("Within Event" > "Across Event", one-tailed $t(27) = 2.0371$, $p = 0 .0258 < 0.05$, 90 % CI = [0.0068, 0.0758]) and the distance effect ("Across One" > "Across Event", one-tailed $t(27) = 3.5854$, $p < 0.001$, 90 % CI = [0.0422, 0.1185]; "Across Two" > "Across Event", one-tailed $t(27) = 4.4945$, $p < 0.001$, 90 % CI = [0.0617, 0.1370]) on temporal order memory.

More importantly, we also replicated the null effect of the boundary number effect. Specifically, accuracy for "Across One" was comparable to that for "Across Two" (Fig. 4g, one-tailed $t(27) = -0.9929$, $p = 0.8352$, Cohen's d = 0.1876). Using parameters that could best fit Pu et al. data (Fig. 4f), the position-coding model predicted equal performance between "Across One" and "Across Two" (one-tailed Student's t-test: $t(27) = -0.8309$, $p = 0.7933$; equivalence test: $t(27) = 5.745$, $p < 0.001$), whereas the context-resetting model predicted better performance for "Across One" than "Across Two" (one-tailed Student's t-test: $t(27) = 3.8183$, $p < 0.001$; equivalence test: $t(27) = 0.001$, $p = 0 .50 > 0.05$). These predictions could be replicated based on 200 iterations (Supplementary material, Fig. S5). Given the effect size predicted by the context-resetting model (i.e., Cohen's d = 1.1670), equivalence test was significant (two one-sided tests, $t(27) = 5.181$, $p < 0.001$, α = 0.05, 90 % CI: [-0.052, 0.014]). Bayesian inference revealed moderate evidence for the hypothesis that "Across One" was equal to "Across Two" ($BF_{01} = 3.187$). These data again supported the position-coding hypothesis but not the context-resetting model.

As predicted, using some combinations of parameter values, the context-resetting model could also predict equivalent performance for "Across One" and "Across Two" conditions in two new experiments (indicated by black dots in Fig. 4h). But these combinations did not overlap with the combination that had the best overall fitting performance of all data including Pu et al. and two new ones (indicated by red pentagram in Fig. 4h). Using the combination of parameter values which had the best overall fitting (Fig. 4h-i), the position-coding model (with $\rho_T = 0.96$ and $\rho_p = 0.92$, GLMM $R^2 = 0.9829$, $p < 0.001$, AIC = 123.7998, BIC = 130.8681) showed smaller AIC and BIC than the context-resetting model (with $\rho_R = 0.96$ and $\lambda = 0.32$, GLMM $R^2 = 0.9929$, $p < 0.001$, AIC = 130.8208, BIC = 137.8891). Furthermore, the position-coding model could capture all behavioral effects, whereas the context-resetting model could not account for the lack of boundary number effect in the two experiments.

To further compare two models, we performed a model recovery simulation to examine whether the models used to simulate the data had the best fit to their own generated data. The results demonstrated that two models were distinct from each other and were able to capture the unique characteristics of the data they produced. Specifically, when fitting the datasets generated by the context-resetting model (Fig. 5a), the least AIC was smaller for the context-resetting model than the position coding model (one-tailed $t(288) = -34.8897$, $p < 0.001$). A reversed pattern was found when fitting the datasets generated by the position-coding model (Fig. 5b, one-tailed $t(288) = 22.3402$, $p < 0.001$).

Finally, we found the Horner et al. (2016) model was unable to capture the full set of behavioral effects in Pu et al. or our two new experiments (Supplementary material, Fig. S6).

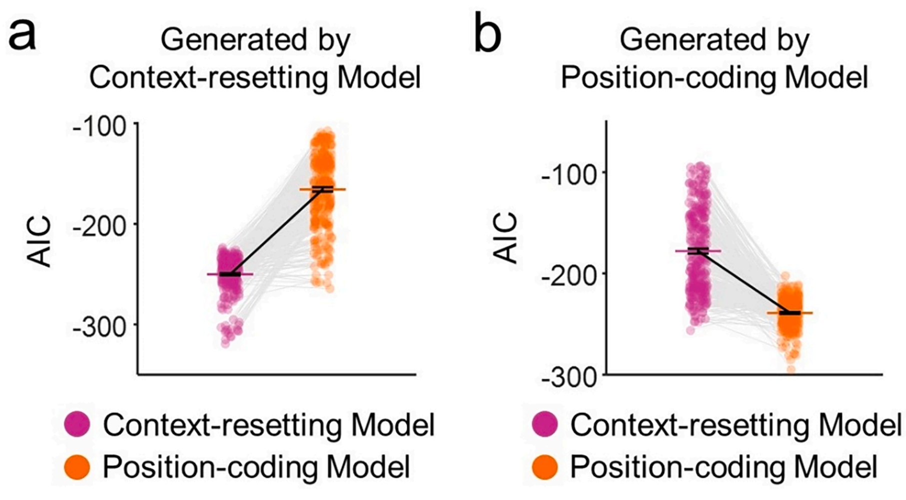


Fig. 5. Model recovery simulation. **a**, When fitting the datasets generated by the context-resetting model, the least AIC values of the context-resetting model (purple) were smaller than those of the position-coding model (orange). **b**, A reversed pattern was found when fitting the datasets generated by the position-coding model. Each translucent dot indicates the least AIC in each simulation and GLMM fitting. Error bar indicates standard error of mean.

4. Discussion

4.1. General discussion and theoretical implications

The current study aimed to provide a cognitively, neurally, and computationally feasible account of the effects of event boundaries on temporal order memory. We developed a new model based on abundant evidence of position coding and found it could explain various behavioral effects. We conducted two new experiments and provided new data for model comparison and development. The position-coding model had a better fitting performance of all the data than the context-resetting model. These results suggest that our position-coding model can explain a variety of effects of event boundaries on temporal order memory.

Our model provides an intuitive explanation of how event boundaries influence temporal order memory. Building on the work of Pu et al. (2022), it proposes that the temporal order of two items is determined by the difference in their contextual representations relative to the context of the first item. An item with a smaller difference in contextual representation from the first item is perceived as occurring earlier. A larger difference in temporal context between two items (measured by the MI metric) correlates with greater accuracy in judging their temporal order. Using the same metric, our model offers a more intuitive explanation for prior findings. When list distances are equal for both within-boundary and cross-boundary pairs, the local position distance for cross-boundary pairs tends to be smaller than for within-boundary pairs. In some cases, the local position sequence is reversed for cross-boundary pairs, meaning post-boundary items appear earlier in local positions than pre-boundary items. This results in a smaller or even negative MI value, leading to poorer memory accuracy for cross-boundary items. Furthermore, as differences in position codes diminish at later local event positions, this may explain the local primacy effect.

Although both the position-coding and context-resetting models demonstrate good overall performance in fitting existing behavioral data, our model has several advantages. First, as stated in the Introduction, numerous behavioral and neural studies have provided evidence for the existence of position coding. In contrast, evidence for reinstatement of the first item's contextual representation at every event boundary is rarely reported.

Second, under most situations, the position-coding and context-resetting models make contrasting predictions regarding the impact of event boundary numbers on temporal order memory. The position-coding model suggests that temporal order memory is determined only by list position and local position, with the number of event boundaries having no effect. In contrast, the context-resetting model generally predicts that more boundaries would be associated with more context resetting, which would decrease MI and impair temporal order memory.

Third, our model assumes two types of temporal information: a slowly drifted one-shot temporal context and position coding, which is consistent with recent neural studies. For example, rodent research has revealed two forms of temporal order representation in the hippocampal–entorhinal system: stable event sequence representation, which strengthen with repeated practice in structured events, and random drift of temporal context that forms automatically through one-shot learning (Tsao et al., 2018). Human neuroimaging studies have similarly demonstrated that both temporal context (Ezzyat & Davachi, 2014; Howard et al., 2005; Jenkins & Ranganath, 2016) and structured event sequences (Bellmund et al., 2019; Deuker et al., 2016; Liu et al., 2022) are represented in the hippocampal–entorhinal system and are crucial for temporal memory.

Position codes can facilitate sequence processing and sequence memory in several ways. First, they provide an effective approach to maintain order information in working memory (Xie et al., 2022). Second, in specific mnemonic strategies, such as the method of loci, position coding provides valuable contextual information to organize memory and preserve temporal order (Liu et al., 2022). Third, even when there is no explicit requirement to remember temporal order, position code can dissociate content from structure (Fan et al., 2021). This factorized representation of content and temporal position facilitates flexible reorganization of information according to rules (Liu et al., 2019) and supports the transfer of knowledge to novel situations (Behrens et al., 2018; Xue, 2022).

4.2. Limitations and future research

Several important questions remain to be addressed in future studies. First, the current study used color frames as contexts to produce experiences of event boundaries. Consistent with previous studies (Heusser et al., 2018; Pu et al., 2022), this paradigm successfully generated several classical event boundary effects. A recent fMRI study using this paradigm also obtained neural evidence of the boundary effect, such as higher similarity for within- than cross-event items in hippocampal CA3 (Bein & Davachi, 2023). Future studies could examine the position-coding model using other approaches to generate event boundaries, such as temporal pauses (Hensen, 1999; Ng & Mayberry, 2002). Second, the models used MI to represent behavioral performances, by comparing the dissimilarity of two items' contexts to that of the first item in the sequence. Alternatively, participants might use the most recent, the first, or other salient items to anchor temporal positions. Future modeling work should explore these possibilities. Third, existing studies have suggested that hierarchical representations of event order may exist, including local and global position codes (Brown et al., 2000; Henson & Burgess, 1997; Yokoi & Diedrichsen, 2019). One existing study has adopted a hierarchical representation of contextual information to explain behavioral effects during serial and free recall (Farrell, 2012). Future modeling work could further examine how different levels of position codes contribute to temporal order memory. Fourth, the current study examined the temporal order of discrete events but not continuous events. Note that position codes would be largely reduced for continuous events (e.g., using a naturalistic paradigm). In this case, our model would predict better memory accuracy for across than within event pairs (see Fig. S3 in Supplementary material). This prediction is consistent with the findings from a recent behavioral study (Zuo et al., 2020). In this study, both monkeys and humans watched a video consisting of two clips and then made a primacy judgement between two frames extracted from the video. The results revealed higher temporal order memory accuracy for cross-clip pairs compared to within-clip pairs, suggesting event boundaries did not impair temporal order memory. Future modeling work should be conducted to test these effects directly. Finally, neither our model nor that of Pu et al. could account for the RT effects in short-term or long-term order memory (Hacker, 1980; Pu et al., 2022). Presumably, forward or backward scanning mechanisms are required to capture the RT data. Future studies incorporating other mechanisms (e.g., chaining model) may be required to account for these RT effects.

In summary, the current study suggests that event boundaries influence temporal order memory by resetting local position coding of items. This provides a computationally simple and neurally feasible account of the role of event boundaries in temporal order memory. Future studies should combine neuroimaging and modeling approaches to test our model under different experimental conditions and across various populations, as well as to uncover the underlying mechanisms.

CRediT authorship contribution statement

Xiaojing Peng: Writing – review & editing, Writing – original draft, Visualization, Software, Methodology, Formal analysis, Data curation, Conceptualization. **Yifei Cao:** Methodology. **Jintao Sheng:** Methodology. **Yu Zhou:** Methodology. **Huinan Hu:** Methodology. **Gui Xue:** Writing – review & editing, Writing – original draft, Supervision, Software, Resources, Methodology, Funding acquisition, Formal analysis, Conceptualization.

Acknowledgement

This work was supported by the National Natural Science Foundation of China (32441110 and 32330039). The studies were not preregistered.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.cogpsych.2025.101714.

Data availability

The simulation code, experimental materials, and experimental data are at this link: https://github.com/HiuzingPaang/TOM_online_materials.git.

References

Baldassano, C., Chen, J., Zadbood, A., Pillow, J. W., Hasson, U., & Norman, K. A. (2017). Discovering Event Structure in Continuous Narrative Perception and Memory. *Neuron, 95*(3), 709–721 e705. https://doi.org/10.1016/j.neuron.2017.06.041

Behrens, T. E. J., Muller, T. H., Whittington, J. C. R., Mark, S., Baram, A. B., Stachenfeld, K. L., & Kurth-Nelson, Z. (2018). What Is a Cognitive Map? *Organizing Knowledge for Flexible Behavior. Neuron, 100*(2), 490–509. https://doi.org/10.1016/j.neuron.2018.10.002

Bein, O., & Davachi, L. (2023). Event integration and temporal differentiation: How hierarchical knowledge emerges in hippocampal subfields through learning. *The Journal of Neuroscience*. , Article JN-RM-0627-0623. https://doi.org/10.1523/jneurosci.0627-23.2023

Bellmund, J. L., Deuker, L., & Doeller, C. F. (2019). Mapping sequence structure in the human lateral entorhinal cortex. *Elife, 8*. https://doi.org/10.7554/eLife.45333

Brodeur, M. B., Dionne-Dostie, E., Montreuil, T., & Lepage, M. (2010). The Bank of Standardized Stimuli (BOSS), a New Set of 480 Normative Photos of Objects to Be Used as Visual Stimuli in Cognitive Research. *Plos One, 5*(5), Article e10773. https://doi.org/10.1371/journal.pone.0010773

Brodeur, M. B., Guerard, K., & Bouras, M. (2014). Bank of Standardized Stimuli (BOSS) Phase II: 930 New Normative Photos. *Plos One, 9*(9). https://doi.org/10.1371/.

Brown, G., Preece, T., & Hulme, C. (2000). Oscillator-based memory for serial order. *Psychol Rev, 107*(1), 127–181.

Clewett, D., DuBrow, S., & Davachi, L. (2019). Transcending time in the brain: How event memories are constructed from experience. *Hippocampus, 29*(3), 162–183. https://doi.org/10.1002/hipo.23074

Conrad, R. (1960). Serial order intrusions in immediate memory. *British Journal of Psychology, 51*, 45–48.

Deuker, L., Bellmund, J. L., Navarro Schroder, T., & Doeller, C. F. (2016). An event map of memory space in the hippocampus. *Elife, 5*. https://doi.org/10.7554/eLife.16534

DuBrow, S., & Davachi, L. (2013). The influence of context boundaries on memory for the sequential order of events. *J Exp Psychol Gen, 142*(4), 1277–1286. https://doi.org/10.1037/a0034024

DuBrow, S., Rouhani, N., Niv, Y., & Norman, K. A. (2017). Does mental context drift or shift? *Curr Opin Behav Sci, 17*, 141–146. https://doi.org/10.1016/j.cobeha.2017.08.003

Estes, W. K. (1955). Statistical theory of spontaneous recovery and regression. *Psychol Rev, 62*(3).

Ezzyat, Y., & Davachi, L. (2014). Similarity breeds proximity: Pattern similarity within and across contexts is related to later mnemonic judgments of temporal proximity. *Neuron, 81*(5), 1179–1189. https://doi.org/10.1016/j.neuron.2014.01.042

Fan, Y., Han, Q., Guo, S., & Luo, H. (2021). Distinct neural representations of content and ordinal structure in auditory sequence memory. *J Neurosci*. https://doi.org/10.1523/JNEUROSCI.0320-21.2021

Farrell, S. (2012). Temporal Clustering and Sequencing in Short-Term Memory and Episodic Memory. *Psychological Review, 119*(2), 223–271. https://doi.org/10.1037/a0027371.supp

Fischer-Baum, S., & McCloskey, M. (2015). Representation of item position in immediate serial recall: Evidence from intrusion errors. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 41*(5), 1426–1446.

Hacker, M. (1980). Speed and Accuracy of Recency Judgments for Events in Short-Term Memory. *J of Exp Psychol: Hum Learn Mem, 6*(6).

Hensen, R. N. (1999). Coding position in short-term memory. *International Journal of Psychology, 34*(5–6), 403–409.

Henson, R. N., & Burgess, N. (1997). *9-11 April). Representations of Serial Order 4th Neural*. London: Computation and Psychology Workshop.

Heusser, A. C., Ezzyat, Y., Shiff, I., & Davachi, L. (2018). Perceptual boundaries cause mnemonic trade-offs between local boundary processing and across-trial associative binding. *J Exp Psychol Learn Mem Cogn, 44*(7), 1075–1090. https://doi.org/10.1037/xlm0000503

Horner, A. J., Bisby, J. A., Wang, A., Bogus, K., & Burgess, N. (2016). The role of spatial boundaries in shaping long-term event representations. *Cognition, 154*, 151–164. https://doi.org/10.1016/j.cognition.2016.05.013

Howard, M. W., Fotedar, M. S., Datey, A. V., & Hasselmo, M. E. (2005). The Temporal Context Model in spatial navigation and relational learning: Toward a common explanation of medial temporal lobe function across domains. *Psychol Rev, 112*(1), 75–116. https://doi.org/10.1037/0033-295X.112.1.75

Howard, M. W., & Kahana, M. J. (2002). A Distributed Representation of Temporal Context. *Journal of Mathematical Psychology, 46*(3), 269–299. https://doi.org/10.1006/jmps.2001.1388

Hsieh, L. T., Gruber, M. J., Jenkins, L. J., & Ranganath, C. (2014). Hippocampal activity patterns carry information about objects in temporal context. *Neuron, 81*(5), 1165–1178. https://doi.org/10.1016/j.neuron.2014.01.015

Jenkins, L. J., & Ranganath, C. (2016). Distinct neural mechanisms for remembering when an event occurred. *Hippocampus, 26*(5), 554–559. https://doi.org/10.1002/hipo.22571

Kurby, C. A., & Zacks, J. M. (2008). Segmentation in the perception and memory of events. *Trends Cogn Sci, 12*(2), 72–79. https://doi.org/10.1016/j.tics.2007.11.004

Lakens, D., Scheel, A. M., & Isager, P. M. (2018). Equivalence Testing for Psychological Research: A Tutorial. *Advances in Methods and Practices in Psychological Science, 1*(2), 259–269. https://doi.org/10.1177/2515245918770963

Lewandowsky, S., & Murdock, B. B. (1989). Memory for Serial Order. *Psychol Rev, 96*(1), 25–57.

Liu, C., Ye, Z., Chen, C., Axmacher, N., & Xue, G. (2022). Hippocampal Representations of Event Structure and Temporal Context during Episodic Temporal Order Memory. *Cereb Cortex, 32*(7), 1520–1534. https://doi.org/10.1093/cercor/bhab304

Liu, Y., Dolan, R. J., Kurth-Nelson, Z., & Behrens, T. E. J. (2019). Human Replay Spontaneously Reorganizes Experience. *Cell, 178*(3), 640–652 e614. https://doi.org/10.1016/j.cell.2019.06.012

Logan, G. D., & Cox, G. E. (2023). Serial order depends on item-dependent and item-independent contexts. *Psychol Rev, 130*(6), 1672–1687. https://doi.org/10.1037/rev0000422

Michelmann, S., Hasson, U., & Norman, K. A. (2023). Evidence That Event Boundaries Are Access Points for Memory Retrieval. *Psychol Sci, 9567976221128206*. https://doi.org/10.1177/09567976221128206

Ng, H. L. H., & Mayberry, M. T. (2002). Grouping in short-term verbal memory: Is position coded temporally? *Quarterly Journal of Experimental Psychology, 55*(2), 391–424.

Osth, A. F., & Dennis, S. (2015). Prior-list intrusions in serial recall are positional. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 41*(6), 1893–1901.

Osth, A. F., & Hurlstone, M. J. (2023). Do item-dependent context representations underlie serial order in cognition? Commentary on Logan (2021). *Psychol Rev, 130*(2), 513–545. https://doi.org/10.1037/rev0000352

Pu, Y., Kong, X. Z., Ranganath, C., & Melloni, L. (2022). Event boundaries shape temporal organization of memory by resetting temporal context. *Nat Commun, 13*(1), 622. https://doi.org/10.1038/s41467-022-28216-9

Radvansky, G. A., & Zacks, J. M. (2017). Event Boundaries in Memory and Cognition. *Curr Opin Behav Sci, 17*, 133–140. https://doi.org/10.1016/j.cobeha.2017.08.006

Sridharan, D., Levitin, D. J., Chafe, C. H., Berger, J., & Menon, V. (2007). Neural dynamics of event segmentation in music: Converging evidence for dissociable ventral and dorsal networks. *Neuron, 55*(3), 521–532. https://doi.org/10.1016/j.neuron.2007.07.003

Tsao, A., Sugar, J., Lu, L., Wang, C., Knierim, J. J., Moser, M. B., & Moser, E. I. (2018). Integrating time from experience in the lateral entorhinal cortex. *Nature, 561*(7721), 57–62. https://doi.org/10.1038/s41586-018-0459-6

Wen, T., & Egner, T. (2022). Retrieval context determines whether event boundaries impair or enhance temporal order memory. *Cognition, 225*, Article 105145. https://doi.org/10.1016/j.cognition.2022.105145

Whittington, J. C. R., Muller, T. H., Mark, S., Chen, G., Barry, C., Burgess, N., & Behrens, T. E. J. (2020). The Tolman-Eichenbaum Machine: Unifying Space and Relational Memory through Generalization in the Hippocampal Formation. *Cell, 183*(5), 1249–1263 e1223. https://doi.org/10.1016/j.cell.2020.10.024

Xie, Y., Hu, P., Li, J., Chen, J., Song, W., Wang, X., & Wang, L. (2022). Geometry of sequence working memory in macaque prefrontal cortex. *Science, 375*, 632–639. https://doi.org/10.1126/science.abm0204

Xue, G. (2022). From remembering to reconstruction: The transformative neural representation of episodic memory. *Prog Neurobiol, 219*, Article 102351. https://doi.org/10.1016/j.pneurobio.2022.102351

Yokoi, A., & Diedrichsen, J. (2019). Neural Organization of Hierarchical Motor Sequence Representations in the Human Neocortex. *Neuron, 103*(6), 1178–1190 e1177. https://doi.org/10.1016/j.neuron.2019.06.017

Zacks, J. M., Braver, T. S., Sheridan, M. A., Donaldson, D. I., Snyder, A. Z., Ollinger, J. M., & Raichle, M. E. (2001). Human brain activity time-locked to perceptual event boundaries. *Nature Neuroscience, 4*(6), 651–655.

Zacks, J. M., Speer, N. K., Swallow, K. M., Braver, T. S., & Reynolds, J. R. (2007). Event perception: A mind-brain perspective. *Psychol Bull, 133*(2), 273–293. https://doi.org/10.1037/0033-2909.133.2.273

Zheng, J., Schjetnan, A. G. P., Yebra, M., Gomes, B. A., Mosher, C. P., Kalia, S. K., & Rutishauser, U. (2022). Neurons detect cognitive boundaries to structure episodic memories in humans. *Nat Neurosci, 25*(3), 358–368. https://doi.org/10.1038/s41593-022-01020-w

Zuo, S., Wang, L., Shin, J. H., Cai, Y., Zhang, B., Lee, S. W., & Kwok, S. C. (2020). Behavioral evidence for memory replay of video episodes in the macaque. *Elife, 9*. https://doi.org/10.7554/eLife.54519