



OPEN ACCESS

EDITED BY

Benjamin Thompson,
University of Waterloo, Canada

REVIEWED BY

Eva Wittenberg,
Central European University, Hungary
Alexis Wellwood,
University of Southern California,
United States

*CORRESPONDENCE

Xinchi Yu
✉ xcyu@umd.edu

RECEIVED 25 August 2023

ACCEPTED 22 December 2023

PUBLISHED 24 January 2024

CITATION

Yu X, Li J, Zhu H, Tian X and Lau E (2024)
Electrophysiological hallmarks for event
relations and event roles in working
memory.
Front. Neurosci. 17:1282869.
doi: 10.3389/fnins.2023.1282869

COPYRIGHT

© 2024 Yu, Li, Zhu, Tian and Lau. This is an
open-access article distributed under the
terms of the [Creative Commons Attribution
License \(CC BY\)](https://creativecommons.org/licenses/by/4.0/). The use, distribution or
reproduction in other forums is permitted,
provided the original author(s) and the
copyright owner(s) are credited and that the
original publication in this journal is cited, in
accordance with accepted academic
practice. No use, distribution or reproduction
is permitted which does not comply with
these terms.

Electrophysiological hallmarks for event relations and event roles in working memory

Xinchi Yu^{1,2*}, Jialu Li^{3,4,5}, Hao Zhu^{3,4,5}, Xing Tian^{3,4,5} and
Ellen Lau^{1,2}

¹Program of Neuroscience and Cognitive Science, University of Maryland, College Park, MD, United States, ²Department of Linguistics, University of Maryland, College Park, MD, United States, ³Division of Arts and Sciences, New York University Shanghai, Shanghai, China, ⁴Shanghai Key Laboratory of Brain Functional Genomics (Ministry of Education), School of Psychology and Cognitive Science, East China Normal University, Shanghai, China, ⁵NYU-ECNU Institute of Brain and Cognitive Science at NYU Shanghai, Shanghai, China

The ability to maintain events (i.e., interactions between/among objects) in working memory is crucial for our everyday cognition, yet the format of this representation is poorly understood. The current ERP study was designed to answer two questions: How is maintaining events (e.g., the tiger hit the lion) neurally different from maintaining item coordinations (e.g., the tiger and the lion)? That is, how is the event relation (present in events but not coordinations) represented? And how is the agent, or initiator of the event encoded differently from the patient, or receiver of the event during maintenance? We used a novel picture-sentence match-across-delay approach in which the working memory representation was “pinged” during the delay, replicated across two ERP experiments with Chinese and English materials. We found that maintenance of events elicited a long-lasting late sustained difference in posterior-occipital electrodes relative to non-events. This effect resembled the negative slow wave reported in previous studies of working memory, suggesting that the maintenance of events in working memory may impose a higher cost compared to coordinations. Although we did not observe significant ERP differences associated with pinging the agent vs. the patient during the delay, we did find that the ping appeared to dampen the ongoing sustained difference, suggesting a shift from sustained activity to activity silent mechanisms. These results suggest a new method by which ERPs can be used to elucidate the format of neural representation for events in working memory.

KEYWORDS

event role, relation, working memory, pinging, object files, object indexicals

1 Introduction

The ability to represent events – not only single objects, but also the interactions between/among them – in working memory is crucial for our everyday cognition. It is clear that we can maintain *ad hoc* events in working memory: for example, if we see a lion hit an elephant, we are able to run subsequent mental computations that make use of the

relations that compose the event representation, even after the hitting itself is over (e.g., inferring that the elephant is in pain or answering a question about who was hit). However, the implementational *format* of events in working memory remains elusive. In the current study, we introduce an ERP (event-related potential) picture – phrase matching paradigm that provides a novel measure of event representation in working memory.

In the following introduction, we first provide an overview of the “neural object indexical theory” (Section 1.1), a theoretical framework for the working memory representation of multiple objects which forms the background for subsequent questions about the neural representation of event relations between objects. We also review previous research on the negative slow wave (NSW), an ERP marker for the number of indexicals in working memory which will be one of the primary components of interest in our current experiment. Then we overview the existing bodies of literature on the neural representation of event relations and event roles (Section 1.2), which together support the research questions of the current paper.

1.1 The neural object indexical theory

1.1.1 Evidence for the neural object indexical theory

The question of how event relations between objects are represented “on the fly” has become more tractable after major advances in the last 30 years in understanding how individual objects are represented and related to the local spatial context in working memory. This line of work suggests that the representation of multiple objects in working memory is supported by a limited set of indexicals or pointers also known as object files (Pylyshyn, 1989, 2001; Kahneman et al., 1992; Xu and Chun, 2007; Carey, 2009; Xie and Zhang, 2017; Brody, 2020; Zhu et al., 2020; Quilty-Dunn and Green, 2021; Quilty-Dunn et al., 2022; Thyer et al., 2022; for a review see Yu and Lau, 2023). The number of discrete object indexicals is limited, classically suggested to have a limit of around 4 (Luck and Vogel, 1997, 2013; Cowan, 2001; Awh et al., 2007; Zhang and Luck, 2008; Xu and Chun, 2009; Ngiam et al., 2022). Various features can be attached to these indexicals, enabling us to tell apart objects with some of the same features, and to distinguish the existence of multiple objects even when their features are identical other than spatial location (Leslie et al., 1998; Treisman, 1998). For example, if two otherwise identical flowers appear at different locations in one’s visual field, typically one is still able to distinguish this from a one-flower situation and to continue to represent the existence of both flowers in the scene, even when direct visual information is interrupted by an occluding screen. This can be understood by positing the maintenance of two object indexicals in working memory, each attaching to its corresponding set of features (in this example, the two sets of features are the same except for location). In contrast, neuropsychological patients with the disorder of simultanagnosia struggle at representing two objects but not one (Coslett and Saffran, 1991; Friedman-Hill et al., 1995; Rafal, 2001), suggesting deficits related to maintaining these object indexicals.

Neuroimaging studies have suggested that these object indexicals are hosted in subregions of the posterior parietal cortex,

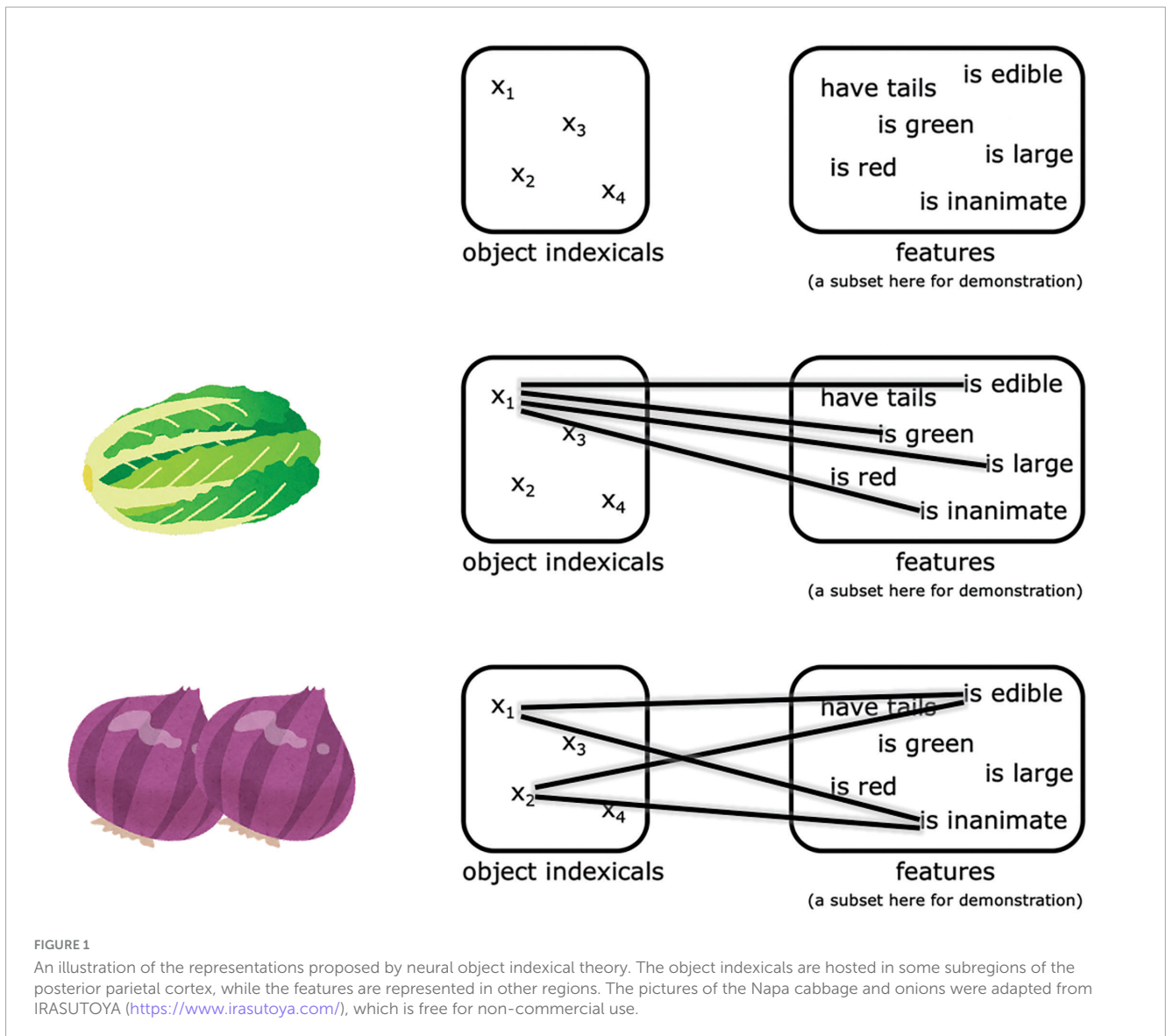
what we will term here the neural object indexical theory (Xu and Chun, 2007, 2009; Balaban et al., 2019). This proposal is based on the observation that activity in some subregions of the posterior parietal cortex reach a plateau beyond 3–4 items (which is the classical working memory capacity), as measured by fMRI (Todd and Marois, 2004; Song and Jiang, 2006; Xu and Chun, 2006; Mitchell and Cusack, 2008; Matsuyoshi et al., 2010; Robitaille et al., 2010; Knops et al., 2014) and MEG (Robitaille et al., 2010). Consistent with the object indexical theory, evidence from neural studies suggest that these posterior parietal representations do not encode object features themselves (Xu and Chun, 2006; Naughtin et al., 2016); rather, these indexicals connect to features represented in other regions, as illustrated in **Figure 1**¹ (Xu, 2007; Naughtin et al., 2016). This theory is also consistent with the proposal that the posterior parietal cortex is crucial to working memory in a more general sense (Bettencourt and Xu, 2016; Jia et al., 2021; for a review see Xu, 2017).

Most visual working memory researchers today assume representations that function like object indexicals, because extreme versions of object-based-all-or-none or feature-only theories for visual working memory cannot account for the behavioral phenomena (for a review see Ngiam, 2023). However, lively debate continues about whether the number of these indexicals is fixed or flexible based on the amount of content that needs to be encoded on each (Brady and Alvarez, 2015; Ngiam et al., 2019), as well as the representational format of features, e.g., whether the features are represented separately or (at least some of them) can be represented conjunctively (Swan and Wyble, 2014; Matthey et al., 2015; Schneegans and Bays, 2017; Hedayati et al., 2022; see Schneegans and Bays, 2019 for a review). While these questions are important, for the purposes of our current investigation of the encoding of event relations between object indexicals, the important shared assumption across all these frameworks is just that increasing the number of indexicals to be represented results in a corresponding increase in processing cost across the maintenance period.

1.1.2 Negative slow wave (NSW): a potential indicator for the object indexical system in working memory

Our methodology for investigating events will center around an ERP component called the negative slow wave (NSW), which is thought to reflect working memory representations related to the object indexical system. Many previous ERP studies of visual working memory have observed sustained negative differences tracking working memory load for multiple objects by using a visual hemifield design that yields a response known as the contralateral delay activity (CDA) (Vogel and Machizawa, 2004). In CDA studies, multiple visual objects are presented on both sides

¹ Note that there are many types of features. For example, one can make the distinction between perceptual (e.g., color, shape) and conceptual (e.g., occupation, edibility) features. It is as yet unclear what type(s) of features can be bound to object indexicals, or that all types of features can be bound to object indexicals. It is also debated whether perceptual and conceptual features are bound to the same set of object indexicals or separate sets of object indexicals (see Brody, 2020; Murez et al., 2020; Yu and Lau, 2023). In the current paper we assume that they are bound to the same set of indexicals: increasing the number of conceptual objects will elicit a similar ERP effect to increasing the number of visual objects (see Section 3.1).



of the visual field, but participants are asked to attend to only one side on any given trial. By hypothesis, ERPs from electrodes contralateral to the attended side will more strongly reflect the working memory representation, so corresponding posterior-occipital electrodes from each side of the scalp are subtracted from each other to compute the contralateral delay activity measure. Many studies have shown that the amplitude of the resulting CDA shows a sustained negativity that is systematically increased with an increasing number of objects to be maintained, up to the individual capacity limit of 3–4 (for an extensive review see Luria et al., 2016). Therefore the CDA has been considered as a potential hallmark for the object indexical system (e.g., Gao et al., 2013; Hakim et al., 2019; Quilty-Dunn and Green, 2021; Cai et al., 2022).

Because our current study was designed to focus on conceptual rather than visuo-spatial representation of events (i.e., the representation shared across linguistic and visual processes), we cannot expect a systematic relationship between working memory representation and visual hemifield. However, another line of visual working memory studies has demonstrated that modulation of the sustained ERP amplitude by number of maintained items can

be observed even without manipulating visual hemifield. Since the measure in these studies does not depend on contralateral-ipsilateral subtraction, the response is called instead the negative slow wave. The NSW has been observed in posterior-occipital electrodes with a similar time course to the CDA (Ruchkin et al., 1992; Klaver et al., 1999; Fukuda et al., 2015; Diaz et al., 2021; for a review see Feldmann-Wüstefeld, 2021).² In visual delay-match-to-sample experiments where the sample objects were presented centrally or bilaterally, the NSW has been found to increase in amplitude (i.e., become more negative) as the number of items to be retained increases (Ruchkin et al., 1992; Mecklinger and Pfeifer, 1996; Klaver et al., 1999; Fukuda et al., 2015; Diaz et al., 2021;

² Note that the exact duration of the NSW effect is under debate; some reports suggested that it could sustain for > 3 s (Ruchkin et al., 1992), while some other reports suggested that it terminates ~ 1 s after memorandum onset (Fukuda et al., 2015). The duration of the CDA is unclear as well (e.g., Perez and Vogel, 2012; Li and Noguchi, 2022). Here we simply note that the NSW and the CDA have broadly similar time courses in the sense that they are both late sustaining components.

Feldmann-Wüstefeld, 2021), and reaches a plateau beyond 3–4 items (Fukuda et al., 2015; Zhou and Thomas, 2015; Feldmann-Wüstefeld, 2021), just like the CDA.³ Therefore, here we use the magnitude of the NSW as an index of working memory load related to the object indexical system.

1.2 What's different about events? Event relations and event roles

What needs to be represented when representing events involving two objects (i.e., participants, entities; for example, “the lion hit the elephant”) in working memory, compared to simply representing two individual objects together (e.g., “the lion and the elephant”)? One piece is simply representing that the two individuals are participating in a particular event of hitting, what we will term here “event relations.” The other piece is representing the particular roles that the individuals are playing in the event (i.e., agents and patients), which we will term here “event roles.” In the current article, our major interest lies in these *conceptual* representations that are shared across visual and linguistic processes (cf. Jackendoff, 1987; Wurm and Caramazza, 2019; Fitch, 2020; Mahon and Kemmerer, 2020). That is, we are interested in the representation that is shared both when one is viewing the scene where a lion hits an elephant, and when one reads a sentence like “the lion hit the elephant.”

1.2.1 The neural representation of event relations

By event relation, we refer to an intuitive distinction between situations where an elephant hits/points at/pushes a lion vs. an elephant and a lion simply standing beside each other and not doing anything to each other (see Figure 2). The difference between these two kinds of situations is whether there is an event relating the two objects. Of course, there are also different *types* of event relations (i.e., event types), e.g., hitting, pointing at, pushing; but the focus of the current article is how our neural system responds to the existence of an event relation at all in working memory, and not how we represent different event types.

Inferring event relations. Event relations can be inferred from both visual and linguistic stimuli. For example, we can use visual information to determine whether there is an event relation for each situation in Figure 2; we can also use linguistic cues to determine whether an event relation is described by phrases like “the lion and the elephant” vs. “the lion hit the elephant.”

Although many visual perception studies have addressed the speed of extracting the event type from a visual image (Dobel et al., 2007b; Hafri et al., 2013; de la Rosa et al., 2014), less is known about how fast the brain detects whether the image depicts a situation including an event relation at all; the only study we are aware of is Hultén et al. (2014). Using MEG, Hultén and colleagues observed

that visual events (e.g., where a lion hits an elephant), compared to visual coordinations (e.g., where a lion and an elephant were simply standing beside each other), triggered an early (~200 ms) difference in left occipital regions. However, based on its location (occipital regions) the effect observed in Hultén et al. (2014) is likely visual rather than conceptual in nature.

Similarly, although many neural studies of language have investigated the representation of event types and event argument structures, examining the contrast between phrases that describe events and those that don't has been less common. Gaston (2020, Experiment 4) investigated the difference between linguistic stimuli including event relations compared to coordination using MEG. Gaston examined the brain responses to serially presented, nonsensical coordinate phrases vs. verb phrases (e.g., the | toasty | tractors | and | the | scenic | cathedrals vs. the | toasty | tractors | entered | the | scenic | cathedrals) and observed a difference from ~150 ms to ~550 ms after the onset of “and” vs. the verb in regions including the transverse temporal sulcus, pSTS (posterior superior temporal sulcus), STS, and ATL (anterior temporal lobe). This effect is in line with the observation of Matchin et al. (2019a), that the presentation of a verb within a sentence elicits a difference starting from ~450 ms upon onset in the angular gyrus, compared to a baseline consisting of phrases without a verb. Such transient effects may reflect a difference in linguistic structure building and/or the representation of event relations at the conceptual level. We should note, however, that the analysis pipelines of Matchin et al. (2019a) and Gaston (2020) may not be sensitive to long-lasting sustained effects.

The maintenance of event relations in working memory. To our knowledge, only two behavioral studies have directly addressed this question (Clevenger and Hummel, 2014; Shen et al., 2021). The results of Clevenger and Hummel (2014) suggested that more relations impose higher working memory load, and the results of Shen et al. (2021) suggested that only two event relations can be maintained in working memory. However, these results remain inconclusive in terms of the *format* of the representation of event relations in working memory; specifically, how are event relations represented with respect to the object indexical system?

One way to approach this question, and the one we pursue in the current study, is to compare the maintenance across a delay period of events like “the tiger hit the elephant” and the maintenance of non-event coordinations like “the tiger and the elephant.” To our knowledge, the MEG study from Hultén et al. (2014) is the only previous neurophysiological study to have examined the maintenance of events compared to coordinations.⁴ In their study, pictures of events and coordinations were presented for 1,500 ms, followed by a speech production task. A numerical difference was observed between events and coordination conditions in a somewhat late time window (from ~300 ms to at least 1 s) time-locked to picture onset in the posterior parietal cortex. Although this difference did not survive statistical

³ In contrast to these studies on visual objects, the results of Morrison et al. (2020) suggest that holding 1, 3, 5 visually-presented *digits* does not drive a sustained parietal effect, at least from visual inspection of their figures, although the electrodes selected in that study (P3, Pz, P4) were less occipital compared to other NSW studies. Digits may encourage phonological rehearsal, whose load is manifested more in frontal and/or central electrodes (Alunni-Menichini et al., 2014; Lefebvre and Jolicoeur, 2016; Simal and Jolicoeur, 2020; Bidelman et al., 2021).

⁴ Note that there have been many fMRI studies comparing events (using pictorial or sentential stimuli) against various types of baselines (e.g., Centelles et al., 2011; Quadflieg et al., 2015; Isik et al., 2017; Okruszek et al., 2018; Matchin et al., 2019b), but due to the relatively poor temporal resolution of fMRI, it is hard to disentangle the extraction and maintenance phases, while the maintenance phase is of our current interest.

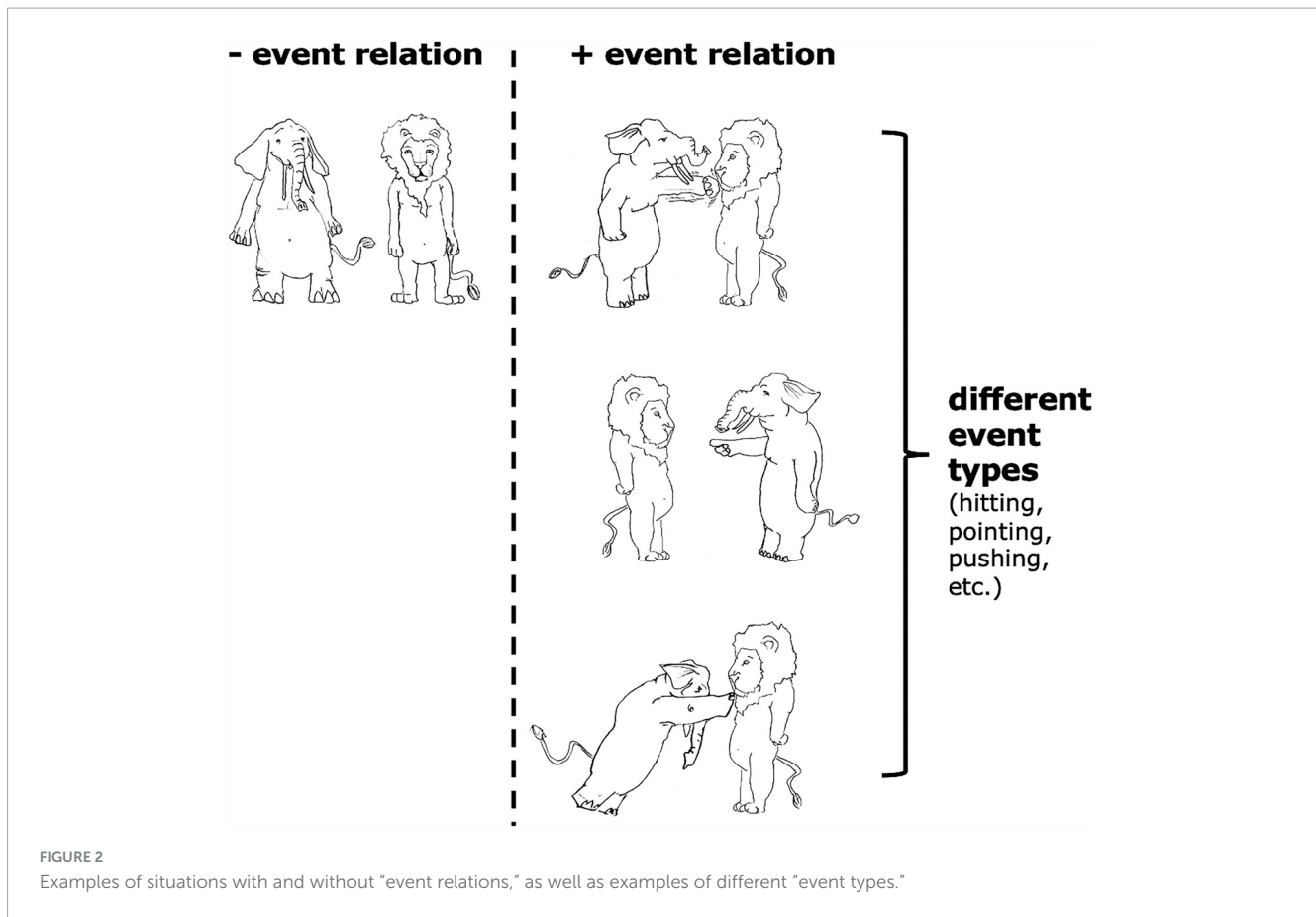


FIGURE 2

Examples of situations with and without “event relations,” as well as examples of different “event types.”

significance in their analysis, this could be due to the small cohort of subjects ($N = 10$).

How should we expect event representation in working memory to manifest differently from simply representing a pair of objects? Through the lens of the object indexical system, we think that there are at least two major possibilities (see Figure 3). One possibility (Figure 3a) is that the event relation may be represented with a separate object indexical from the participants (cf. Clewenger and Hummel, 2014; Shen et al., 2021). In other words, adding an event relation increases working memory cost in exactly the same way as adding another object, and counts against the same capacity limit as objects. For example, in the case of “the tiger hit the elephant,” apart from the two object indexicals for the tiger and the elephant, there is a separate object indexical representing the hitting relation, possibly by connecting to the concept HITTING and other features of hitting.⁵ In this case, there will be an NSW-like difference between the maintenance of events and coordinations. Another possibility (Figure 3b) is that the object indexicals in working memory are true to their name and stand in only for objects, not event relations between objects, as implied for example in Altmann and Ekves’ (2019) theory. The relational event encoding might then be accomplished

by other posterior parietal circuits, or by other brain regions altogether, such as the hippocampus (e.g., Konkel and Cohen, 2009; Eichenbaum and Cohen, 2014). The object indexical working memory representation might only code such event relations implicitly and indirectly in an object-centered fashion, for example by connecting agent and patient “features” to the object indexicals (for further explanation of agents and patients see Section 1.2.2 below). In this case, there is no reason to predict that there will be an NSW-like difference between the maintenance of events and coordinations.

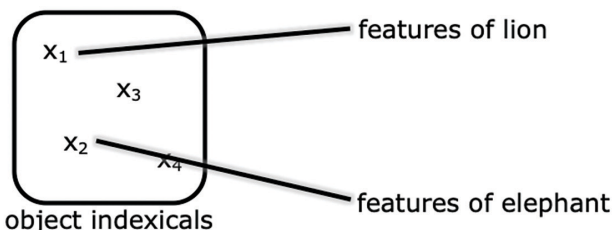
1.2.2 The neural representation of event roles

Many modern theories assume the existence of generalized roles like agent (the initiator of the event) and patient (the receiver of the event), which characterize the involvement of participants across many different types of events (see Strickland, 2017; Rissman and Majid, 2019; Ünal et al., 2021; Wilson et al., 2022). These roles may be understood to be encoded at the linguistic semantic level or at the non-linguistic conceptual level (see Williams, 2015, 2021 for extensive discussion). How such event roles (also known as “semantic” roles or participant roles) are neurally coded in working memory remains unclear.

Inferring event roles. Conceptual event roles can be inferred from visual and linguistic stimuli. In visual behavioral studies (i.e., studies using visual stimuli), it has been observed that the event role of objects can be inferred after only a brief visual presentation (Dobel et al., 2007b; Hafri et al., 2013). In an electroencephalography (EEG) study by Cohn et al. (2017), subjects

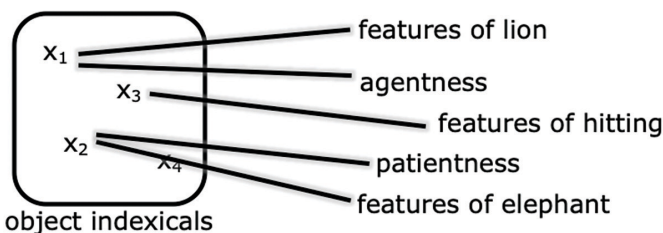
⁵ Note that, just as the posterior parietal indexicals for objects are not thought to represent the features of the objects themselves but to connect to those features represented elsewhere, according to this hypothesis the indexicals for event relations connect to relevant event features and knowledge elsewhere.

the lion and the elephant: 2 object indexicals



the lion hit the elephant

Possibility (a): 3 object indexicals



Possibility (b): 2 object indexicals

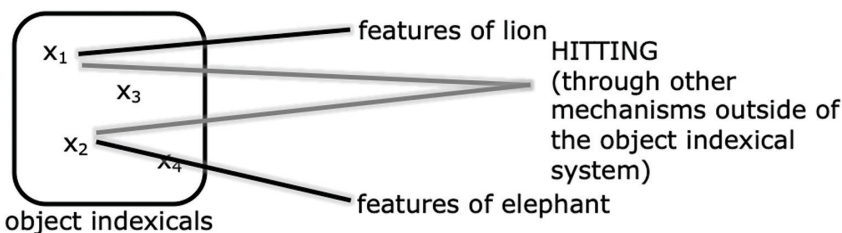


FIGURE 3

An illustration of the two possibilities of how event relations are represented. Possibility a: the event relation is encoded by another object indexical. Possibility b: the event relation is represented by other mechanisms in, e.g., other regions of the posterior parietal cortex or the hippocampus.

were presented with comic strips (modified from *the Peanuts*) sequentially, and the ERP responses to frames illustrating the to-be-agent and to-be-patient character were compared. Of the to-be-agent characters, some were making actions suggesting their agentness (“preparatory agents”), while some were not (“non-preparatory agents”). Relative to the patient condition and the non-preparatory agent condition, the preparatory agent condition elicited a left anterior negativity in 300–600 ms and a slightly leftward fronto-central positivity in 500–900 ms.

One way that languages can convey event role assignments is through linguistic case-marking. For example, in Japanese, the conceptual content of *John gave Mary a book* can be realized as “John-NOM Mary-DAT book-ACC gave,” where -NOM, -DAT, -ACC are nominative, dative and accusative case markers *ga*, *ni*, and *o*. Languages can also convey event roles through linear order as well as other structural cues, as in the English active/passive alternation “The cat chased the dog” and “The dog was chased by the cat,” where linear order of the arguments alternate but the event roles stay constant. Prior fMRI work has reported some success in decoding the binding between event roles and objects from the neural response to these kinds of alternations in English (e.g., Frankland and Greene, 2015, 2020; Wang et al., 2016; Lalisie and Smolensky, 2021). However, given the poor temporal resolution of

fMRI, these studies were not aimed at disentangling the extraction and maintenance processes of event roles.

The maintenance of event roles. A number of behavioral studies have investigated what factors govern accurate maintenance of event roles across time. At least one study suggests that maintaining more event role-object bindings reduces the accuracy of subsequent memory performance (Jessop and Chang, 2020), although it is unclear whether this effect was driven by an increase in the number of objects held in working memory or an increase in the number of event roles bound to these objects (or both). A subsequent study has shown that representing conflicting event roles (e.g., being the agent in one event but the patient in another event) compromises memory performance (Jessop and Chang, 2022). Peng et al. (2018), through a behavioral experiment using the dense sampling method, suggested that event roles are represented at (or even by) different phases of the alpha band (~ 10 Hz) rhythm; we discuss this possibility in further detail in the Discussion.

To our knowledge, no prior electrophysiological study has examined the maintenance of event roles in working memory. In fact, it is a practical challenge to disentangle the separate representation of the agent and the patient in working memory, given that they participate in the same event. In the current study we introduce an exploratory paradigm adapted from the

“pinging” paradigm used in recent working memory studies (Wolff et al., 2015, 2017, 2019, 2020; Fornaciai and Park, 2020; ten Oever et al., 2020; Huang et al., 2021; Duncan et al., 2023; Fan and Luo, 2023). Wolff et al. (2015) conducted a standard delay-match-to-sample working memory task in which the match in the orientation of two Gabor patches needed to be evaluated across a relatively long delay. During the delay period, a high-contrast visual bull’s-eye-shaped stimulus with no orientation information (i.e., the ping) was presented in an attempt to probe the encoding of the orientation information being maintained in memory, similar to the idea of sonar or echolocation. It was found that the otherwise undecodable memory content at that time (the orientation of the sample Gabor patch in this case) was decodable if the ping was presented. In this study and others that followed, the main idea was to probe different working memory contents during retention with the same impulse perturbation (i.e., the ping).

In the current study, we adapt the pinging paradigm to probe different parts of the same working memory content during retention using different pings (see also Fan et al., 2021). While event representations are being maintained in working memory across the delay period we “ping” the event role with the name of the entity to which it was bound, as illustrated in Figure 4. We ping the agent in half of the event trials and the patient in the other half, which gives us the opportunity to compare the ERP corresponding to an impulse perturbation of agents vs. patients in working memory. Notably, since each animal can be the agent or patient, the agent pings and patient pings are perceptually the same across trials. In order to keep the trial structure consistent throughout the experiment, we also pinged one of the two animals in the coordination condition, even though in this case, absent event roles, both entities have the same status in working memory.

Although little is known thus far about how event roles are neurally encoded, one possibility that we explored in this study is that event roles are represented as a kind of “magnitude”.⁶ Behaviorally, agent-patient event roles have sometimes been observed to interact with spatial position (i.e., left-right), which can be treated as a “magnitude” (e.g., Chatterjee et al., 1995; Maass and Russo, 2003; Dobel et al., 2007a; but see Geminiani et al., 1995; Barrett et al., 2002; Altmann et al., 2006; Kazandjian et al., 2011; Dobel et al., 2014 for conflicting results). Some authors have proposed that regions of parietal cortex represent diverse kinds of scalar and quantity information in a common neural implementation format (Walsh, 2003; Buetti and Walsh, 2009; Summerfield et al., 2020; for a critical review see Martin et al., 2017). One kind of neural evidence taken to support this idea is the finding that magnitudes (including numerosity, magnitude of reward, and magnitude of evidence during decision-making) drive an ERP effect in central-parietal electrodes, that is, the central parietal positivity (CPP; Yeung and Sanfey, 2004; O’Connell et al., 2012; Spitzer et al., 2017; Luyckx et al., 2019; for an overview see O’Connell and Kelly, 2021). For transient (i.e., not constantly-changing) stimuli, the CPP is a late component at around 300 to 700 ms post stimulus-onset (Spitzer et al., 2017; Luyckx et al., 2019), with larger magnitudes being more positive. The hypothesis that event roles are represented as magnitudes would predict an ERP effect with a similar scalp distribution to the CPP.

⁶ Other possibilities will be discussed in Section 3.2.

1.2.3 The current study

Our study was designed to investigate neural measures of the maintenance of both event relations and event roles. We ran two experiments with a similar design in two sites, one in Chinese in China (Experiment 1; $N = 19$) and one in English in the US (Experiment 2; $N = 16$). The choice of these two sites/languages was not driven by particular features of the languages, but simply because of practical constraints experienced by the investigators during the COVID-19 pandemic. However, this did allow us to demonstrate that the results extend across two different languages and participant populations. We take posterior-occipital electrodes P7, P8 as the NSW region of interest (ROI) for event relation maintenance, and central-parietal electrodes CP1, CP2, Pz (Experiment 1) and Cz, CPz, Pz (Experiment 2) as the CPP ROI for event role maintenance (see Figure 5).

2 Results

2.1 Experiment 1

2.1.1 Behavioral results

Across the 19 subjects included in the analyses, the mean proportion of correct responses was 96% (range 89–99%). Mean proportion of correct responses was 93% (when the agent was pinged), 95% (when the patient was pinged) and 98% (in coordination trials). Based on the high accuracy, all trials were analyzed for the EEG analysis.

The mean (\pm SD) reaction time of correct responses was $1,241 \pm 288$ ms when the agent was pinged, $1,230 \pm 275$ ms when the patient was pinged, $1,236 \pm 278$ ms for all event trials and 903 ± 233 ms for all coordination trials. The reaction times reported here were first averaged within subjects for all the correct trials, then averaged across subjects. Paired t -test (two-tailed) indicated no significant difference between the reaction times for agent ping trials and patient ping trials, $t(18) = 0.6$, $p = 0.56$. Paired t -test (two-tailed) revealed a significant difference between the reaction times for event trials and coordination trials, $t(18) = 11.7$, $p < 0.001$. It has been observed with the delay-match-to-sample paradigm that the more objects one needs to maintain in working memory, the slower one reacts to the probe stimuli (Hyun et al., 2009; Mitchell and Cusack, 2011; Park et al., 2017). Therefore, the longer reaction time to the probe (in our experiment the linguistic expression) for event trials observed in our current experiment is in line with our interpretation of our EEG results, that the maintenance of events imposes a higher working memory load compared to the maintenance of coordinations, similar to representing an extra object (see Section 3).

2.1.2 EEG results

Event-related potential responses to event pictures and coordination pictures are illustrated in Figure 6. We expected initial evoked responses to the event pictures and coordination pictures to differ because the pictures in the two conditions were physically different across various dimensions, and that responses to the sentences in the two conditions might differ because the form of the correct response differed in the two conditions. Our

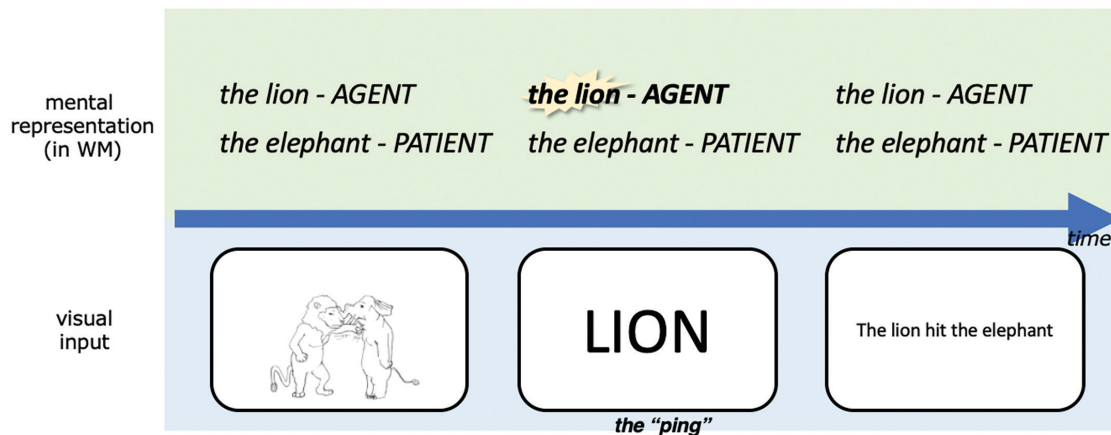


FIGURE 4

A simplified schematic illustration of our experimental design, highlighting our adapted pinging paradigm. The task was to respond to whether the linguistic expression matched with the picture in this trial upon presentation of the linguistic expression, this was designed to encourage subjects to represent the event *conceptually* during the retention period. WM, working memory. For detailed parameters of the experiment, see Materials and methods.

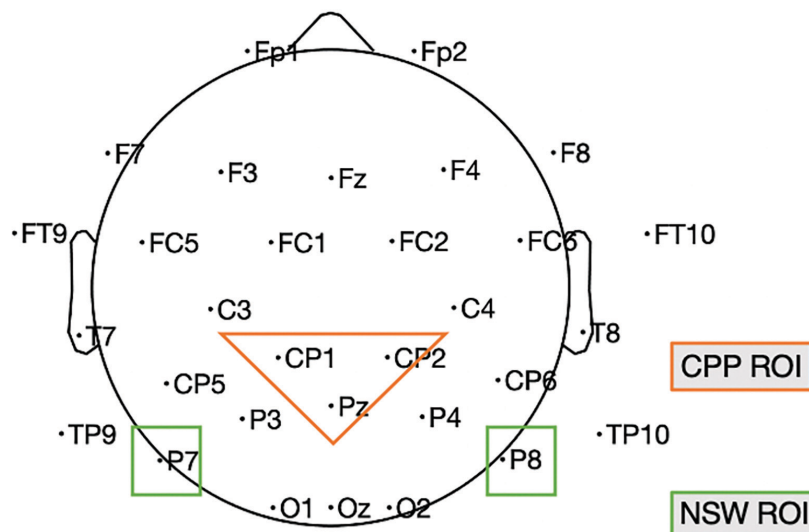


FIGURE 5

Illustration of the electrodes of our two ROIs in Experiment 1; in Experiment 2 the NSW ROI included Cz, CPz, and Pz because of a different layout. CPP, centro-parietal positivity; NSW, negative slow wave; ROI, region of interest. This is based on the standard 10–20 map provided by the DIPFIT plug-in of EEGLAB.

question of interest was whether we would observe differences across the delay between the picture presentation and the sentence presentation in the posterior electrodes associated with the negative slow wave.

A temporal cluster-based permutation test (Maris and Oostenveld, 2007) was performed on the responses to the event picture and the coordination picture in the time window of $-200:3500$ ms for the NSW ROI using customized MATLAB scripts. First, potential clusters were identified with point-to-point paired t -test at a threshold of $\alpha = 0.05$ (two-tailed). Then, the t values of these potential clusters were compared against a permuted (permutations = 9,999) distribution of t values, with a threshold of $\alpha = 0.05$ (two-tailed). A cluster was identified in the negative slow wave (NSW) ROI, extending from 438 to 2,469 ms after

picture onset ($p = 1 \times 10^{-4}$, Figure 6). The difference manifested as an increased negativity for the event pictures relative to the coordination pictures in these posterior electrodes. Although our analyses focused on the pre-determined NSW ROI, the scalp topography of the cluster suggests that this posterior negativity may have been accompanied by a corresponding (central-)frontal positivity.

Event-related potential responses to the delay-period agent ping and the patient ping in the event condition are illustrated in Figure 6. Cluster-based permutation test (Maris and Oostenveld, 2007) was performed comparing the agent ping vs. patient ping conditions in the time window of $-200:1800$ ms using customized MATLAB scripts using the same approach described above for the event picture comparison. No significant clusters distinguishing the

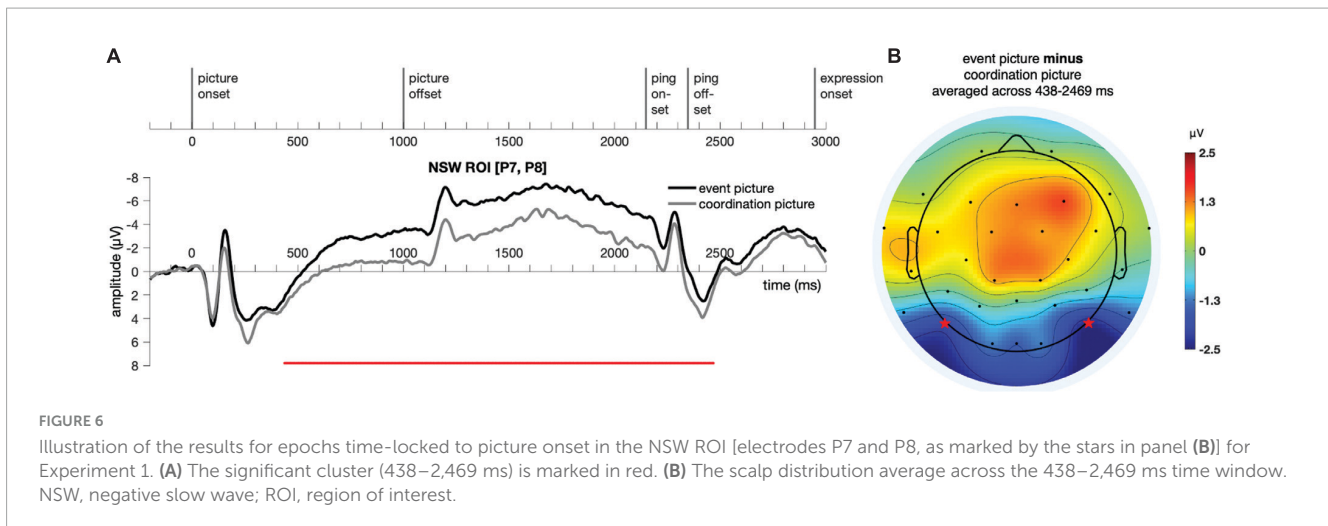


FIGURE 6 Illustration of the results for epochs time-locked to picture onset in the NSW ROI [electrodes P7 and P8, as marked by the stars in panel (B)] for Experiment 1. (A) The significant cluster (438–2,469 ms) is marked in red. (B) The scalp distribution average across the 438–2,469 ms time window. NSW, negative slow wave; ROI, region of interest.

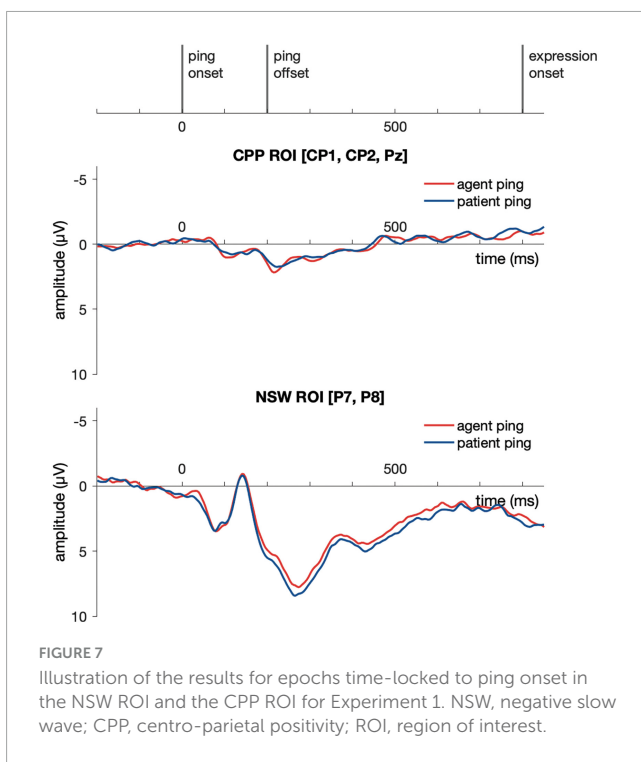


FIGURE 7 Illustration of the results for epochs time-locked to ping onset in the NSW ROI and the CPP ROI for Experiment 1. NSW, negative slow wave; CPP, centro-parietal positivity; ROI, region of interest.

agent ping and patient ping were identified in the CPP ROI, nor in the NSW ROI (Figure 7).

2.1.3 Rationale for experiment 2

In Experiment 1, we observed that maintaining events in working memory, compared to maintaining coordinations, elicited a sustained NSW (negative slow wave) effect. However, during data analysis we discovered an issue in the accuracy of the timing parameters that could have led us to overestimate the duration of the sustained negativity (see Section 4.1.6 for more detail). In Experiment 2, we aimed to replicate the effect in a different population (Experiment 1 was run in Shanghai with stimuli in Chinese and Experiment 2 was run in the US with stimuli in English), with corrected presentation parameters that allow more precise estimation of the duration of the effect.

As visual inspection of the Experiment 1 effects suggested that the presentation of the ping may have acted to reduce or eliminate the sustained negativity, we also took Experiment 2 as an opportunity to evaluate this possibility by presenting the ping somewhat earlier in the trial. We presented the pictures for shorter duration in Experiment 2 (400 ms) compared to Experiment 1 (1,000 ms), as an additional measure to ensure that the sustained effects we observed did indeed reflect delay-period activity rather than continued visual processing.

2.2 Experiment 2

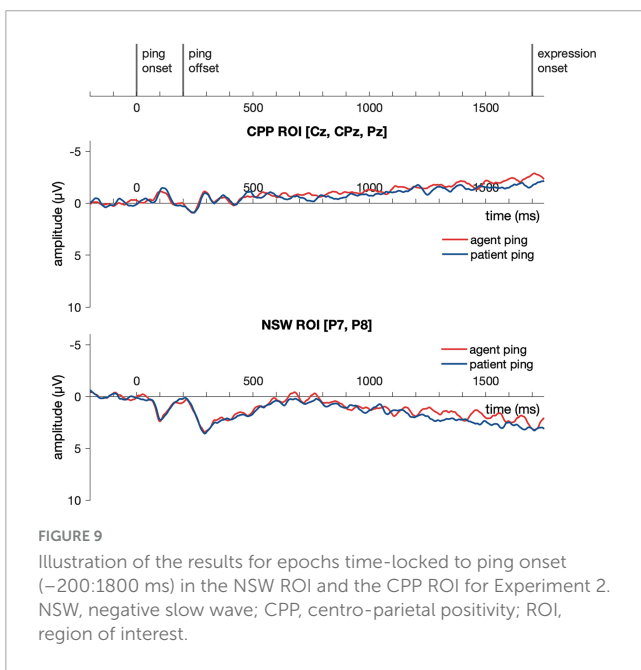
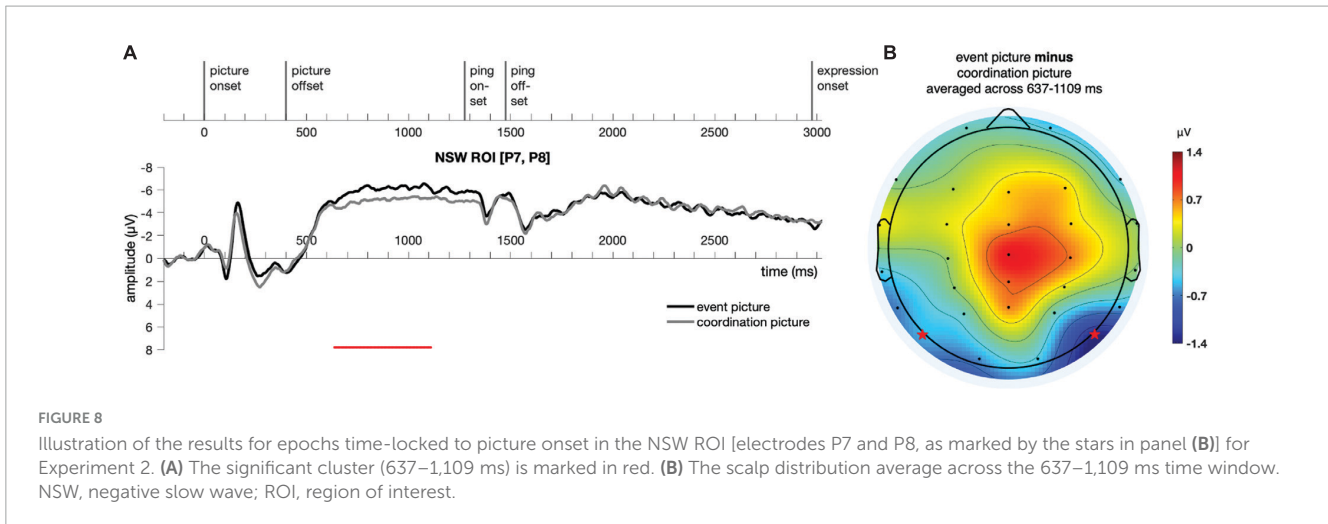
2.2.1 Behavioral results

Across the 16 subjects included in the analyses, the mean proportion of correct responses was 92% (range 83–97%). Mean proportion of correct responses was 88% (when the agent was pinged), 91% (when the patient was pinged) and 94% (in coordination trials). Based on the high accuracy, all trials were analyzed for the EEG analysis.

The mean (\pm SD) reaction time of correct responses was $1,617 \pm 345$ ms when the agent was pinged, $1,616 \pm 357$ ms when the patient was pinged, $1,617 \pm 345$ ms for all event trials and $1,149 \pm 349$ ms for all coordination trials. The reaction times reported here were first averaged within subjects for all the correct trials, then averaged across subjects. Paired *t*-test (two-tailed) indicated no significant difference between the reaction times for agent ping trials and patient ping trials, $t(15) = 0.6$, $p = 0.99$. Paired *t*-test (two-tailed) revealed a significant difference between the reaction times for event trials and coordination trials, $t(15) = 9.7$, $p < 0.001$, similar to the results in Experiment 1.

2.2.2 EEG results

Similar to Experiment 1, a cluster-based permutation test (Maris and Oostenveld, 2007) was performed on the responses to the event picture and the coordination picture in the time window of $-200:3500$ ms for the NSW ROI using customized MATLAB scripts. A cluster was identified in the negative slow wave (NSW) ROI, extending from 637 to 1,109 ms after picture onset ($p = 0.006$, Figure 8). As in Experiment 1, the difference



manifested as an increased negativity for the event pictures relative to the coordination pictures in these posterior electrodes. We note that the amplitude of the sustained effect was somewhat smaller in Experiment 2 than Experiment 1, perhaps because of the shorter picture duration for encoding in Experiment 2 (see Brady et al., 2016, for a similar correlation between encoding duration and effect size). The scalp topography of this NSW effect was generally similar to the one obtained in Experiment 1. The duration of the sustained effect appeared to extend just up through the onset of the ping screen, after which no significant differences were observed.

Event-related potential responses to the delay-period agent ping and the patient ping in the event condition are illustrated in Figure 9. Similar to Experiment 1, a cluster-based permutation test (Maris and Oostenveld, 2007) was performed comparing the agent ping vs. patient ping conditions in the time window of –200:1800 ms using the same approach described above for the event picture comparison. No significant clusters distinguishing the

agent ping and patient ping were identified in the CPP ROI, nor in the NSW ROI.

3 Discussion

Our current experiments constitute one of the first electrophysiological studies directly addressing how event relations and event roles (i.e., agent and patient) are maintained in working memory. We report two key findings. First, across Experiments 1 and 2, we saw that the working memory retention of events (e.g., a tiger hit an elephant) relative to simple coordination (e.g., a tiger and an elephant) was hallmarked by a response similar to a negative slow wave (NSW) in posterior-occipital electrodes during the delay period. Maintenance of events resulted in a more negative sustained response compared to coordinations, suggesting higher working memory load based on the standard interpretations of the negative slow wave in prior studies. Our second finding was that in a novel agent vs. patient “pinging” manipulation during the delay, we observed no reliable difference over centro-parietal positivity electrodes or negative slow wave electrodes for agent as compared to patient. Interestingly, however, the sustained negative slow wave effect seemed to be largely dampened after the presentation of the ping, as observed to track the shift of the ping presentation timing from Experiment 1 to Experiment 2.

3.1 Representing event relations in working memory

The major difference between the event and coordination conditions was in whether or not an event relation needed to be represented across the delay period. The NSW effect in Experiment 1 onset at ~ 450 ms and extended even until ~ 2,500 ms; for Experiment 2, it onset at ~ 600 ms and extended until ~ 1,100 ms (approximately when it was interrupted by the ping). This late, *sustained* effect across the two conditions suggests that event relations *can* be represented by sustained activity in working memory. Our current study, in demonstrating a neural measure sensitive to the maintenance of event relations,

thus marks a crucial first step toward further elucidating the implementation-level question of *how* event relations are represented in working memory.

What brain regions contributed to the NSW effects in our current study? Although future work with source localization in MEG is needed, we believe that posterior parietal cortex is the most likely candidate. The relatively late onset of the sustained effect here is comparable to the trend toward a sustained effect for event representation in Hultén et al.'s (2014) MEG study, which they localized to posterior parietal cortex. Of course, a posterior parietal source would also be consistent with the previous work reviewed in the introduction emphasizing the contribution of this region to working memory (Robitaille et al., 2010; Fan et al., 2012; Becke et al., 2015; Bettencourt and Xu, 2016; Jia et al., 2021) and event representation (Thompson et al., 2007; Centelles et al., 2011; Wang et al., 2016; Williams et al., 2017; Matchin et al., 2019a,b). Moreover, the posterior parietal cortex has been observed to be involved in both linguistic and visual processing, suggesting that it is hosting conceptual representations (Menenti et al., 2012; Jouen et al., 2015; Baldassano et al., 2017; Zadbood et al., 2017; Liu et al., 2020; Ivanova et al., 2021). Previous work (Summerfield et al., 2020) has proposed that the posterior parietal cortex plays a crucial role in representing “relational structures” (see also O'Reilly et al., 2022), yet the format of this representation remains elusive. One possibility consistent with the negative slow wave (NSW) effect during the maintenance of events vs. coordinations is that the event relation in an event is represented by a separate object indexical apart from the two participants. These “relation indexicals” could be qualitatively different from the indexicals for objects; a possibility that could be tested in future experiments. However, given the absence of localization information in the current study, it is also possible that the NSW we observed was driven by other posterior parietal circuits (or even other brain regions). Further work is needed to disentangle these possibilities.

Interestingly, Lapinskaya et al. (2016) observed a late (from 500 ms to at least 800 ms) parietal-occipital ERP effect between object nouns vs. event nouns and verbs after the presentation of the second word in a two-word similarity judgment task. This effect had a similar scalp distribution to our current effect based on visual inspection. The object nouns were more positive over the posterior electrodes compared to verbs and event nouns, which mirrors the current results in which coordinations elicited a less negative response than events. However, Lapinskaya et al. (2016) used a single-word similarity judgment paradigm, which is quite different than the paradigm of our current study, and did not examine responses over an extended maintenance delay.

We should also note that some behavioral studies have shown that when two objects are involved in an event, participants can actually perform better on working memory tasks than when they are maintaining two separate objects (Stahl and Feigenson, 2014; Ding et al., 2017; Lu et al., 2022; Paparella and Papeo, 2022). This seeming contradiction with our current results remains to be resolved in future research. One possibility is that the number of working memory indexicals used to encode the same input may vary according to task environments, as suggested by studies on visual grouping (Balaban and Luria, 2016; Rabbitt et al., 2017; McCants et al., 2020). Another possibility is that different numbers of indexicals may be used during encoding and maintenance

phases, as suggested for visual grouping (Peterson et al., 2015; Thyer et al., 2022).

3.2 Representing event roles in working memory

In this study we introduced a novel “pinging” paradigm to attempt to observe evidence of the differential coding of event roles in working memory across the delay. One hypothesis was that event roles are coded similarly to magnitudes, and that this would be associated with differences in the electrodes previously associated with a centro-parietal positivity electrodes suggested to reflect magnitude coding. We failed to find evidence for this hypothesis, as we did not observe a significant effect of the pinged event role over CPP electrodes across the two experiments. This is not definitive evidence against magnitude coding of event roles, as the CPP may not be a general hallmark for all magnitudes. We also note that much of the existing pinging literature has relied on decoding rather than comparing mean differences. For example, Fornaciai and Park (2020), in a numerosity working memory task, identified a significant ERP difference upon sample presentation across conditions, yet failed to identify significant ERP differences upon ping presentation during the retention period across these conditions. In line with prior pinging literature, they did, however, observe that ping presentation was associated with above-chance decodability. Further research is needed to investigate whether ERP amplitudes are generally insensitive to the perturbation responses elicited by pinging.

However, our failure to find agent-patient differences in the ping paradigm may also reflect the fact that magnitude coding is not the format by which event roles are related to object indexicals in working memory. There are clear theoretical challenges for representing event roles as magnitudes. For example, if we consider tools or instruments (e.g., Mary opened the bottle with *the opener*), it is hard to think of this role as being somewhere on a 1-D scale along the same line of agent and patient. This is different from other magnitudes (e.g., numerosity, magnitude of reward, and magnitude of evidence during decision-making), which can be understood as being represented on a 1-D magnitude scale.

What are alternative encoding mechanisms beyond magnitude that could support event role representation? There are currently at least two other major hypotheses about the implementation of event roles. One is that different event roles, as features, are represented by different neural firing patterns. The other is that agent and patient roles are signaled by oscillatory phase. For example, it could be that when an object indexical is representing the agent object, it is most excited at one phase, while when the same object indexical is representing the patient object, it is most excited at another phase. The brain may make use of both representations for different purposes. Some evidence for the first hypothesis comes from an fMRI decoding study of Wang et al. (2016), which successfully decoded (above chance) from the response in a wide range of brain regions whether a certain animal was the agent or the patient in a visual scene. Versions of the second hypothesis have long been discussed in the computational literature (Shastri and Ajanagadde, 1993; Shastri, 1999, 2002, 2007; Hummel

and Holyoak, 2003; Martin and Dumas, 2017), yet have only recently been tested by Peng et al. (2018). Using the dense sampling method for identifying behavioral oscillations, Peng and colleagues found that the object indexicals for agents and patients are most excited at a ~ 90 degree lag in the alpha band (~ 10 Hz). Results from several other studies are also in line with the possibility that when holding multiple objects (≥ 2) in working memory, different object indexicals are most excited at different oscillatory phases (Jia et al., 2017; Peters et al., 2021; Pomper and Ansorge, 2021; Balestrieri et al., 2022; Chota et al., 2022). Our current results do not provide strong evidence either for or against these hypotheses.

Another dispute over the representation of event roles is whether they are independent or dependent from the event type (see Williams, 2015, 2021). By relation-independent, we refer to abstract roles like agent and patient; by relation-dependent, we refer to action-specific roles such as hitter and hittee, hugger and huggee, etc. On the assumption that roles are relation-dependent, one possibility is that the perception of an independent “agent” role may come from different relation-dependent roles being aligned on some graded dimension of “agentivity” (for example, hitters may be more “agenty” than huggers). While a large body of literature in formal semantics assumes the existence of relation-independent roles (for discussions see Williams, 2021), Dowty (1991) suggested that (proto-)agents and (proto-)patients may rather be characterized by varying numbers of “typical” features, which correspond to graded responses that some prior studies have empirically quantified with subjective ratings (Kako, 2006a,b; Reisinger et al., 2015; Rissman and Lupyan, 2022, for studies with a similar logic using other ratings see Johnson, 1967; Madnani et al., 2010). More research is still needed to resolve this dispute, which in turn will inform whether neural studies should expect to find for consistent “agent” or “patient” representations across event types.

3.3 Sustained activity vs. activity-silent mechanisms

Our current results are consistent with the idea that event relations can be maintained in working memory in the form of sustained neural activity, given that we observe a sustained ERP difference between the retention of events and coordinations. Although sustained activity has long been seen as a key mechanism for working memory representation (for reviews see Leavitt et al., 2017; Wang, 2021), at least two lines of research have argued that working memory retention can also sometimes be implemented by “activity-silent mechanisms” (Lewis-Peacock et al., 2012; Stokes, 2015; Rose et al., 2016; Kamiński and Rutishauser, 2020; Beukers et al., 2021) or other mechanisms that do not induce sustained changes in neural metabolism (e.g., Wutz et al., 2020; Lundqvist et al., 2021).

One line of research has utilized visual stimuli, and observed that the amplitude of CDA after presenting two consecutive memory arrays is not always reflective of the load of the first memory array (Berggren and Eimer, 2016; Feldmann-Wüstefeld et al., 2018; Zhang et al., 2021) and that a disruption in sustained activity is not necessarily accompanied by a huge impairment in memory performance (Kreither et al., 2022). Another line of research has utilized linguistic stimuli, and has observed that the

maintenance of syntactic dependency is not always accompanied by the presence of the sustained component of SAN (sustained anterior negativity) (McKinnon and Osterhout, 1996; Kaan et al., 2000; Lau and Liao, 2018; Yano and Koizumi, 2018, 2021; Lo and Brennan, 2021; Sprouse et al., 2021; Cruz Heredia et al., 2022; Lau and Yang, 2023; for an earlier review see Lau, 2018).

Our current experiment provides one of the first reports of how event relations are neurally represented across time in working memory. The sustained negative slow wave effect across events and coordinations in Experiment 1 and 2 demonstrated that the working memory maintenance of event relations *can* be realized by sustained activity. However, as Experiment 2 confirmed, this sustained NSW effect was largely dampened by the ping, and yet participants continued to demonstrate high accuracy on the behavioral memory task. This is in line with previous findings in visual working memory studies that an interruption during retention dampened CDA (Berggren and Eimer, 2016; Feldmann-Wüstefeld et al., 2018; Zhang et al., 2021), yet resulted in little detriment to behavioral accuracy (Kreither et al., 2022).⁷ Together, these results suggest that the representational format of the memorandum may shift from sustained activity to activity-silent representation upon interruption. Therefore, our observation that the sustained difference was reduced after the ping suggests that although the working memory maintenance of event relations *can* be realized by sustained activity, it can also be represented in an activity silent format.

3.4 Event type and social interactions

In the current study, we focused on how the brain reacts to whether or not there is an event relation at all in working memory, but not how the event type itself (e.g., hitting, kicking) is represented. Previous fMRI studies have converged on the posterior lateral temporal cortex (PLTC, Wurm and Lingnau, 2015; Wurm et al., 2016; Hafri et al., 2017; Walbrin and Koldewyn, 2019; Wurm and Caramazza, 2019) as hosting the representation for event types. Following our interpretation of the negative slow wave effect, it is possible that an indexical for event relations in the PPC (posterior parietal cortex) points to the event type represented in PLTC during working memory maintenance. However, evidence remains mixed on whether the event type representations in PLTC reflect the extraction of event types or the maintenance of event types in working memory, or both (Lu et al., 2016; Zhou et al., 2022).

The NSW effect in our current study could also be interpreted as reflecting social cognition computations in particular, related to the presence of social interaction for events. Future research would be needed to determine whether all event types can drive an NSW effect, or only social ones. Interestingly, some recent studies have suggested that some sub-regions in PLTC may favor the representation of social interactions over non-social ones (Wurm et al., 2017; Wurm and Caramazza, 2019; Yang et al., 2020).

⁷ Notably, the reaction times seemed to be longer for Experiment 2 compared to Experiment 1, despite similar stimulus-onset asynchrony across the picture and the linguistic probe. This is in line with the observation in Zhang et al. (2021): although holding content in the activity-silent mode for a longer time does not always come with a lower behavioral accuracy, it can come with a longer reaction time to the probe.

3.5 Conclusion

The current study provides one of the first attempts to discover ERP hallmarks for the conceptual representation of events vs. coordination and agent vs. patient roles in working memory. We observed a sustained difference during a delay period for maintaining conceptual information about events relative to coordinations in working memory, which resembled the posterior-occipital negative slow wave (NSW) effect observed in previous visual working memory studies. Interestingly, a novel event role “pinging” manipulation also revealed that the NSW effect seemed to be largely dampened after the presentation of the ping.

Much future work still needs to be done to gain a better understanding of these electrophysiological hallmarks observed in our current study, such that they can be used to investigate the format of event representation in working memory. The current study is an exploratory, yet necessary first step.

4 Materials and methods

4.1 Experiment 1

4.1.1 Participants

In total, 24 subjects participated in this experiment; they were balanced in terms of the list used in the main experiment. Of these subjects, data from 5 subjects were excluded: 1 of these 5 subjects was excluded because of incomplete data. Data from the other 4 of the 5 subjects were excluded based on a procedure detailed in Section 4.1.5. The remaining 19 subjects analyzed had an age of 18–27 ($M = 22$, 11 female), all right-handed as measured by a translated version of the Edinburgh Handedness Inventory (Oldfield, 1971). Subjects were recruited from the New York University Shanghai and East China Normal University community. Subjects received monetary reimbursement for their participation. Written consent was acquired from each subject, and the procedures were approved by the Institutional Review Board of the University of Maryland and New York University Shanghai.

4.1.2 Stimuli

A total of 144 “event pictures” about events involving two characters were constructed, adapted from stimuli used in Hultén et al. (2014). The Hultén et al. (2014) stimuli originally consisted of cartoons depicting a single animal conducting a certain action or standing still (side-facing, used as the patients in event pictures, or front-facing, used in coordination pictures; for the definition of coordination pictures see below).

The characters used in the stimuli set were four animals (the elephant 大象, the lion 狮子, the hippo 河马 and the sheep 绵羊, therefore, there are 6 combinations). Each event involved one of 6 distinct actions (hitting 打, pulling 拉拽, kissing 亲, pointing at 指, pushing 推, scolding 骂). From each of the 6 character duos (e.g., the elephant and the lion), each of the 6 actions (e.g., hitting), and each of the 2 ways of allocating event roles (e.g., the elephant hitting the lion vs. the lion hitting the elephant), we constructed 72 unique event pictures. Note that due to the nature of the Hultén et al. (2014) stimuli (the animals performing an action were all right-facing), the agent was always to the left in these 72 “original”

pictures. Therefore, we created another 72 mirror-flipped pictures by mirror-flipping the 72 original pictures to balance the relative visual position of the two characters, thus resulting in 144 event pictures in total.

Another 12 “coordination pictures” were also constructed, with two characters (also within the elephant, the lion, the hippo and the crocodile) standing next to each other. For each animal pair (X,Y), we created separate pictures in which animal X is to the left or animal Y is to the left.

The same event (e.g., lion hitting elephant) was only presented once for each subject. In order to balance relative right/left positions, we distributed the 144 original and mirror-flipped pictures across two lists (list A and B), each containing 72 pictures representing unique events; in each list, a certain character had an equal probability of appearing to the left as agent, to the left as patient, to the right as agent, and to the right as patient. Besides, for each action, the probability of the agent being to the left was the same as the probability of the agent being to the right. Both list A and list B contained an equal amount of original (i.e., 36) and mirror-flipped (i.e., 36) event pictures. Each list also included 72 coordination pictures (6 repetitions of each of the 12 coordination pictures). That is, each of the two lists consisted of 144 pictures in total (72 event pictures and 72 coordination pictures); this corresponds to 144 trials that each subject went through (72 event trials and 72 coordination trials). For more examples of our picture stimuli, see [Supplementary Figure 1](#).

Each trial contained a single word “ping” stimulus during the delay between the picture and the probe linguistic expression. These stimuli were arranged so that within each list, (a) there was an equal probability for the agent on the left, the agent on the right, the patient on the left and the patient on the right to be pinged, (b) for each action, the agent and patient were equally likely to be pinged, (c) for each pinged character, there was an equal possibility that this character was the agent or the patient.

The format of probe linguistic expressions is illustrated in [Table 1](#). In trials where an event picture (e.g., a picture depicting a lion hitting an elephant) had been presented (i.e., event trials), the linguistic expression (here, a sentence) either matched with the event picture (50% probability, e.g., “the lion hit the elephant”/“the elephant was hit by the lion”) or was different in event role assignment compared to the event picture (50% probability, e.g., “the elephant hits the lion”/“the lion was hit by the elephant”). The sentence was determined randomly to be in active or passive form. In trials where a coordination picture (e.g., a picture depicting a lion standing next to an elephant) had been presented (i.e., coordination trials), this linguistic expression (here, a phrase) either matched with the picture (50% probability, e.g., “the lion and the elephant”/“the elephant and the lion”) or was different in character compared to the coordination picture (50% probability, e.g., “the lion and the hippo,” “the elephant and the hippo,” etc.). The order of the two animals was determined randomly.

4.1.3 Procedure

Each trial started with a “+” in the middle of the screen (for 500 ms) signaling the beginning of a trial. This “+” was followed by the presentation of an event picture for 1,000 ms (for example, a picture depicting the lion hitting the elephant) or a coordination picture (for example, a picture depicting the lion standing next to

TABLE 1 Example of linguistic expressions presented in Experiment 1.

Event condition	Active expression	狮子	打了	大象		
	<i>Pinyin</i> transcription	shizi	da-le	daxiang		
	English translation	lion	hit-ASP	elephant		
		“(the/a) lion hit (the/an) elephant”				
	Passive expression	大象	被	狮子	打了	
	<i>Pinyin</i> transcription	daxiang	bei	shizi	da-le	
	English translation	elephant	BEI	lion	hit-ASP	
		“(the/an) elephant was hit by (the/a) lion”				
Coordination condition	Coordination expression	狮子	和	大象		
	<i>Pinyin</i> transcription	shizi	he	daxiang		
	English translation	lion	and	elephant		
		“(the/an) elephant and (the/a) lion”				

the elephant). After 900 ms of a blank screen, a brief “pinging” word in larger font (250 pt) referring to one of the two characters (e.g., “lion” or “elephant”) was presented for 200 ms, in order to highlight the just-encoded character-role pairing maintained in working memory. This ping was followed by a blank screen for 600 ms, resulting in a total SOA (stimulus onset asynchrony) of 800 ms between ping presentation and the subsequent probe linguistic expression (a sentence or a phrase). After the delay period, a linguistic expression (see Table 1) appeared at the center of the screen, which either matched or did not match with the picture. The whole linguistic expression was presented on the same screen.

Subjects were asked to make a response on the keyboard whether the linguistic expression matched with the picture in the same trial as fast and accurately as possible; subjects were allocated into two response hand groups: in one group subjects pressed “z” with the left hand when matching and pressed “m” with the right hand when not; in the other group subjects pressed “m” with the right hand when matching and pressed “z” with the left hand when not. Then a feedback screen displaying whether the choice was right (in green) or wrong (in red) was presented for 400 ms; if the subjects did not make the response within 5 s, the feedback screen would automatically appear, showing “no response” (in white). After the feedback screen, there was an inter-trial interval (ITI) uniformly random across 1,000–2,000 ms before the next trial. Within the ITI, a line of asterisks appeared on the screen; before the experiment, the subjects were encouraged to blink during this period, in order to reduce blinks within trials. Subjects were instructed to always fixate on the center of the screen within a trial. For a schematic illustration of one trial in Experiment 1 see Figure 10.

Before the main experiment, subjects were given 24 practice trials (12 event pictures, 12 coordination pictures) randomly, constructed from two animals not involved in the experiment (which is, the cow 奶牛 and the crocodile 鳄鱼) and all 6 actions. Two balanced lists of event pictures (each containing 12 event pictures) were similarly created. The practice trials were used to familiarize subjects with the verbs used for the actions, as well as the response for coordination sentences, etc. These 24 practice trials were administered repetitively until the subjects reached 100% accuracy on all 24 practice trials. Note that in practice trials, the unmatching linguistic expressions for coordination pictures were constructed by substituting one of the characters (i.e., the cow or

the crocodile) for another animal that was not involved in this experiment (which is, the squirrel 松鼠). This difference in the probe phrases compare to our main experiment was unlikely to affect our results, as this does not change essentially that in the coordination conditions, subjects were representing the two characters as two grouped entities.

In the main experiment, five self-paced breaks were given every 24 trials (there were in total 144 trials). The experiment was run with Psychtoolbox 3 (Brainard, 1997; Kleiner et al., 2007) in MATLAB 2020b. The background of the screen was gray, and the texts were white. The pictures were created in 1,600 × 1,200 px, and were presented on the screen as 24° × 18°. All texts, except for the ping, were presented with a font size of 100 pt; the pings were presented with a font size of 250 pt. Subjects sat at a distance of 80 cm before the screen.

4.1.4 Behavioral data analysis

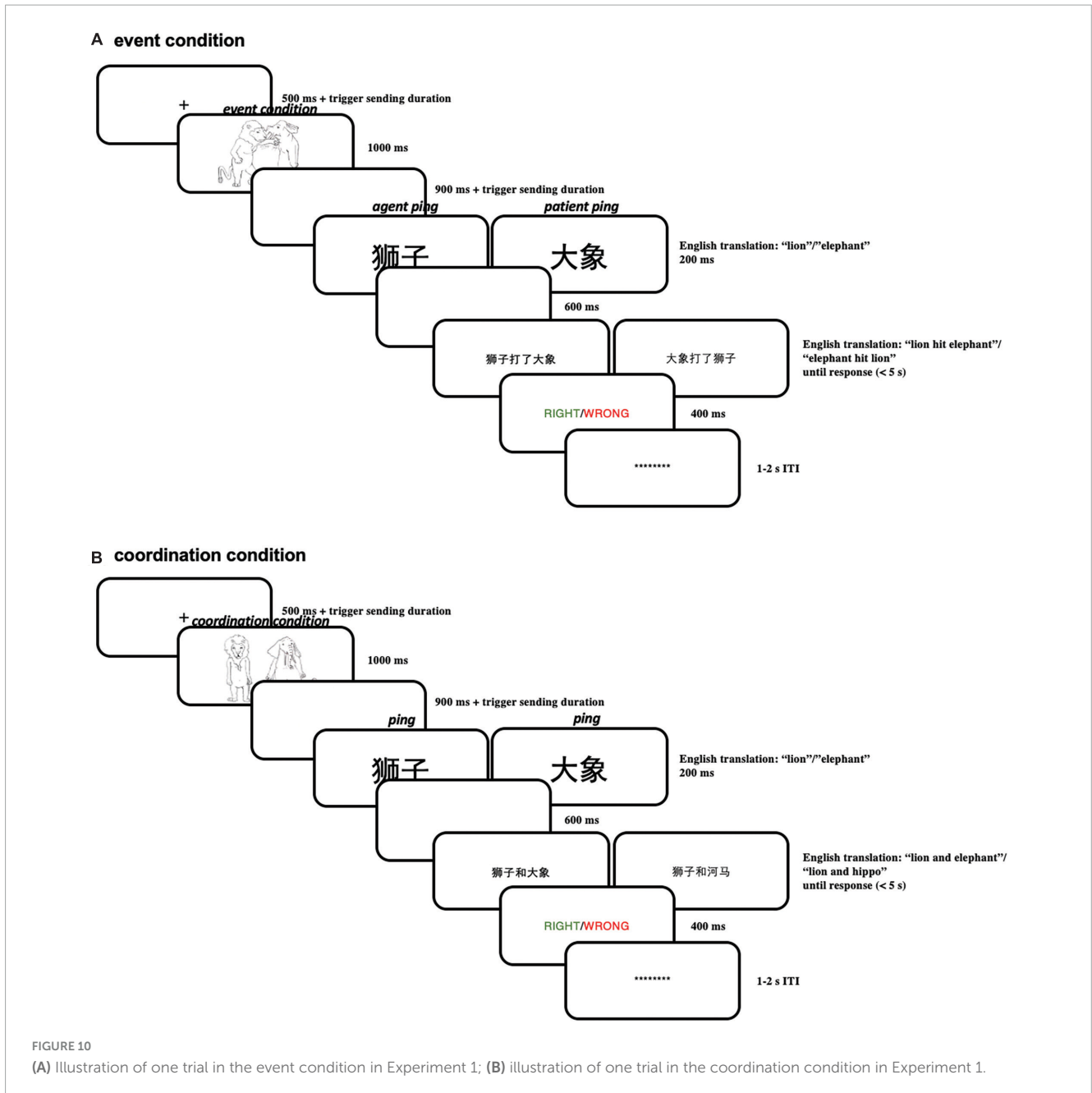
The reaction times were first averaged within subjects for all the correct trials, then paired *t*-tests were conducted across conditions in JASP 0.16.1 (JASP Team, 2022).

4.1.5 EEG recording

EEG signal was recorded using a 32-channel active electrode system (Brain Vision actiCHamp; Brain Products) with a 1,000 Hz sampling rate in an electromagnetically shielded and sound-proof room. Electrodes were placed on an EasyCap, on which electrode holders were arranged according to the 10–20 international electrode system. The impedance of each electrode was kept below 10 kΩ. The data were referenced online to electrode Cz. Two additional EOG electrodes (HEOG and VEOG) were attached for monitoring ocular activity. The EEG data were acquired with Brain Vision PyCoder software and filtered online between DC and 200 Hz with a notch filter at 50 Hz.

4.1.6 EEG data analysis

EEG data were analyzed using customized MATLAB (version: R2020b) scripts based on EEGLAB v2021.0 (Delorme and Makeig, 2004) and ERPLAB v8.20 (Lopez-Calderon and Luck, 2014). Bad channels (identified by visual inspection, 0–1 channels for each



subject) were first interpolated (spherical); throughout the whole analyzing pipeline, only scalp electrodes were included. Then the EEG data were re-referenced using average (because the ICLLabel plug-in was trained on average-referenced data, Pion-Tonachini et al., 2019), then down-sampled to 256 Hz, using functions in EEGLAB. Then the data were filtered by an IIR Butterworth filter with a cut-off at 0.01 Hz and 40 Hz using functions in ERPLAB (DC bias was also removed). Then ICA was conducted with the runica algorithm, and components that were labeled (using the ICLLabel plug-in, Pion-Tonachini et al., 2019) with labels "Muscle," "Eye," "Heart," "Line Noise," or "Channel Noise" with a confidence of $\geq 90\%$ were removed from data. Each subject had 0–2 components removed.

Epochs of $-500:3500$ ms were extracted (time locked to picture onset or ping onset), baseline corrected using the period -200 ms to 0 before picture onset. Epochs in which at least one electrode had a range of more than $100 \mu V$ during the time interval of interest (picture: $-200:3500$ ms, ping: $-200:1800$ ms) were excluded. For each subject, if there was at least one condition (i.e., event picture, coordination picture, agent ping, patient ping) that had fewer than 20 epochs, this subject was excluded. This criterion excluded the 4 subjects mentioned in Section 4.1, resulting in 19 subjects analyzed.

For the event picture vs. coordination picture comparison, we selected only $-200:3500$ ms epochs time-locked to picture onset that passed the above exclusion criteria (a range of no more than $100 \mu V$) in these 19 subjects. We first averaged the ERPs

within subjects across epochs for each electrode. For the NSW ROI analysis, we averaged the response in the NSW ROI (P7 and P8) for each subject, then conducted cluster-based permutation test across time (Maris and Oostenveld, 2007) by the `permutest()` function (Gerber, 2021) comparing the response for the NSW ROI for event picture vs. coordination picture conditions. In the cluster-based permutation tests, first, potential clusters were identified with point-to-point paired *t*-test at a threshold of $\alpha = 0.05$ (two-tailed). Then, the *t* values of these potential clusters were compared against a permuted (permutations = 9,999) distribution of *t* values, with a threshold of $\alpha = 0.05$ (two-tailed).

For the agent ping vs. patient ping comparison, we selected only $-200:1800$ ms epochs time-locked to ping onset that passed the above exclusion criteria (a range of no more than $100 \mu\text{V}$) in these 19 subjects. We first averaged the ERPs within subjects across epochs for each electrode. For the CPP ROI analysis, we averaged the response in the CPP ROI (CP1, CP2, Pz) for each subject, then conducted cluster-based permutation test (Maris and Oostenveld, 2007) by the `permutest()` function (Gerber, 2021) comparing the response for the CPP ROI for agent ping vs. patient ping conditions. Because we did not observe significant clusters in this pre-determined analysis, we then exploratorily conducted a cluster-based permutation test across time with the same parameters in each of the electrodes separately.

After beginning data analysis, we discovered an inaccuracy in the timing parameters of Experiment 1 which introduced jitter between the actual presentation of the picture and the recording of the event by the EEG system. Fortunately, a photodetector signal accompanied the ping presentation, allowing a precise estimation of its actual timing. We used this validated time estimate for the ping presentation to infer the timing of the picture presentation, but this means the estimate of the picture onset timing is less direct. In the Experiment 2 replication we used a different presentation code and oscilloscope testing to ensure precise temporal alignment of the trigger codes for both picture presentation and ping presentation.

4.2 Experiment 2

4.2.1 Participants

In total, 23 subjects participated in this experiment. Of these subjects, data from 7 subjects were excluded: 2 subjects were excluded because of incomplete data; 1 subject was excluded because of low behavioral accuracy (63%); 4 subjects were excluded based on the same procedure as Experiment 1, detailed in Section 4.1.5. The remaining 16 subjects had an age of 18–30 ($M = 24$, 8 female). 13 of them were right-handed and 3 of them were left-handed, as measured by the Edinburgh Handedness Inventory (Oldfield, 1971). They were balanced in terms of the list used in the main experiment, the list used in the practice session, and response hand. Subjects were recruited from the University of Maryland, College Park community. Subjects received monetary reimbursement for their participation. Written consent was acquired from each subject, and the procedures were approved by the Institutional Review Board of the University of Maryland.

4.2.2 Stimuli

The picture stimuli in Experiment 2 were the same as Experiment 1. The probe linguistic expressions in Experiment 2 were the same, except that they were in English: e.g., “the lion pointed at the elephant,” “the sheep was hit by the lion,” “the lion and the hippo.” The pings in Experiment 2 were English words in capital letters, e.g., “LION.”

4.2.3 Procedure

The procedure of Experiment 2 is generally similar to Experiment 1, except for the timings and font sizes (modified because of a smaller screen). Each trial started with a “+” in the middle of the screen (for 500 ms) signaling the beginning of a trial. This “+” was followed by the presentation of an event picture for 400 ms or a coordination picture. After 875 ms of a blank screen, a brief “pinging” word in capital letters and larger font (100 pt) referring to one of the two characters (e.g., “LION” or “ELEPHANT”) was presented for 200 ms, in order to highlight the just-encoded character-role pairing maintained in working memory. This ping was followed by a blank screen for 1,500 ms. After the delay period, a linguistic expression in English appeared at the center of the screen, which either matched or did not match with the picture. The whole linguistic expression was presented on the same screen.

Subjects were asked to make a response on the keyboard whether the linguistic expression matched with the picture in the same trial as fast and accurately as possible. Then a feedback screen displaying whether the choice was right (in green) or wrong (in red) was presented for 400 ms; if the subjects did not make the response within 5 s, the feedback screen would automatically appear, showing “no response” (in white). After the feedback screen, there was an inter-trial interval (ITI) uniformly random across 1,000–2,000 ms before the next trial. Within the ITI, a line of asterisks appeared on the screen; before the experiment, the subjects were encouraged to blink during this period, in order to reduce blinks within trials. Subjects were instructed to always fixate on the center of the screen within a trial. For a schematic illustration of one trial in Experiment 2 see Figure 11.

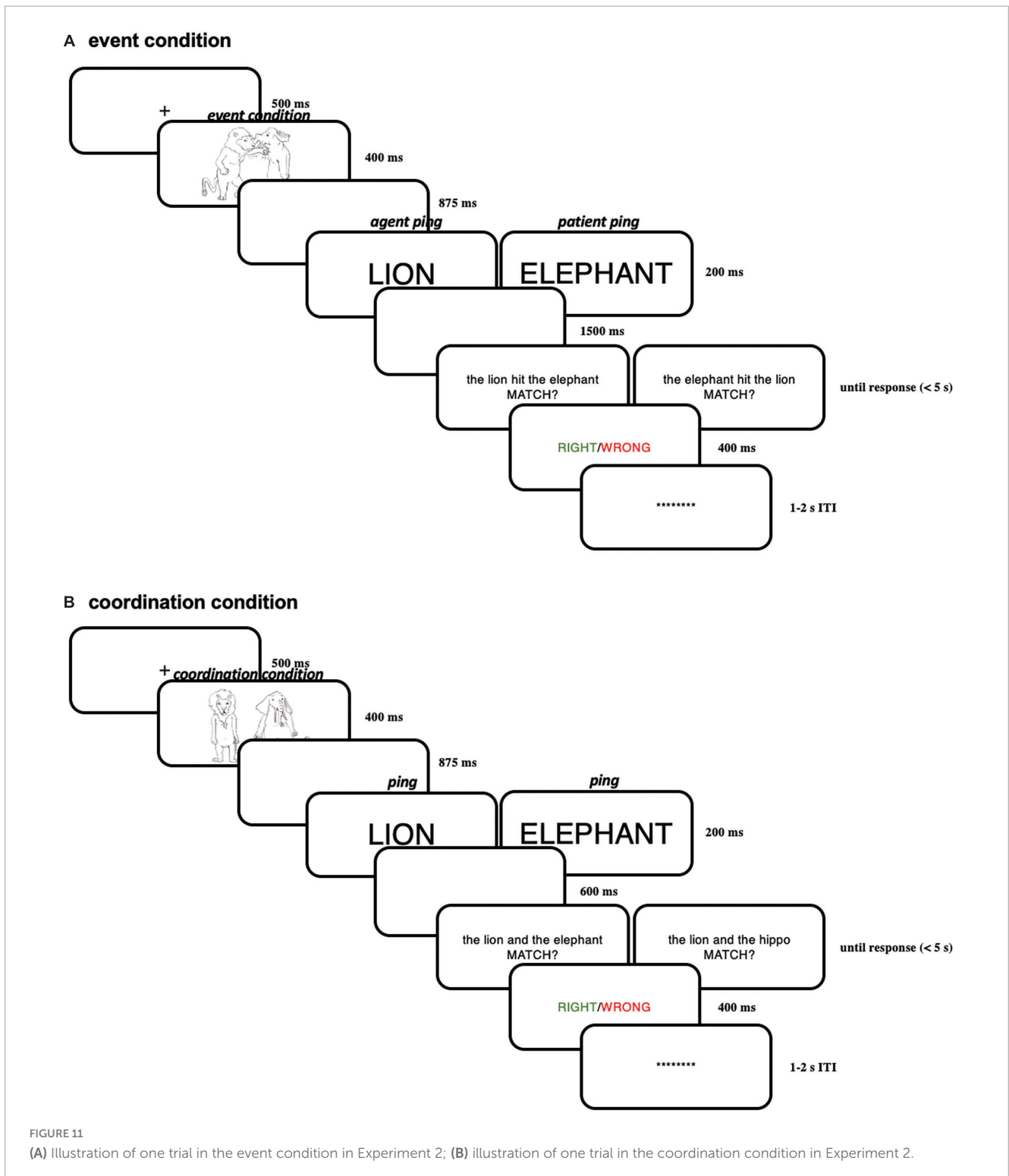
The background of the screen was gray, and the texts were white. The pictures were presented on the screen as $24^\circ \times 18^\circ$. All texts, except for the ping, were presented with a font size of 30 pt; the pings were presented with a font size of 100 pt. The distance between the subject and the screen was measured individually.

4.2.4 Behavioral data analysis

Same as Experiment 1.

4.2.5 EEG recording

Twenty-nine tin electrodes were held in place on the scalp by an elastic cap (Electro-Cap International, Inc., Eaton, OH) in a 10–20 configuration. Bipolar electrodes were placed above and below the left eye and at the outer canthus of both eyes to monitor vertical and horizontal eye movements. Additional electrodes were placed over the left and right mastoids. Scalp electrodes were referenced online to the left mastoid. The ground electrode was positioned on the scalp in front of Fz. Impedances were maintained at less than $20 \text{ k}\Omega$ for all scalp electrode sites, and less than $10 \text{ k}\Omega$ for mastoid and



ocular electrodes. The EEG signal was amplified by a NeuroScan SynAmps® Model 5083 (NeuroScan, Inc., Charlotte, NC) with a bandpass of 0.05–100 Hz and was continuously sampled at 500 Hz by an analog-to-digital converter.

4.2.6 EEG data analysis

The analysis pipeline was the same as Experiment 1 (Section 4.1.6), except that for Experiment 2, the CPP ROI was Cz, CPz and

Pz. There were 0–3 interpolated bad channels for each subject, and 0–1 ICA component was removed for each subject.

Data availability statement

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation.

Ethics statement

The studies involving humans were approved by the Institutional Review Board of the University of Maryland and New York University Shanghai. The studies were conducted in accordance with the local legislation and institutional requirements. The participants provided their written informed consent to participate in this study.

Author contributions

XY: Conceptualization, Data curation, Formal analysis, Investigation, Methodology, Software, Visualization, Writing—original draft, Writing—review and editing. JL: Data curation, Investigation, Software, Writing—review and editing. HZ: Data curation, Investigation, Software, Writing—review and editing. XT: Resources, Supervision, Writing—review and editing. EL: Conceptualization, Funding acquisition, Methodology, Resources, Supervision, Writing—original draft, Writing—review and editing.

Funding

The author(s) declare financial support was received for the research, authorship, and/or publication of this article. This work was supported by NSF #1749407 (to EL).

References

- Altmann, G., and Ekves, Z. (2019). Events as intersecting object histories: a new theory of event representation. *Psychol. Rev.* 126:817. doi: 10.1037/rev0000154
- Altmann, L. J., Saleem, A., Kendall, D., Heilman, K. M., and Rothi, L. J. G. (2006). Orthographic directionality and thematic role illustration in English and Arabic. *Brain Lang.* 97, 306–316. doi: 10.1016/j.bandl.2005.12.003
- Alunni-Menichini, K., Guimond, S., Bermudez, P., Nolden, S., Lefebvre, C., and Jolicoeur, P. (2014). Saturation of auditory short-term memory causes a plateau in the sustained anterior negativity event-related potential. *Brain Res.* 1592, 55–64. doi: 10.1016/j.brainres.2014.09.047
- Awh, E., Barton, B., and Vogel, E. K. (2007). Visual working memory represents a fixed number of items regardless of complexity. *Psychol. Sci.* 18, 622–628. doi: 10.1111/j.1467-9280.2007.01949.x
- Balaban, H., Drew, T., and Luria, R. (2019). Neural evidence for an object-based pointer system underlying working memory. *Cortex* 119, 362–372. doi: 10.1016/j.cortex.2019.05.008
- Balaban, H., and Luria, R. (2016). Object representations in visual working memory change according to the task context. *Cortex* 81, 1–13. doi: 10.1016/j.cortex.2016.04.004
- Baldassano, C., Chen, J., Zadbood, A., Pillow, J. W., Hasson, U., and Norman, K. A. (2017). Discovering event structure in continuous narrative perception and memory. *Neuron* 95, 709–721. doi: 10.1016/j.neuron.2017.06.041
- Balestrieri, E., Ronconi, L., and Melcher, D. (2022). Shared resources between visual attention and visual working memory are allocated through rhythmic sampling. *Eur. J. Neurosci.* 55, 3040–3053. doi: 10.1111/ejn.15264
- Barrett, A. M., Kim, M., Crucian, G. P., and Heilman, K. M. (2002). Spatial bias: effects of early reading direction on Korean subjects. *Neuropsychologia* 40, 1003–1012. doi: 10.1016/s0028-3932(01)00147-6
- Becke, A., Müller, N., Vellage, A., Schoenfeld, M. A., and Hopf, J. M. (2015). Neural sources of visual working memory maintenance in human parietal and ventral extrastriate visual cortex. *Neuroimage* 110, 78–86. doi: 10.1016/j.neuroimage.2015.01.059
- Berggren, N., and Eimer, M. (2016). Does contralateral delay activity reflect working memory storage or the current focus of spatial attention within visual working memory? *J. Cogn. Neurosci.* 28, 2003–2020.
- Bettencourt, K. C., and Xu, Y. (2016). Decoding the content of visual short-term memory under distraction in occipital and parietal areas. *Nat. Neurosci.* 19, 150–157. doi: 10.1038/nn.4174
- Beukers, A. O., Buschman, T. J., Cohen, J. D., and Norman, K. A. (2021). Is activity silent working memory simply episodic memory? *Trends Cogn. Sci.* 25, 284–293. doi: 10.1016/j.tics.2021.01.003
- Bidelman, G. M., Brown, J. A., and Bashivan, P. (2021). Auditory cortex supports verbal working memory capacity. *Neuroreport* 32:163.
- Brady, T. F., and Alvarez, G. A. (2015). No evidence for a fixed object limit in working memory: spatial ensemble representations inflate estimates of working memory capacity for complex objects. *J. Exp. Psychol. Learn. Memory Cogn.* 41:921. doi: 10.1037/xlm0000075
- Brady, T. F., Störmer, V. S., and Alvarez, G. A. (2016). Working memory is not fixed-capacity: more active storage capacity for real-world objects than for simple stimuli. *Proc. Natl. Acad. Sci. U.S.A.* 113, 7459–7464.

Acknowledgments

We would like to thank Matti Laine for sharing the stimuli from Hultén et al. (2014).

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

The author(s) declared that they were an editorial board member of Frontiers, at the time of submission. This had no impact on the peer review process and the final decision.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Supplementary material

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fnins.2023.1282869/full#supplementary-material>

- Brainard, D. H. (1997). The psychophysics toolbox. *Spat. Vis.* 10, 433–436.
- Brody, G. (2020). *Indexing objects in vision and communication*. Doctoral dissertation. Vienna: Central European University.
- Bueti, D., and Walsh, V. (2009). The parietal cortex and the representation of time, space, number and other magnitudes. *Philos. Trans. R. Soc. B Biol. Sci.* 364, 1831–1840.
- Cai, Y., Fulvio, J. M., Samaha, J., and Postle, B. R. (2022). Context-binding in visual working memory is reflected in bilateral event-related potentials, but not in contralateral delay activity. *eNeuro* 9, 1–15. doi: 10.1523/ENEURO.0207-22.2022
- Carey, S. (2009). *The origin of concepts*. Oxford: Oxford University Press.
- Centelles, L., Assaiante, C., Nazarian, B., Anton, J. L., and Schmitz, C. (2011). Recruitment of both the mirror and the mentalizing networks when observing social interactions depicted by point-lights: a neuroimaging study. *PLoS One* 6:e15749. doi: 10.1371/journal.pone.0015749
- Chatterjee, A., Maher, L. M., and Heilman, K. M. (1995). Spatial characteristics of thematic role representation. *Neuropsychologia* 33, 643–648. doi: 10.1016/0028-3932(94)00134-b
- Chota, S., Leto, C., van Zantwijk, L., and Van der Stigchel, S. (2022). Attention rhythmically samples multi-feature objects in working memory. *Sci. Rep.* 12:14703. doi: 10.1038/s41598-022-18819-z
- Clevenger, P. E., and Hummel, J. E. (2014). Working memory for relations among objects. *Atten. Percept. Psychophys.* 76, 1933–1953.
- Cohn, N., Paczynski, M., and Kutas, M. (2017). Not so secret agents: event-related potentials to semantic roles in visual event comprehension. *Brain Cogn.* 119, 1–9. doi: 10.1016/j.bandc.2017.09.001
- Coslett, H. B., and Saffran, E. (1991). Simultanagnosia: to see but not two see. *Brain* 114, 1523–1545.
- Cowan, N. (2001). The magical number 4 in short-term memory: a reconsideration of mental storage capacity. *Behav. Brain Sci.* 24, 87–114.
- Cruz Heredia, A. A., Dickerson, B., and Lau, E. (2022). Towards understanding sustained neural activity across syntactic dependencies. *Neurobiol. Lang.* 3, 87–108. doi: 10.1162/nol_a_00050
- de la Rosa, S., Choudhery, R. N., Curio, C., Ullman, S., Assif, L., and Buelthoff, H. H. (2014). Visual categorization of social interactions. *Vis. Cogn.* 22, 1233–1271.
- Delorme, A., and Makeig, S. (2004). EEGLAB: an open source toolbox for analysis of single-trial EEG dynamics including independent component analysis. *J. Neurosci. Methods* 134, 9–21.
- Diaz, G. K., Vogel, E. K., and Awh, E. (2021). Perceptual grouping reveals distinct roles for sustained slow wave activity and alpha oscillations in working memory. *J. Cogn. Neurosci.* 33, 1354–1364. doi: 10.1162/jocn_a_01719
- Ding, X., Gao, Z., and Shen, M. (2017). Two equals one: two human actions during social interaction are grouped as one unit in working memory. *Psychol. Sci.* 28, 1311–1320. doi: 10.1177/0956797617707318
- Dobel, C., Diesendruck, G., and Bölte, J. (2007a). How writing system and age influence spatial representations of actions: a developmental, cross-linguistic study. *Psychol. Sci.* 18, 487–491. doi: 10.1111/j.1467-9280.2007.01926.x
- Dobel, C., Enriquez-Geppert, S., Zwitserlood, P., and Bölte, J. (2014). Literacy shapes thought: the case of event representation in different cultures. *Front. Psychol.* 5:290. doi: 10.3389/fpsyg.2014.00290
- Dobel, C., Gummior, H., Bölte, J., and Zwitserlood, P. (2007b). Describing scenes hardly seen. *Acta Psychol.* 125, 129–143.
- Dowty, D. (1991). Thematic proto-roles and argument selection. *Language* 67, 547–619.
- Duncan, D. H., van Moorselaar, D., and Theeuwes, J. (2023). Pinging the brain to reveal the hidden attentional priority map using encephalography. *Nat. Commun.* 14:4749. doi: 10.1038/s41467-023-40405-8
- Eichenbaum, H., and Cohen, N. J. (2014). Can we reconcile the declarative memory and spatial navigation views on hippocampal function? *Neuron* 83, 764–770.
- Fan, Y., and Luo, H. (2023). Reactivating ordinal position information from auditory sequence memory in human brains. *Cereb. Cortex* 33, 5924–5936.
- Fan, Y., Han, Q., Guo, S., and Luo, H. (2021). Distinct neural representations of content and ordinal structure in auditory sequence memory. *J. Neurosci.* 41, 6290–6303. doi: 10.1523/JNEUROSCI.0320-21.2021
- Fan, Z., Muthukumaraswamy, S. D., Singh, K. D., and Shapiro, K. (2012). The role of sustained posterior brain activity in the serial chaining of two cognitive operations: a MEG study. *Psychophysiology* 49, 1133–1144. doi: 10.1111/j.1469-8986.2012.01391.x
- Feldmann-Wüstefeld, T. (2021). Neural measures of working memory in a bilateral change detection task. *Psychophysiology* 58:e13683. doi: 10.1111/psyp.13683
- Feldmann-Wüstefeld, T., Vogel, E. K., and Awh, E. (2018). Contralateral delay activity indexes working memory storage, not the current focus of spatial attention. *J. Cogn. Neurosci.* 30, 1185–1196.
- Fitch, W. T. (2020). Animal cognition and the evolution of human language: why we cannot focus solely on communication. *Philos. Trans. R. Soc. B* 375:20190046. doi: 10.1098/rstb.2019.0046
- Fornaciai, M., and Park, J. (2020). Neural dynamics of serial dependence in numerosity perception. *J. Cogn. Neurosci.* 32, 141–154. doi: 10.1162/jocn_a_01474
- Frankland, S. M., and Greene, J. D. (2015). An architecture for encoding sentence meaning in left mid-superior temporal cortex. *Proc. Natl. Acad. Sci. U.S.A.* 112, 11732–11737. doi: 10.1073/pnas.1421236112
- Frankland, S. M., and Greene, J. D. (2020). Two ways to build a thought: distinct forms of compositional semantic representation across brain regions. *Cereb. Cortex* 30, 3838–3855. doi: 10.1093/cercor/bhaa001
- Friedman-Hill, S. R., Robertson, L. C., and Treisman, A. (1995). Parietal contributions to visual feature binding: evidence from a patient with bilateral lesions. *Science* 269, 853–855.
- Fukuda, K., Mance, I., and Vogel, E. K. (2015). α power modulation and event-related slow wave provide dissociable correlates of visual working memory. *J. Neurosci.* 35, 14009–14016. doi: 10.1523/JNEUROSCI.5003-14.2015
- Gao, Z., Ding, X., Yang, T., Liang, J., and Shui, R. (2013). Coarse-to-fine construction for high-resolution representation in visual working memory. *PLoS One* 8:e57913. doi: 10.1371/journal.pone.0057913
- Gaston, P. E. (2020). *The role of syntactic prediction in auditory word recognition*. Doctoral dissertation. College Park, MD: University of Maryland.
- Geminiani, G., Bisiach, E., Berti, A., and Rusconi, M. L. (1995). Analogical representation and language structure. *Neuropsychologia* 33, 1565–1574.
- Gerber, E. M. (2021). *Permutest. MATLAB central file exchange*. Available online at: <https://www.mathworks.com/matlabcentral/fileexchange/71737-permutest> (accessed December 29, 2021).
- Hafri, A., Papafragou, A., and Trueswell, J. C. (2013). Getting the gist of events: recognition of two-participant actions from brief displays. *J. Exp. Psychol. Gen.* 142:880. doi: 10.1037/a0030045
- Hafri, A., Trueswell, J. C., and Epstein, R. A. (2017). Neural representations of observed actions generalize across static and dynamic visual input. *J. Neurosci.* 37, 3056–3071. doi: 10.1523/JNEUROSCI.2496-16.2017
- Hakim, N., Adam, K. C., Günseli, E., Awh, E., and Vogel, E. K. (2019). Dissecting the neural focus of attention reveals distinct processes for spatial attention and object-based storage in visual working memory. *Psychol. Sci.* 30, 526–540.
- Hedayati, S., O'Donnell, R. E., and Wyle, B. (2022). A model of working memory for latent representations. *Nat. Hum. Behav.* 6, 709–719.
- Huang, Q., Zhang, H., and Luo, H. (2021). Sequence structure organizes items in varied latent states of working memory neural network. *ELife* 10:e67589. doi: 10.7554/eLife.67589
- Hultén, A., Karvonen, L., Laine, M., and Salmelin, R. (2014). Producing speech with a newly learned morphosyntax and vocabulary: an magnetoencephalography study. *J. Cogn. Neurosci.* 26, 1721–1735. doi: 10.1162/jocn_a_00558
- Hummel, J. E., and Holyoak, K. J. (2003). A symbolic-connectionist theory of relational inference and generalization. *Psychol. Rev.* 110:220. doi: 10.1037/0033-295x.110.2.220
- Hyun, J. S., Woodman, G. F., Vogel, E. K., Hollingworth, A., and Luck, S. J. (2009). The comparison of visual working memory representations with perceptual inputs. *J. Exp. Psychol. Hum. Percept. Perform.* 35, 1140–1160.
- Isik, L., Koldewyn, K., Beeler, D., and Kanwisher, N. (2017). Perceiving social interactions in the posterior superior temporal sulcus. *Proc. Natl. Acad. Sci. U.S.A.* 114, E9145–E9152.
- Ivanova, A. A., Mineroff, Z., Zimmerer, V., Kanwisher, N., Varley, R., and Fedorenko, E. (2021). The language network is recruited but not required for nonverbal event semantics. *Neurobiol. Lang.* 2, 176–201.
- Jackendoff, R. (1987). On beyond zebra: the relation of linguistic and visual information. *Cognition* 26, 89–114. doi: 10.1016/0010-0277(87)90026-6
- JASP Team (2022). *JASP (Version 0.16.1) [Computer software]*.
- Jessop, A., and Chang, F. (2020). Thematic role information is maintained in the visual object-tracking system. *Q. J. Exp. Psychol.* 73, 146–163. doi: 10.1177/1747021819882842
- Jessop, A., and Chang, F. (2022). Thematic role tracking difficulties across multiple visual events influences role use in language production. *Vis. Cogn.* 30, 151–173.
- Jia, J., Liu, L., Fang, F., and Luo, H. (2017). Sequential sampling of visual objects during sustained attention. *PLoS Biol.* 15:e2001903. doi: 10.1371/journal.pbio.2001903
- Jia, K., Li, Y., Gong, M., Huang, H., Wang, Y., and Li, S. (2021). Perceptual learning beyond perception: mnemonic representation in early visual cortex and intraparietal sulcus. *J. Neurosci.* 41, 4476–4486. doi: 10.1523/JNEUROSCI.2780-20.2021
- Johnson, M. G. (1967). Syntactic position and rated meaning. *J. Verbal Learn. Verbal Behav.* 6, 240–246.
- Jouen, A. L., Ellmore, T. M., Madden, C. J., Pallier, C., Dominey, P. F., and Ventre-Dominey, J. (2015). Beyond the word and image: characteristics of a common meaning system for language and vision revealed by functional

- and structural imaging. *Neuroimage* 106, 72–85. doi: 10.1016/j.neuroimage.2014.11.024
- Kaan, E., Harris, A., Gibson, E., and Holcomb, P. (2000). The P600 as an index of syntactic integration difficulty. *Lang. Cogn. Process.* 15, 159–201. doi: 10.1016/j.cogbrainres.2004.09.012
- Kahneman, D., Treisman, A., and Gibbs, B. J. (1992). The reviewing of object files: object-specific integration of information. *Cogn. Psychol.* 24, 175–219. doi: 10.1016/0010-0285(92)90007-0
- Kako, E. (2006a). Thematic role properties of subjects and objects. *Cognition* 101, 1–42.
- Kako, E. (2006b). The semantics of syntactic frames. *Lang. Cogn. Process.* 21, 562–575.
- Kamiński, J., and Rutishauser, U. (2020). Between persistently active and activity-silent frameworks: novel vistas on the cellular basis of working memory. *Ann. N. Y. Acad. Sci.* 1464, 64–75. doi: 10.1111/nyas.14213
- Kazandjian, S., Gaash, E., Love, I. Y., Zivotofsky, A. Z., and Chokron, S. (2011). Spatial representation of action phrases among bidirectional readers: the effect of language environment and sentence complexity. *Soc. Psychol.* 42:249.
- Klaver, P., Smid, H. G., and Heinze, H. J. (1999). Representations in human visual short-term memory: an event-related brain potential study. *Neurosci. Lett.* 268, 65–68.
- Kleiner, M., Brainard, D., and Pelli, D. (2007). What's new in Psychtoolbox-3? *Perception* 36:14.
- Knops, A., Piazza, M., Sengupta, R., Eger, E., and Melcher, D. (2014). A shared, flexible neural map architecture reflects capacity limits in both visual short-term memory and enumeration. *J. Neurosci.* 34, 9857–9866. doi: 10.1523/JNEUROSCI.2758-13.2014
- Konkel, A., and Cohen, N. J. (2009). Relational memory and the hippocampus: representations and methods. *Front. Neurosci.* 3:166–174. doi: 10.3389/neuro.01.023.2009
- Kreither, J., Papaioannou, O., and Luck, S. J. (2022). Active working memory and simple cognitive operations. *J. Cogn. Neurosci.* 34, 313–331.
- Lalisse, M., and Smolensky, P. (2021). Distributed neural encoding of binding to thematic roles. *arXiv* [preprint]. arXiv:2110.12342.
- Lapinskaya, N., Uzomah, U., Bedny, M., and Lau, E. (2016). Electrophysiological signatures of event words: dissociating syntactic and semantic category effects in lexical processing. *Neuropsychologia* 93, 151–157. doi: 10.1016/j.neuropsychologia.2016.10.014
- Lau, E. (2018). “Neural indices of structured sentence representation: state of the art,” in *Psychology of learning and motivation*, eds K. Federmeier and E. Schotter (Amsterdam: Elsevier), 117–142.
- Lau, E., and Liao, C. H. (2018). Linguistic structure across time: ERP responses to coordinated and uncoordinated noun phrases. *Lang. Cogn. Neurosci.* 33, 633–647.
- Lau, E., and Yang, X. (2023). “Neural response to coordination reflects interpretive processes, not syntactic connectedness [conference presentation],” in *Proceedings of the human sentence processing 2023*, Pittsburgh, PA.
- Leavitt, M. L., Mendoza-Halliday, D., and Martinez-Trujillo, J. C. (2017). Sustained activity encoding working memories: not fully distributed. *Trends Neurosci.* 40, 328–346.
- Lefebvre, C., and Joliceur, P. (2016). Memory for pure tone sequences without contour. *Brain Res.* 1640, 222–231.
- Leslie, A. M., Xu, F., Tremoulet, P. D., and Scholl, B. J. (1998). Indexing and the object concept: developing ‘what’ and ‘where’ systems. *Trends Cogn. Sci.* 2, 10–18.
- Lewis-Peacock, J. A., Drysdale, A. T., Oberauer, K., and Postle, B. R. (2012). Neural evidence for a distinction between short-term memory and the focus of attention. *J. Cogn. Neurosci.* 24, 61–79.
- Li, Y., and Noguchi, Y. (2022). Neural correlates of a load-dependent decline in visual working memory. *Cereb. Cortex Commun.* 3:tgac015.
- Liu, C., Wang, R., Li, L., Ding, G., Yang, J., and Li, P. (2020). Effects of encoding modes on memory of naturalistic events. *J. Neurolinguistics* 53:100863.
- Lo, C. W., and Brennan, J. R. (2021). EEG correlates of long-distance dependency formation in mandarin Wh-Questions. *Front. Hum. Neurosci.* 15:591613. doi: 10.3389/fnhum.2021.591613
- Lopez-Calderon, J., and Luck, S. J. (2014). ERPLAB: an open-source toolbox for the analysis of event-related potentials. *Front. Hum. Neurosci.* 8:213. doi: 10.3389/fnhum.2014.00213
- Lu, X., Dai, A., Guo, Y., Shen, M., and Gao, Z. (2022). Is the social chunking of agent actions in working memory resource-demanding? *Cognition* 229:105249. doi: 10.1016/j.cognition.2022.105249
- Lu, X., Huang, J., Yi, Y., Shen, M., Weng, X., and Gao, Z. (2016). Holding biological motion in working memory: an fMRI study. *Front. Hum. Neurosci.* 10:251. doi: 10.3389/fnhum.2016.00251
- Luck, S. J., and Vogel, E. K. (1997). The capacity of visual working memory for features and conjunctions. *Nature* 390, 279–281.
- Luck, S. J., and Vogel, E. K. (2013). Visual working memory capacity: from psychophysics and neurobiology to individual differences. *Trends Cogn. Sci.* 17, 391–400.
- Lundqvist, M., Rose, J., Warden, M. R., Buschman, T., Miller, E. K., and Herman, P. (2021). A hot-coal theory of working memory. *bioRxiv* [preprint]. doi: 10.1101/2020.12.30.424833
- Luria, R., Balaban, H., Awh, E., and Vogel, E. K. (2016). The contralateral delay activity as a neural measure of visual working memory. *Neurosci. Biobehav. Rev.* 62, 100–108.
- Luyckx, F., Nili, H., Spitzer, B., and Summerfield, C. (2019). Neural structure mapping in human probabilistic reward learning. *ELife* 8:e42816.
- Maass, A., and Russo, A. (2003). Directional bias in the mental representation of spatial events: nature or culture? *Psychol. Sci.* 14, 296–301.
- Madnani, N., Boyd-Graber, J., and Resnik, P. (2010). “Measuring transitivity using untrained annotators,” in *Proceedings of the NAACL HLT 2010 workshop on creating speech and language data with Amazon’s Mechanical Turk*, Los Angeles, CA, 188–194.
- Mahon, B. Z., and Kemmerer, D. (2020). Interactions between language, thought, and perception: cognitive and neural perspectives. *Cogn. Neuropsychol.* 37, 235–240.
- Maris, E., and Oostenveld, R. (2007). Nonparametric statistical testing of EEG and MEG-data. *J. Neurosci. Methods* 164, 177–190.
- Martin, A. E., and Doumas, L. A. (2017). A mechanism for the cortical computation of hierarchical linguistic structure. *PLoS Biol.* 15:e2000663. doi: 10.1371/journal.pbio.2000663
- Martin, B., Wiener, M., and van Wassenhove, V. (2017). A Bayesian perspective on accumulation in the magnitude system. *Sci. Rep.* 7:630. doi: 10.1038/s41598-017-00680-0
- Matchin, W., Brodbeck, C., Hammerly, C., and Lau, E. (2019a). The temporal dynamics of structure and content in sentence comprehension: evidence from fMRI-constrained MEG. *Hum. Brain Mapp.* 40, 663–678. doi: 10.1002/hbm.24403
- Matchin, W., Liao, C. H., Gaston, P., and Lau, E. (2019b). Same words, different structures: an fMRI investigation of argument relations and the angular gyrus. *Neuropsychologia* 125, 116–128. doi: 10.1016/j.neuropsychologia.2019.01.019
- Matsuyoshi, D., Ikeda, T., Sawamoto, N., Kakigi, R., Fukuyama, H., and Osaka, N. (2010). Task-irrelevant memory load induces inattention blindness without temporo-parietal suppression. *Neuropsychologia* 48, 3094–3101. doi: 10.1016/j.neuropsychologia.2010.06.021
- Matthey, L., Bays, P. M., and Dayan, P. (2015). A probabilistic palimpsest model of visual short-term memory. *PLoS Comput. Biol.* 11:e1004003. doi: 10.1371/journal.pcbi.1004003
- McCants, C. W., Katus, T., and Eimer, M. (2020). Task goals modulate the activation of part-based versus object-based representations in visual working memory. *Cogn. Neurosci.* 11, 92–100. doi: 10.1080/17588928.2019.1642864
- McKinnon, R., and Osterhout, L. (1996). Constraints on movement phenomena in sentence processing: evidence from event-related brain potentials. *Lang. Cogn. Process.* 11, 495–524.
- Mecklinger, A., and Pfeifer, E. (1996). Event-related potentials reveal topographical and temporal distinct neuronal activation patterns for spatial and object working memory. *Cogn. Brain Res.* 4, 211–224. doi: 10.1016/s0926-6410(96)00034-1
- Menenti, L., Segaert, K., and Hagoort, P. (2012). The neuronal infrastructure of speaking. *Brain Lang.* 122, 71–80.
- Mitchell, D. J., and Cusack, R. (2008). Flexible, capacity-limited activity of posterior parietal cortex in perceptual as well as visual short-term memory tasks. *Cereb. Cortex* 18, 1788–1798. doi: 10.1093/cercor/bhm205
- Mitchell, D. J., and Cusack, R. (2011). The temporal evolution of electromagnetic markers sensitive to the capacity limits of visual short-term memory. *Front. Hum. Neurosci.* 5:18. doi: 10.3389/fnhum.2011.00018
- Morrison, C., Kamal, F., Le, K., and Taler, V. (2020). Monolinguals and bilinguals respond differently to a delayed matching-to-sample task: an ERP study. *Biling. Lang. Cogn.* 23, 858–868.
- Murez, M., Smortchkova, J., and Strickland, B. (2020). “The mental files theory of singular thought,” in *Singular thought and mental files*, eds R. Goodman, J. Genone, and N. Kroll (Oxford: Oxford University Press), 107–142. doi: 10.1111/tops.12110
- Naughtin, C. K., Mattingley, J. B., and Dux, P. E. (2016). Distributed and overlapping neural substrates for object individuation and identification in visual short-term memory. *Cereb. Cortex* 26, 566–575. doi: 10.1093/cercor/bhu212
- Ngiam, W. X. Q. (2023). Mapping visual working memory models to a theoretical framework. *Psychon. Bull. Rev.* [Epub ahead of print]. doi: 10.3758/s13423-023-02356-5

- Ngiam, W. X., Foster, J. J., Adam, K. C., and Awh, E. (2022). Distinguishing guesses from fuzzy memories: further evidence for item limits in visual working memory. *Attent. Percept. Psychophys.* 85, 1695–1709.
- Ngiam, W. X., Khaw, K. L., Holcombe, A. O., and Goodbourn, P. T. (2019). Visual working memory for letters varies with familiarity but not complexity. *J. Exp. Psychol. Learn. Memory Cogn.* 45, 1761–1775.
- O'Connell, R. G., and Kelly, S. P. (2021). Neurophysiology of human perceptual decision-making. *Annu. Rev. Neurosci.* 44, 495–516.
- O'Connell, R. G., Dockree, P. M., and Kelly, S. P. (2012). A supramodal accumulation-to-bound signal that determines perceptual decisions in humans. *Nat. Neurosci.* 15, 1729–1735. doi: 10.1038/nn.3248
- O'Reilly, R. C., Ranganath, C., and Russin, J. L. (2022). The structure of systematicity in the brain. *Curr. Dir. Psychol. Sci.* 31, 124–130.
- Okruszek, Ł., Wordecha, M., Jarkiewicz, M., Kossowski, B., Lee, J., and Marchewka, A. (2018). Brain correlates of recognition of communicative interactions from biological motion in schizophrenia. *Psychol. Med.* 48, 1862–1871.
- Oldfield, R. C. (1971). The assessment and analysis of handedness: the Edinburgh inventory. *Neuropsychologia* 9, 97–113.
- Paparella, I., and Papeo, L. (2022). Chunking by social relationship in working memory. *Vis. Cogn.* 30, 1–17.
- Park, H. B., Zhang, W., and Hyun, J. S. (2017). Dissociating models of visual working memory by reaction-time distribution analysis. *Acta Psychol.* 173, 21–31.
- Peng, Y., Javangula, P. R., Lu, H., and Holyoak, K. J. (2018). “Behavioral oscillations in verification of relational role bindings,” in *Proceedings of the 40th annual conference of the cognitive science society*, eds C. Kalish, M. Rau, J. Zhu, and T. T. Rogers (Austin, TX: Cognitive Science Society).
- Perez, V. B., and Vogel, E. K. (2012). “What ERPs can tell us about working memory,” in *The Oxford handbook of event-related potential components*, eds S. Luck and E. Kappenman (Oxford: Oxford University Press), 361–372. doi: 10.1177/1073858405280524
- Peters, B., Kaiser, J., Rahm, B., and Bledowski, C. (2021). Object-based attention prioritizes working memory contents at a theta rhythm. *J. Exp. Psychol. Gen.* 150, 1250–1256. doi: 10.1037/xge0000994
- Peterson, D. J., Gözenman, F., Arciniega, H., and Berryhill, M. E. (2015). Contralateral delay activity tracks the influence of Gestalt grouping principles on active visual working memory representations. *Attent. Percept. Psychophys.* 77, 2270–2283. doi: 10.3758/s13414-015-0929-y
- Pion-Tonachini, L., Kreutz-Delgado, K., and Makeig, S. (2019). ICLabel: an automated electroencephalographic independent component classifier, dataset, and website. *Neuroimage* 198, 181–197.
- Pomper, U., and Ansoorge, U. (2021). Theta-rhythmic oscillation of working memory performance. *Psychol. Sci.* 32, 1801–1810.
- Pylshyn, Z. (1989). The role of location indexes in spatial perception: a sketch of the FINST spatial-index model. *Cognition* 32, 65–97. doi: 10.1016/0010-0277(89)90014-0
- Pylshyn, Z. W. (2001). Visual indexes, preconceptual objects, and situated vision. *Cognition* 80, 127–158.
- Quadflieg, S., Gentile, F., and Rossion, B. (2015). The neural basis of perceiving person interactions. *Cortex* 70, 5–20. doi: 10.1016/j.cortex.2014.12.020
- Quilty-Dunn, J., and Green, E. J. (2021). Perceptual attribution and perceptual reference. *Philos. Phenomenol. Res.* 106, 273–298.
- Quilty-Dunn, J., Porot, N., and Mandelbaum, E. (2022). The best game in town: the re-emergence of the language of thought hypothesis across the cognitive sciences. *Behav. Brain Sci.* 46:e261.
- Rabbitt, L. R., Roberts, D. M., McDonald, C. G., and Peterson, M. S. (2017). Neural activity reveals perceptual grouping in working memory. *Int. J. Psychophysiol.* 113, 40–45.
- Rafal, R. (2001). “Balint's syndrome,” in *Handbook of neuropsychology*, 2nd Edn, Vol. 4, ed. M. Behrmann (New York, NY: Elsevier Science), 121–141.
- Reisinger, D., Rudinger, R., Ferraro, F., Harman, C., Rawlins, K., and Van Durme, B. (2015). Semantic proto-roles. *Trans. Assoc. Comput. Linguist.* 3, 475–488. doi: 10.1016/j.cognition.2005.08.002
- Rissman, L., and Lupyan, G. (2022). A dissociation between conceptual prominence and explicit category learning: evidence from agent and patient event roles. *J. Exp. Psychol. Gen.* 151, 1707–1732. doi: 10.1037/xge0001146
- Rissman, L., and Majid, A. (2019). Thematic roles: core knowledge or linguistic construct? *Psychon. Bull. Rev.* 26, 1850–1869.
- Robitaille, N., Marois, R., Todd, J., Grimault, S., Cheyne, D., and Jolicoeur, P. (2010). Distinguishing between lateralized and nonlateralized brain activity associated with visual short-term memory: fMRI, MEG, and EEG evidence from the same observers. *Neuroimage* 53, 1334–1345. doi: 10.1016/j.neuroimage.2010.07.027
- Rose, N. S., LaRocque, J. J., Riggall, A. C., Gosseries, O., Starrett, M. J., Meyerling, E. E., et al. (2016). Reactivation of latent working memories with transcranial magnetic stimulation. *Science* 354, 1136–1139.
- Ruchkin, D. S., Johnson, R. Jr., Grafman, J., Canoune, H., and Ritter, W. (1992). Distinctions and similarities among working memory processes: an event-related potential study. *Cogn. Brain Res.* 1, 53–66.
- Schneegans, S., and Bays, P. M. (2017). Neural architecture for feature binding in visual working memory. *J. Neurosci.* 37, 3913–3925. doi: 10.1523/JNEUROSCI.3493-16.2017
- Schneegans, S., and Bays, P. M. (2019). New perspectives on binding in visual working memory. *Br. J. Psychol.* 110, 207–244.
- Shastri, L. (1999). Advances in Shruti—A neurally motivated model of relational knowledge representation and rapid inference using temporal synchrony. *Appl. Intell.* 11, 79–108.
- Shastri, L. (2002). Episodic memory and cortico-hippocampal interactions. *Trends Cogn. Sci.* 6, 162–168.
- Shastri, L. (2007). “SHRUTI: a neurally motivated architecture for rapid, scalable inference,” in *Perspectives of neural-symbolic integration*, eds B. Hammer and P. Hitzler (Berlin: Springer), 183–203.
- Shastri, L., and Ajanagadde, V. (1993). From simple associations to systematic reasoning: a connectionist representation of rules, variables and dynamic bindings using temporal synchrony. *Behav. Brain Sci.* 16, 417–451.
- Shen, M., Chen, J., Yang, X., Dong, H., Chen, H., and Zhou, J. (2021). The storage mechanism of dynamic relations in visual working memory. *Cognition* 209:104571.
- Simal, A., and Jolicoeur, P. (2020). Scanning acoustic short-term memory: evidence for two subsystems with different time-course and memory strength. *Int. J. Psychophysiol.* 155, 105–117. doi: 10.1016/j.ijpsycho.2020.06.004
- Song, J. H., and Jiang, Y. (2006). Visual working memory for simple and complex features: an fMRI study. *Neuroimage* 30, 963–972.
- Spitzer, B., Waschke, L., and Summerfield, C. (2017). Selective overweighting of larger magnitudes during noisy numerical comparison. *Nat. Hum. Behav.* 1, 1–8. doi: 10.1038/s41562-017-0145
- Sprouse, J., Kucerova, I., Park, J., Cerrone, P., Schrum, N., and Lau, E. F. (2021). Sustained anterior negativities and movement dependencies in syntax. Storrs, CT: Department of Linguistics, University of Connecticut.
- Stahl, A. E., and Feigenson, L. (2014). Social knowledge facilitates chunking in infancy. *Child Dev.* 85, 1477–1490. doi: 10.1111/cdev.12217
- Stokes, M. G. (2015). ‘Activity-silent’ working memory in prefrontal cortex: a dynamic coding framework. *Trends Cogn. Sci.* 19, 394–405.
- Strickland, B. (2017). Language reflects “core” cognition: a new theory about the origin of cross-linguistic regularities. *Cogn. Sci.* 41, 70–101. doi: 10.1111/cogs.12332
- Summerfield, C., Luyckx, F., and Sheahan, H. (2020). Structure learning and the posterior parietal cortex. *Prog. Neurobiol.* 184:101717.
- Swan, G., and Wyble, B. (2014). The binding pool: a model of shared neural resources for distinct items in visual working memory. *Attent. Percept. Psychophys.* 76, 2136–2157. doi: 10.3758/s13414-014-0633-3
- ten Oever, S., De Weerd, P., and Sack, A. T. (2020). Phase-dependent amplification of working memory content and performance. *Nat. Commun.* 11:1832. doi: 10.1038/s41467-020-15629-7
- Thompson, C. K., Bonakdarpour, B., Fix, S. C., Blumenfeld, H. K., Parrish, T. B., Gitelman, D. R., et al. (2007). Neural correlates of verb argument structure processing. *J. Cogn. Neurosci.* 19, 1753–1767.
- Thyer, W., Adam, K. C. S., Diaz, G. K., Velázquez Sánchez, I. N., Vogel, E. K., and Awh, E. (2022). Storage in visual working memory recruits a content-independent pointer system. *Psychol. Sci.* 33, 1680–1694. doi: 10.1177/09567976221090923
- Todd, J. J., and Marois, R. (2004). Capacity limit of visual short-term memory in human posterior parietal cortex. *Nature* 428, 751–754.
- Treisman, A. (1998). Feature binding, attention and object perception. *Philos. Trans. R. Soc. Lond. Series B Biol. Sci.* 353, 1295–1306.
- Ünal, E., Ji, Y., and Papafragou, A. (2021). From event representation to linguistic meaning. *Top. Cogn. Sci.* 13, 224–242.
- Vogel, E. K., and Machizawa, M. G. (2004). Neural activity predicts individual differences in visual working memory capacity. *Nature* 428, 748–751.
- Walbrin, J., and Koldewyn, K. (2019). Dyadic interaction processing in the posterior temporal cortex. *Neuroimage* 198, 296–302.
- Walsh, V. (2003). A theory of magnitude: common cortical metrics of time, space and quantity. *Trends Cogn. Sci.* 7, 483–488. doi: 10.1016/j.tics.2003.09.002
- Wang, J., Cherkassky, V. L., Yang, Y., Chang, K. M. K., Vargas, R., Diana, N., et al. (2016). Identifying thematic roles from neural representations measured by functional magnetic resonance imaging. *Cogn. Neuropsychol.* 33, 257–264. doi: 10.1080/02643294.2016.1182480
- Wang, X. J. (2021). 50 years of mnemonic persistent activity: quo vadis? *Trends Neurosci.* 44, 888–902. doi: 10.1016/j.tins.2021.09.001

- Williams, A. (2015). *Arguments in syntax and semantics*. Cambridge: Cambridge University Press.
- Williams, A. (2021). "Events in semantics," in *The cambridge handbook of the philosophy of language*, ed. P. Stalmaszczyk (Cambridge: Cambridge University Press), 366–386.
- Williams, A., Reddigari, S., and Pyllkänen, L. (2017). Early sensitivity of left perisylvian cortex to relationality in nouns and verbs. *Neuropsychologia* 100, 131–143. doi: 10.1016/j.neuropsychologia.2017.04.029
- Wilson, V. A., Zuberbühler, K., and Bickel, B. (2022). The evolutionary origins of syntax: event cognition in nonhuman primates. *Sci. Adv.* 8:eabn8464. doi: 10.1126/sciadv.abn8464
- Wolff, M. J., Ding, J., Myers, N. E., and Stokes, M. G. (2015). Revealing hidden states in visual working memory using electroencephalography. *Front. Syst. Neurosci.* 9:123. doi: 10.3389/fnsys.2015.00123
- Wolff, M. J., Jochim, J., Akyürek, E. G., and Stokes, M. G. (2017). Dynamic hidden states underlying working-memory-guided behavior. *Nat. Neurosci.* 20, 864–871. doi: 10.1038/nn.4546
- Wolff, M. J., Jochim, J., Akyürek, E. G., Buschman, T. J., and Stokes, M. G. (2020). Drifting codes within a stable coding scheme for working memory. *PLoS Biol.* 18:e3000625. doi: 10.1371/journal.pbio.3000625
- Wolff, M. J., Kandemir, G., Stokes, M. G., and Akyürek, E. G. (2019). Impulse responses reveal unimodal and bimodal access to visual and auditory working memory. *bioRxiv* [preprint]. doi: 10.1101/623835
- Wurm, M. F., and Caramazza, A. (2019). Distinct roles of temporal and frontoparietal cortex in representing actions across vision and language. *Nat. Commun.* 10:289. doi: 10.1038/s41467-018-08084-y
- Wurm, M. F., and Lingnau, A. (2015). Decoding actions at different levels of abstraction. *J. Neurosci.* 35, 7727–7735.
- Wurm, M. F., Ariani, G., Greenlee, M. W., and Lingnau, A. (2016). Decoding concrete and abstract action representations during explicit and implicit conceptual processing. *Cereb. Cortex* 26, 3390–3401. doi: 10.1093/cercor/bhv169
- Wurm, M. F., Caramazza, A., and Lingnau, A. (2017). Action categories in lateral occipitotemporal cortex are organized along sociality and transitivity. *J. Neurosci.* 37, 562–575. doi: 10.1523/JNEUROSCI.1717-16.2016
- Wutz, A., Zazio, A., and Weisz, N. (2020). Oscillatory bursts in parietal cortex reflect dynamic attention between multiple objects and ensembles. *J. Neurosci.* 40, 6927–6937. doi: 10.1523/JNEUROSCI.0231-20.2020
- Xie, W., and Zhang, W. (2017). Dissociations of the number and precision of visual short-term memory representations in change detection. *Memory Cogn.* 45, 1423–1437. doi: 10.3758/s13421-017-0739-7
- Xu, Y. (2007). The role of the superior intraparietal sulcus in supporting visual short-term memory for multifeature objects. *J. Neurosci.* 27, 11676–11686. doi: 10.1523/JNEUROSCI.3545-07.2007
- Xu, Y. (2017). Reevaluating the sensory account of visual working memory storage. *Trends Cogn. Sci.* 21, 794–815. doi: 10.1016/j.tics.2017.06.013
- Xu, Y., and Chun, M. M. (2006). Dissociable neural mechanisms supporting visual short-term memory for objects. *Nature* 440, 91–95.
- Xu, Y., and Chun, M. M. (2007). Visual grouping in human parietal cortex. *Proc. Natl. Acad. Sci. U.S.A.* 104, 18766–18771.
- Xu, Y., and Chun, M. M. (2009). Selecting and perceiving multiple visual objects. *Trends Cogn. Sci.* 13, 167–174.
- Yang, H., He, C., Han, Z., and Bi, Y. (2020). Domain-specific functional coupling between dorsal and ventral systems during action perception. *Sci. Rep.* 10:21200. doi: 10.1038/s41598-020-78276-4
- Yano, M., and Koizumi, M. (2018). Processing of non-canonical word orders in (in) felicitous contexts: evidence from event-related brain potentials. *Lang. Cogn. Neurosci.* 33, 1340–1354.
- Yano, M., and Koizumi, M. (2021). The role of discourse in long-distance dependency formation. *Lang. Cogn. Neurosci.* 36, 711–729.
- Yeung, N., and Sanfey, A. G. (2004). Independent coding of reward magnitude and valence in the human brain. *J. Neurosci.* 24, 6258–6264.
- Yu, X., and Lau, E. (2023). The binding problem 2.0: beyond perceptual features. *Cogn. Sci.* 47:e13244. doi: 10.1111/cogs.13244
- Zadbood, A., Chen, J., Leong, Y. C., Norman, K. A., and Hasson, U. (2017). How we transmit memories to other brains: constructing shared neural representations via communication. *Cereb. Cortex* 27, 4988–5000. doi: 10.1093/cercor/bhx202
- Zhang, J., Ye, C., Sun, H. J., Zhou, J., Liang, T., Li, Y., et al. (2021). The passive state: a protective mechanism for information in working memory tasks. *J. Exp. Psychol. Learn. Mem. Cogn.* 48, 1235–1248. doi: 10.1037/xlm0001092
- Zhang, W., and Luck, S. J. (2008). Discrete fixed-resolution representations in visual working memory. *Nature* 453, 233–235. doi: 10.1038/nature06860
- Zhou, H., Su, C., Wu, J., Li, J., Lu, X., Gong, L., et al. (2022). A domain-general frontoparietal network interacts with domain-preferential intermediate pathways to support working memory task. *Cereb. Cortex* 33, 2774–2787. doi: 10.1093/cercor/bhac241
- Zhou, L., and Thomas, R. D. (2015). Principal component analysis of the memory load effect in a change detection task. *Vis. Res.* 110, 1–6.
- Zhu, S. D., Zhang, L. A., and von der Heydt, R. (2020). Searching for object pointers in the visual cortex. *J. Neurophysiol.* 123, 1979–1994. doi: 10.1152/jn.00112.2020