



Research paper

Language-derived mental object load suppresses posterior alpha power

Xinchi Yu^{a,b,c,*} , Utku Türk^{b,1} , Allison Dods^{b,1}, Xing Tian^{d,e,f} , Ellen Lau^{a,b} ^a Program in Neuroscience and Cognitive Science, University of Maryland, College Park, MD, USA^b Department of Linguistics, University of Maryland, College Park, MD, USA^c Institute of Medical Psychology, Ludwig Maximilian University of Munich, Munich, Germany^d Division of Arts and Sciences, New York University Shanghai, Shanghai, China^e Shanghai Key Laboratory of Brain Functional Genomics (Ministry of Education), School of Psychology and Cognitive Science, East China Normal University, Shanghai, China^f NYU-ECNU Institute of Brain and Cognitive Science at NYU Shanghai, Shanghai, China

ARTICLE INFO

Keywords:

Posterior alpha power

Vision

Language

Object load

Mental imagery

ABSTRACT

Language and vision are two primary source modalities of information for humans. To uncover shared and unique mental representations across linguistic and visual inputs, one approach is to test the extent to which a certain neural signature is shared. In the current EEG study, we tested whether the response profile of posterior alpha power – a neural signature well-studied for visual working memory – generalizes to non-visual linguistic input. Employing a novel imagine-one-then-more paradigm in a large EEG sample ($N = 52$), we found that increasing the number of imagined objects derived from a non-visual linguistic source – auditory speech during eye closure – led to greater suppression of left posterior alpha power, and that this suppression plateaued beyond ~3 objects. This response profile closely mirrors effects previously observed for visually-derived representations. As posterior alpha power likely reflects the representational load of slots/pointers or spatial attention in moment-to-moment cognition, our study thus provides novel insights into the specific representational facets shared across language and vision.

1. Introduction

Unlike our close primate relatives, a unique endowment in humans is our full-fledged language system (Hauser et al., 2002; Hockett, 1960). For primates in general, vision is the primary information source (Etting, 2023); for humans in particular, language is a second key modality of information (Kim et al., 2019). One question of longstanding interest is therefore the extent to which the representations and computations are shared across the neuro-cognitive processing of visual and linguistic inputs (Brodbeck et al., 2016; Brody & Csibra, 2026; Givón, 2002; Yu, 2025).

Multiple possibilities have been proposed in the literature. In some cases the representation of linguistic inputs is likely to rely on “recycling” existing visual neural circuitry and associated computations, which are evolutionarily more primitive (Dehaene & Cohen,

* Corresponding author. Program in Neuroscience and Cognitive Science, University of Maryland, College Park, MD, USA.

E-mail addresses: xcyu@umd.edu, xinchiyuneuro@gmail.com (X. Yu).¹ Utku Türk and Allison Dods share second-authorship.

2007; Hickok & Poeppel, 2004, 2007). For example, the neural processing of speech may rely on “recycling” the pre-existing dual-stream architecture developed for vision (Hickok & Poeppel, 2007). In other cases language – being a unique faculty in humans – may rely on uniquely expanded or developed regions or architectures in humans (Becker et al., 2022, 2025; Friederici, 2017, 2023; Gallardo et al., 2023). For example, an expanded Broca’s area and strengthened white matter fibers in humans – compared to our close primate relatives – has been argued to be key to language processing (Gallardo et al., 2023). Indeed, recent neuroimaging studies suggest that some brain regions are unique to vision or language, and some other brain regions may be shared by both (Ivanova et al., 2021; Jouen et al., 2015; Popham et al., 2021).

One approach to shed light on this question is to study whether a neural signature that has been observed for one modality could also be observed for the other modality (Brodbeck et al., 2016; Styrkowiec et al., 2025; Yu, 2025). In the current study, we tested whether the response profile of posterior alpha power, a neural signature well-studied in vision research (specifically, visual working memory research), generalizes to non-visual linguistic input.

1.1. Posterior alpha power in visual working memory studies

A prior body of research has focused on posterior alpha power in the visual working memory literature. In a trial of a typical visual working memory (VWM) study (e.g., Fukuda, Mance, & Vogel, 2015), participants first see a memorandum consisting of one or more objects. After a short delay, they need to respond to whether a probe matches with the memorandum. Responding accurately to the probe requires holding the memorandum objects in mind across the delay. In this case, multiple objects in the mind (i.e., mental objects) are derived from a direct visual source in recent history (i.e., the presented memorandum). In other words, these mental objects are visually-derived. The posterior alpha power – alpha power (8–12 Hz) in posterior EEG electrodes – has been found to decrease as the number of visual objects increases, and to plateau beyond ~ 4 items, the typical VWM capacity (Chen et al., 2024; Diaz et al., 2021; Fukuda, Mance, & Vogel, 2015; Fukuda & Woodman, 2017).

Recent years have seen advances in understanding the nature of mental “slots”, particularly in the domain of visually-derived mental objects. Advancing from the somewhat vague conception of “slots”, recent theories propose that an object is mentally registered by a content-free pointer (or token, indexical), acting as a mental stand-in for this object and binding its features together (Awh & Vogel, 2025; Balaban, 2026; Balaban et al., 2019, 2024; Balaban & Luria, 2025; Jones, Diaz, et al., 2024; Jones, Thyer, et al., 2024; Ngiam, 2024; Park, 2025; Swan & Wyble, 2014; Thyer et al., 2022; Wei et al., 2024; Yu & Lau, 2023, 2025). Critically, in this view, pointers are represented separately from features, corresponding to object-based load and feature-based load respectively. Here, “object-based load” refers to the total number of objects, and “feature-based load” refers to the total number of features to be represented. Empirically, object-based load can be isolated by controlling for the amount of features (Balaban & Luria, 2017; Naughtin et al., 2016). For example, four objects with the same color are expected to tax comparable feature-based load as one object with this color, while differing in object-based load, i.e., the number of pointers taxed (Gao et al., 2011; Naughtin et al., 2016; Xu, 2009).

Recent evidence supports that posterior alpha power is modulated by object-based load (i.e., pointer-based load), when feature-based load and the spatial extent of attention are controlled for (Chen et al., 2024). This allows us to pinpoint whether a specific part of representation is shared across processing visual and linguistic inputs. In Chen et al. (2024), the authors tested posterior alpha power during the VWM delay period across three conditions. In all three conditions, the memorandum consisted of six Pacman-like shapes with various colors and orientations. In the “grouped” condition, the “mouths” of the six shapes formed a star shape which encouraged participants to group all six Pacman shapes as one item. In the “partially grouped” condition, three of the six shapes formed a Kanizsa triangle, thus encouraging participants to group these three shapes as one item. In the “ungrouped” condition, the orientation of the “mouths” did not encourage the perception of a larger Gestalt. Chen and colleagues discovered that posterior alpha power was most suppressed for the ungrouped condition (which had the most subjectively-perceived objects), followed by the partially grouped condition, and least suppressed for the grouped condition (which had the fewest subjectively-perceived objects). As the number of features in total and the spatial extent of attention in the memorandum were controlled for across the three conditions, the modulation of posterior alpha power is best accounted for by object-based load. However, posterior alpha power may additionally reflect the spatial extent of attended area (Diaz et al., 2021), which still requires further investigation. For example, Jones, Diaz and colleagues (2024) observed both a significant alpha power effect of number of locations attended and a marginally significant alpha power effect of number of objects. Additionally, the Chen et al. (2024) results can be alternatively interpreted as reflecting attentional control processes, that are conceptually separable from the spatial extent of attention (Adam et al., 2015; Fukuda, Woodman, & Vogel, 2015).

Although evoked sustained responses have also been relatively well-studied in vision, these responses are less suitable for our current experiment. For example, an evoked response – the negative slow wave (NSW) – appears to track visual working memory load (Feldmann-Wüstefeld, 2021; Fukuda, Mance, & Vogel, 2015; Yu et al., 2024). However, a trial in the current mental imagery study necessarily takes multiple seconds due to the time scale of mental imagery formation (Berliner, 1918; D’Angiulli & Reeves, 2002), and evoked sustained responses over longer EEG epochs are easily derailed by much larger yet non-neural drifts (Kappenman & Luck, 2010; Luck, 2014). In contrast, analysis of the alpha band is unlikely to be affected by slow drifts at multi-second level.

1.2. The current study: examining posterior alpha power for linguistic input

Unlike existing studies with visual stimuli, in the current study, we examine a non-visual scenario – using verbal instructions to generate visual imagery – and we test whether object-based load in this case modulates posterior alpha power in a similar fashion to visually-induced object load. Specifically, we asked participants to imagine objects (identical in features) in the mind based on auditory speech instructions during eye closure. These objects were not visually presented to the participants at all (Fig. 1A).

Intriguingly, we found that speech-induced mental object load also suppresses posterior alpha power, similar to what was previously reported for visually-induced mental object load. This suggests that both visually and non-visually induced objects rely on a shared set of pointers in the mind.

Note that our current study was designed to examine language-derived mental representations. We want to highlight here a distinction between vision-vs. language-derived representations and visual vs. linguistic representations. The former is only concerned with whether the representation is induced by visual or linguistic input (Bi, 2021; Dirani & Pylkkänen, 2024), while the latter concerns the nature of the representations. The representation that we were probing in the current study involves visual imagery that is induced by auditory speech input. While our current study can only address questions about vision-vs. language-derived representations, we hope our current study could serve as a stepping stone towards answering the question about visual vs. linguistic representations in the future. In our analysis, left and right posterior channels are considered different ROIs (regions of interest), due to the long-standing debate on the hemispheric lateralization of mental imagery (Brown & Kosslyn, 1993; Ehrlichman & Barrett, 1983; Liu et al., 2022; Spagna et al., 2021; see also Section 4.1).

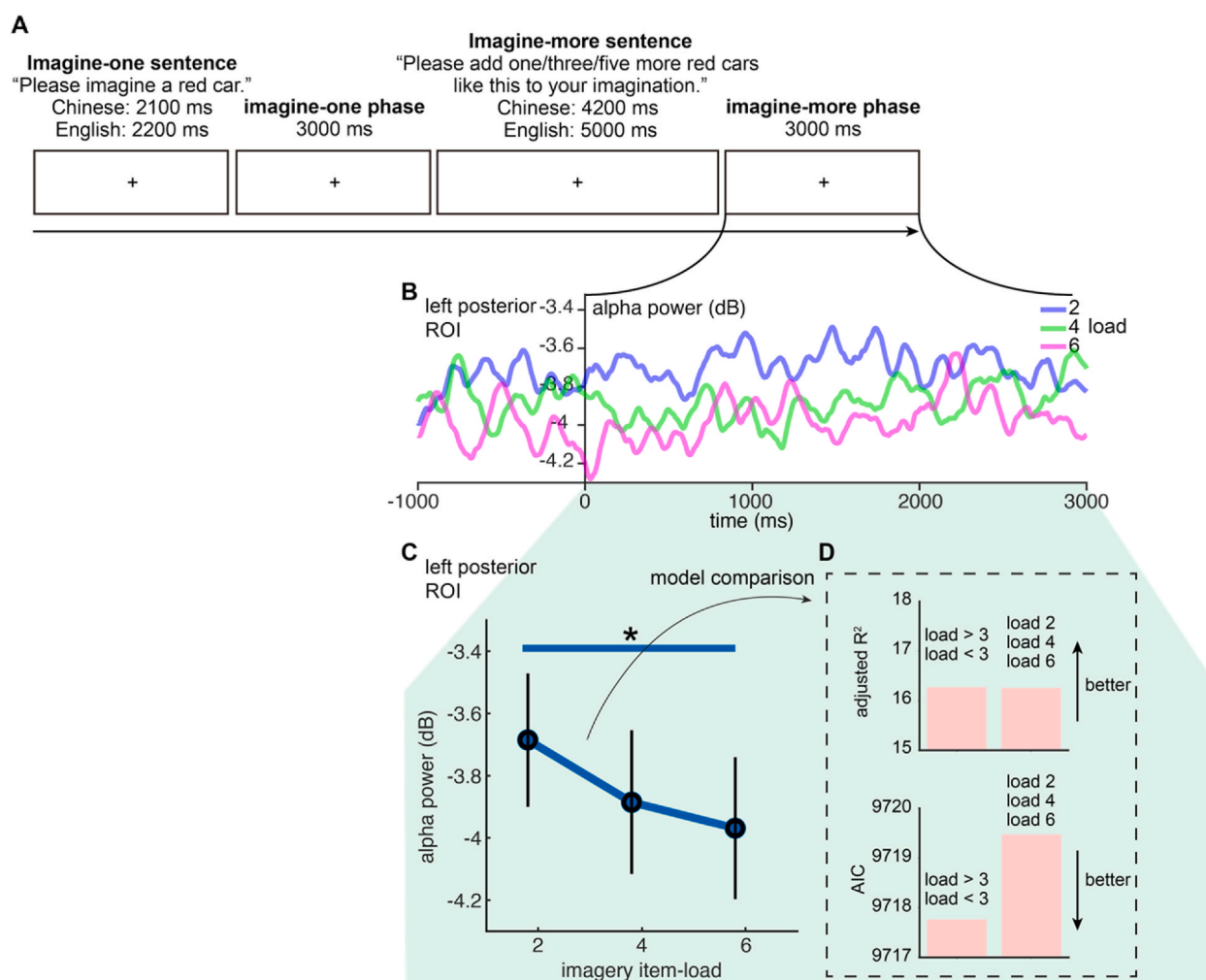


Fig. 1. Left posterior alpha power inversely predicts speech-induced mental imagery load. (A) Illustration of the procedure of a trial in our experiment. (B) Average alpha power during the imagine-more phase and the preceding 1000 ms in the left posterior ROI. Data were all resampled to 500 Hz and smoothed with a sliding 100 ms window for visualization. For the data pattern across the entire epoch see [Supplementary Fig. S1](#); for this plot with 95% confidence interval see [Supplementary Figure S2](#). Note that it is normal that the divergence across load conditions emerged before time zero, as the number word onset 2100 ms (Shanghai) or 4200 ms (Maryland) prior to time zero. (C) Average alpha power across the imagine-more phase for the left posterior ROI. Error bars represent standard error across participants. *: $p < 0.05$. (D) Model comparison for the left ROI, between a GLME model with three load levels and one with only two load levels (load < and > 3), on adjusted R^2 and AIC.

2. Materials and Methods

2.1. Participants

The entire dataset consists of 52 participants (age 18-29, 31 female), half tested in Shanghai, and half tested in Maryland: Shanghai, 26 participants, age 19-29, $M \pm SD = 23.0 \pm 2.2$, 17 female; Maryland, 26 participants, age 18-24, $M \pm SD = 19.5 \pm 1.3$, 14 female. Across the two sites, there was no significant difference in the distribution of genders (Fisher's exact test, log odds ratio = -0.47 , $p = 0.57$, two-tailed), but there was a difference in age distribution (Welch's t -test administered due to violation of the equal variance assumption): Welch's $t(40.1) = -6.97$, $p < 0.001$ (two-tailed). We do not expect the age difference to systematically affect our results, since all participants were young adults. The sample sizes for each testing site were identical, and are comparable to previous studies on posterior alpha power for visually-induced object load (Fukuda, Mance, & Vogel, 2015; Fukuda & Woodman, 2017). The participants received monetary reimbursement or course credits for their participation. Two additional participants were excluded from analysis because of not completing all trials ($N = 1$) or because a reference electrode became dislodged during the recording ($N = 1$). Participants in Shanghai were administered stimuli in Mandarin Chinese, and participants in Maryland were administered stimuli in English; they were all proficient in the corresponding language. The procedures are approved by the IRBs of New York University Shanghai and University of Maryland College Park.

2.2. Stimuli

For each language (Mandarin Chinese and English), we constructed 28 adjective-noun phrases by crossing 6 monosyllabic color adjectives with 5 monosyllabic concrete nouns. 2 phrases out of the 30 combinations were excluded for each language for semantic ambiguities (e.g., "red shirt" may also mean "non-essential"). The adjectives and nouns for Chinese are hong2 (red); lü4 (green); huang2 (yellow); lan2 (blue); bai2 (white); hei1 (black); hua1 (flower); zhi3 (paper); bi3 (pen); che1 (car); wan3 (bowl). The adjectives and nouns for English are black, blue, gray, green, red, white, car, mug, shirt, spoon, bowl. The two lists are necessarily not exactly the same in meaning as some words are only monosyllabic in one language.

For each phrase, we generated an "imagine-one" sentence in the following form: *Please imagine a [adjective] [noun]* (Chinese: 请想象 — [classifier] [adjective] [noun]). The imagine-one sentence spans 2100 ms in Chinese and 2200 ms in English. For each phrase, we also generated an "imagine-more" sentence in the following form: *Please add [one/three/five] more [adjective] [noun] like this to your imagination* (Chinese: 请在想象中增加 [one/three/five] [classifier] 这样的 [adjective] [noun]). The imagine-more sentence had a duration of 4200 ms in Chinese and 5000 ms in English, and the number word had an onset at 2100 ms for Chinese and 800 ms for English. The sentences were synthesized with built-in Mac OS voice with consistent speech rate and syllable durations (Chinese voice: "Ting-Ting"; English voice: "Alex"). Each synthesized segment was normalized to the same intensity in Praat (Boersma, 2001).

2.3. Procedure

In each trial, the participants first saw a fixation cross for 1000 ms. Then they saw an instruction of "close your eyes then press space" on the screen, where they were instructed to first close their eyes, and subsequently press the spacebar to start a trial. After 1200 ms (Shanghai) or 8200 ms (Maryland), they heard one of the imagine-one sentences, e.g., "Please imagine a red car". We chose to extend the time period between eye closing and onset of imagine instructions for the Maryland dataset because of recent observations that posterior alpha power may take several seconds to plateau after eye closure (Geller et al., 2014; Wöstmann et al., 2020). The participants were given a 3000 ms period to continue imagining the object, which we term the 'imagine-one phase'. Then, they heard one of the imagine-more sentences with the same adjective-noun combination, e.g., "Please add three more red cars like this to your imagery". This imagine-one-then-more procedure ensures that each imagined object was comparable in size across the three load conditions, mirroring the case of typical visual working memory experiments (Fukuda, Mance, & Vogel, 2015; Fukuda & Woodman, 2017; Xu & Chun, 2006). Following the imagine-more instruction, participants were given a 3000 ms period to continue imagining the object, which we term the 'imagine-more phase'. Participants were explicitly instructed at the start of the experiment to imagine multiple objects *simultaneously* in the mind. Afterwards, they heard the command "stop", and they were instructed to open their eyes and stop imagining. A self-paced inter-trial interval then followed before the start of the next trial. In line with the 84 imagine-more sentences, each participant went through 84 trials accordingly in random order. Prior to the main experiment, they were administered three practice trials with the adjective-noun combination of "green birds" (not present in the main experiment) to familiarize them with the procedure. The participants were also familiarized with the experimental word list beforehand to ensure they had no trouble comprehending the synthesized speech. Auditory stimuli were delivered binaurally via earphones for the Shanghai participants, and via a loudspeaker placed centrally in front of the Maryland participants.

2.4. EEG recording

Shanghai site. EEG was recorded from 32 electrodes mounted in an electronic cap (EasyCap; Brain Vision actiCHamp; Brain Products) according to the 10/20 system (FP1, Fz, F3, F7, FT9, FC5, FC1, C3, T7, TP9, CP5, CP1, Pz, P3, P7, O1, Oz, O2, P4, P8, TP10, CP6, CP2, C4, T8, FT10, FC6, FC2, F4, F8, FP2, Cz) at a sampling rate of 1000 Hz. Two additional EOG electrodes (HEOG and VEOG) were attached for monitoring ocular activity. The data were referenced online to electrode Cz. Impedances were maintained below 25 kΩ. The EEG signals were filtered online between DC and 200 Hz with a notch filter at 50 Hz.

Maryland site. EEG was recorded from 29 electrodes mounted in an electronic cap (Electro-Cap International Incorporated; SynAmps; Neuroscan) according to the 10/20 system (FP1, FP2, F7, F3, FZ, F4, F8, FT7, FC3, FCZ, FC4, FT8, T7, C3, CZ, C4, T8, TP7, CP3, CPZ, CP4, TP8, P7, P3, PZ, P4, P8, O1, O2) at a sampling rate of 500 Hz, with an additional electrode placed anterior to Fz as ground. Bipolar HEOG and VEOG electrodes were attached for monitoring ocular activity. The data were referenced online to the right mastoid. Impedances were maintained at less than 25 k Ω for all scalp electrode sites, and less than 10 k Ω for mastoid and ocular electrodes. The EEG signals were filtered online between 0.05 and 100 Hz.

2.5. EEG data processing

EEG data were analyzed with customized MATLAB scripts (with EEGLAB and ERPLAB). Bad channels, if any, were interpolated spherically (0-1 for each participant, mean = 0.05). After removing DC bias, raw data from the scalp electrodes were re-referenced to their average, and then passed through a 0.1-40 Hz (cutoff frequencies) Butterworth filter. After that, ICA was conducted with the runica algorithm, and components that were labeled with “Muscle,” “Eye,” “Heart,” “Line Noise,” or “Channel Noise” with a confidence higher than 90% were removed from data, based on ICLabel (Pion-Tonachini et al., 2019). On average 0.4 components were removed for each participant.

To extract alpha power activities, data was subsequently bandpass-filtered with cutoff frequencies of 8 and 12 Hz with a Butterworth filter (order = 2). Alpha power for each time point and electrode was calculated by squaring the analytic signal derived from Hilbert transformation (Adam et al., 2015; Foster et al., 2015). Epochs were extracted from 1000 ms before the offset of imagine-one sentences, until the end of the imagine-more phase. A very small proportion of epochs (0.27%) containing incidental presentation-timing fluctuations were excluded from the Chinese dataset. For each participant, outlier trials were excluded in case any electrode had a range more than 3 standard deviations from the mean range of that electrode over all trials. This resulted in 1763 remaining (81% of all trials; two: 578; four: 597; six: 588) epochs in the Chinese dataset and 1847 remaining (85% of all trials; two: 600; four: 625; six: 622) epochs in the English dataset. To facilitate comparisons across participants, alpha power in arbitrary units was then converted to decibels (dB) relative to the mean amplitude of the 1000 ms preceding the imagine-one phase (de Vries et al., 2017; Zeng et al., 2024). This period can be treated as a reference because at this point, participants did not know the number of objects they would be asked to imagine during the imagine-more sentence and the imagine-more phase. Additionally, this period was just prior to the imagine-one phase, enabling us to perform sanity checks on the imagine-one phase as well (see Supplementary Analyses). Guided by prior studies suggesting hemispheric differences in alpha power during mental imagery (Benedek et al., 2014; Moore & Lux, 1987), we treated the left (O1, P7) and right (O2, P8) posterior channels as separate ROIs. The time window of interest was the imagine-more phase, the imagination period immediately following the offset of the imagine-more sentence. Posterior alpha power was calculated by averaging the responses within the time windows and ROI channels for each trial.

To validate that the effects are confined to alpha power rather than neighboring power bands, we also bandpass-filtered data with Butterworth filter with a higher order (corresponding to steeper cutoffs; order = 4) for beta (13-30 Hz) and theta (4-7 Hz) bands. A steeper cutoff was applied to avoid contamination from alpha band activities. Data for beta and theta bands were subject to the same processing pipeline as alpha band activities, which is detailed above. For beta and theta band analyses, a similar number of trials as for alpha band analysis was retained after removing outlier trials > 3 SD following the analysis pipeline for alpha band activities (beta band: Chinese, 1731 trials, English, 1837 trials; theta band: Chinese, 1757 trials, English, 1794 trials).

2.6. Statistical analyses

To quantify the effect of load (2, 4 and 6) and ROI (left, right) on posterior alpha power during the imagine-more phase, we employed a generalized linear mixed-effect (GLME) model on trial-level data, with load, ROI and their interaction as fixed effects. Trial number was included as a random effect along with participant ID and recording site (Shanghai or Maryland) in this first model in order to account for trial-level relationships between the left and right ROI measurements: *Posterior alpha power* $\sim 1 + \text{load} + \text{ROI} + \text{load} \times \text{ROI} + (1|\text{participant}:\text{site}:\text{trial})$. The response variable was modeled with a normal distribution and linked with an identity function. Beta and theta band activities were analyzed with the same GLME model.

After establishing the interaction effect between load and ROI, we then analyzed the data for the left and right ROIs separately. As there was only one value in each trial, in this model we included only participant and recording site as random effects: *Posterior alpha power* $\sim 1 + \text{load} + (1|\text{participant}:\text{site})$. Load and alpha power were z-scored prior to analysis. To complementarily examine “pairwise differences” between load conditions, we administered complementary GLME analyses, where load was instead coded as dummy variables: (a) two variables representing the effect of load 4 relative to 2, and load 6 relative to 2, or (b) two variables representing the effect of load 2 relative to 6, and load 4 relative to 6.

In order to evaluate whether there was a plateau beyond ~ 4 items for the left ROI, we compared across two GLME models: one with three categorical load levels, and one with only a distinction between load < 3 and load > 3. The models were compared with a log-likelihood ratio test, and also on AIC (Akaike information criterion) and adjusted R^2 (cf. Yu et al., 2025). All reported p -values are two-tailed.

Additionally as a sanity check, to directly evaluate whether recording site (coded as a categorical variable) qualitatively affected our results, we also ran a GLME model on trial-level alpha power, with site and all interactions involving site as fixed effects: *Posterior alpha power* $\sim 1 + \text{load} + \text{ROI} + \text{site} + \text{load} \times \text{ROI} + \text{load} \times \text{site} + \text{site} \times \text{ROI} + \text{load} \times \text{ROI} \times \text{site} + (1|\text{participant}:\text{trial})$.

3. Results

3.1. Main results

GLME analysis revealed a significant interaction effect between ROI (left, right) and load: $\beta = 0.023$, $SE = 0.011$, $t = 2.13$, $p = 0.033$. Based on this interaction effect, we then analyzed the left and right ROIs separately. A significant effect of load was found in the left ROI, $\beta = -0.031$, $SE = 0.015$, $t = -2.01$, $p = 0.044$, but not in the right ROI, $\beta = -0.007$, $SE = 0.015$, $t = -0.46$, $p = 0.64$ (Fig. 1B and C). This lateralization effect was consistent with the visual impression from scalp distributions (Fig. 2). Complementary GLME analysis on the left ROI revealed a significant difference between load 2 and 6 ($\beta = -0.075$, $SE = 0.037$, $t = -2.01$, $p = 0.044$), but a non-significant numerical difference between load 2 and 4 ($\beta = -0.055$, $SE = 0.037$, $t = -1.48$, $p = 0.14$). Coding the dummy variable instead relative to the load 6 condition, we similarly found a significant difference between load 2 and 6 ($\beta = 0.075$, $SE = 0.037$, $t = 2.01$, $p = 0.044$), but not between load 4 and 6 ($\beta = 0.020$, $SE = 0.037$, $t = 0.54$, $p = 0.59$).

A GLME model with three load levels (adjusted $R^2 = 16.25\%$, AIC = 9719.5) did not outperform a similar model only with two load levels (with only a distinction above and below 3, adjusted $R^2 = 16.27\%$, AIC = 9717.8), mirroring the results of the log-likelihood ratio test, $\chi^2(1) = 0.29$, $p = 0.59$. Rather, the three-load-level model had numerically lower adjusted R^2 and higher AIC (Fig. 1D). Taken together, modeling load conditions as three levels appeared not to be necessary, compared to the two-load-level model.

3.2. Sanity checks

No evidence supports a qualitative difference in the pattern of our results across the two recording sites. Treating recording site and interaction effects involving site as fixed effects (*Posterior alpha power* $\sim 1 + \text{load} + \text{ROI} + \text{site} + \text{load} \times \text{ROI} + \text{load} \times \text{site} + \text{site} \times \text{ROI} + \text{load} \times \text{ROI} \times \text{site} + (1|\text{participant:trial})$), there was indeed a main effect of recording site which is understandable due to for example the different recording systems across the two sites ($\beta = 0.26$, $SE = 0.033$, $t = 8.03$, $p < 0.001$). However crucially, the interaction effects involving site were all non-significant (p 's > 0.29), and the interaction effect between ROI and load was preserved ($\beta = 0.031$, $SE = 0.016$, $t = 1.96$, $p = 0.050$). Note that we did not expect the recording sites to interact with the load effect, since the neural circuits for visual and language processing are largely shared across populations (Malik-Moraleda et al., 2022), and we are not aware of an existing theory that would predict a language-based difference in the neural basis of mental object load. Taken together, the inclusion

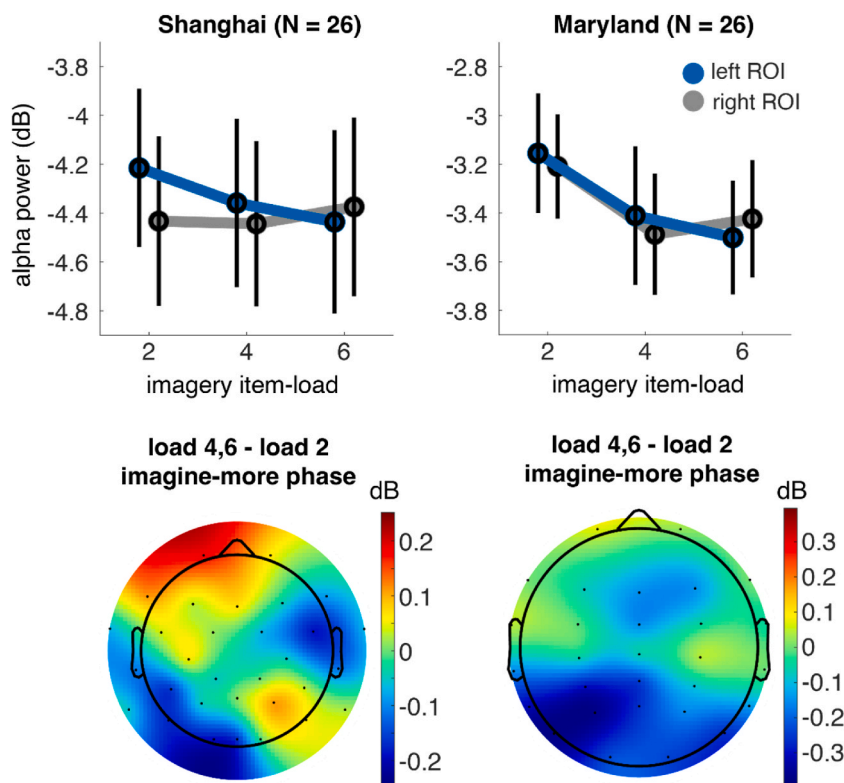


Fig. 2. Posterior alpha power during the imagine-more phase and scalp distribution for the Shanghai and Maryland participants respectively. Upper panels: Mean posterior alpha power in the left and right ROIs respectively. Error bars represent standard error across participants. Lower panels: Scalp distribution of alpha power averaged across load 4 and 6 conditions, against that for the load 2 condition, during the imagine-more phase.

of both recording sites underscores the generalizability of our observations, despite the differences in population and slight differences in experimental procedures (e.g., different duration of the imagine-more sentences, see the Materials and Methods section).

Our results are confined to alpha band activities, but not neighboring frequency bands. For beta band (13–30 Hz) and theta band (4–7 Hz) activities respectively, we conducted a similar GLME analysis as alpha band power ($\text{Posterior beta or theta power} \sim 1 + \text{load} + \text{ROI} + \text{load} \times \text{ROI} + (1|\text{participant:site:trial})$). For beta band activities, there was a significant main effect for ROI ($\beta = -0.032$, $\text{SE} = 0.011$, $t = -2.91$, $p = 0.004$), but the main effect of load and the interaction effect between ROI and load were not statistically significant (p 's > 0.13). For theta band activities, neither the main effects nor the interaction effect was statistically significant (p 's > 0.11).

4. Discussion

4.1. Left posterior alpha power reflects object-based mental object load

With a novel imagine-one-then-more paradigm leveraging 3610 EEG trials from 52 participants, we observed that left posterior alpha power decreases as object-based imagery load increases, when these mental objects were derived from a non-visual source: auditory speech instructions for imagery during eye closure. This load effect was object-based rather than feature-based, since participants were asked to imagine identical objects with shared features. Moreover, the alpha power appeared to plateau beyond ~ 3 items. This load-dependent and capacity-limited pattern is in line with what has been observed for visually-induced object-based load (Chen et al., 2024; Fukuda, Mance, & Vogel, 2015; Fukuda & Woodman, 2017).

Intriguingly, our current EEG effect was left-lateralized. Indeed, the laterality of mental imagery has long been debated (Brown & Kosslyn, 1993; Ehrlichman & Barrett, 1983; Liu et al., 2022). Interest in this topic traces back to earlier neuropsychological investigations with e.g., split-brain patients whose corpus callosum was removed. For example, Farah and colleagues (1985) tested one such patient on a task where they would see an uppercase letter, and respond to whether the corresponding lower case letter only occupied the “middle” vertical range (e.g., a, c, e) or extended into upper or lower spaces (e.g., b, d, y). This task was believed to reflect mental imagery as one would need to imagine the exact shape of the lowercase letters. Strikingly, the patient could only perform such task above chance when the uppercase letter was presented in the right visual field, corresponding to the left hemisphere. In contrast, other studies suggest a primary role of the right hemisphere for face imagery (Bowers et al., 1991). As mental imagery is a multi-component construct consisting of at least the representation of pointers and features, as well as temporally distinct phases such as formation and maintenance (Farah, 1984; Kosslyn & Shwartz, 1977), the answer to the laterality question is unlikely simple. In other words, it is unlikely that mental imagery *only* relies on the left or right hemisphere. The exact brain regions recruited and their laterality may depend on a range of factors including the content that is imagined (e.g., objects vs. faces, Liu et al., 2022) and the cognitive computations required by the task (Hu & Yu, 2023; Kosslyn, 1988). It is also possible that the lateralization of the load effect observed in the current study might be attributable to the left-lateralization of language processing (Malik-Moraleda et al., 2022; Morillon et al., 2010), if mental imagery were to share a certain extent of lateralization with its source origin. Given the limited spatial resolution of EEG, our current study was not designed to adjudicate debates over lateralization of the underlying neural generators. Future studies employing neuroimaging methods with higher spatial resolution (e.g., MEG, magnetoencephalography) will be critical in clarifying the laterality of neural sources for this effect.

4.2. The imagine-one-then-more paradigm

In the current study, we employed a new paradigm to investigate mental imagery load, namely, the imagine-one-then-more paradigm. We first asked participants to imagine one object, and then asked them to add more objects identical to the initial one to their imagery. This paradigm has some advantages that we believe can be useful for future investigations of non-visually-induced mental imagery.

First, as a majority of visual working memory load studies used arrays varying in the number of objects but keep the size of each object consistent (Luria et al., 2016; Zhang & Luck, 2008), we wanted to also control for the size of each object in the current imagery task. A phenomenological intuition is that, if one is asked to imagine two vs. six items all at once, the size per item would be smaller for six vs. two items. In the imagine-one-then-more paradigm, participants did not know how many more items they would need to imagine during the imagine-one phase, and the additional items were instructed to be of the same size as the initial one. In this way the paradigm introduced here allows the size of each object to be maximally controlled across object-based load conditions.

Second, imagining all objects all at once would also encourage participants to group all items together as one single item or set, and therefore would diminish any possible effects related to the number of objects across conditions. By asking the participants to first imagine one item and then add more items, and to imagine different items in each trial so that the task is more ad hoc, the strategy of grouping all items as a single item or set appears to be less encouraged. While we cannot guarantee that this grouping strategy was prevented in all participants, we consider our current paradigm to be advantageous compared to imagining all items all at once.

Based on these factors, we suggest that the imagine-one-then-more paradigm can be a useful tool for studying object-based load effects in mental imagery, matching visual imagery tasks with canonical visual working memory tasks better and discouraging grouping to some extent.

4.3. The capacity of mental imagery

Our current results suggest that posterior alpha power plateaued beyond ~ 3 items, in line with the well-established object-based

capacity limits for visual working memory (~3 items; Cowan, 2001; Awh et al., 2007; Zhang & Luck, 2008; Bengson & Luck, 2016). This is similarly consistent with the hypothesis that mental imagery makes use of the same object-based pointer system as standard visual working memory.

Interestingly, a recent study by Balaban and Ullman (2025) argued that participants could only simulate gravity-constrained movements of obscured visual items serially, even comparing two items to one item. In that study, participants needed to respond to when balls – based on gravity – would fall onto the ground, where only the initial part of the trajectories was shown. This means that the participants needed to mentally continue the trajectories temporally. For one ball, participants' temporal estimate aligned linearly with the actual timing. However, when two balls were presented, although participants' estimate for each item correlated well with the actual timing, one of the balls was consistently estimated as falling late – by about 600 ms. The authors took this to reflect a serial process instead of a parallel simulation process, implying a one-item capacity in this task. This intriguing result poses a seeming discrepancy with our current results, as a capacity limit of one item was suggested, instead of ~3. We hypothesize that the one-item capacity limit observed in Balaban and Ullman (2025) arose from the communication bottleneck between object-based representations and the intuitive physics system. In our current task, participants only needed to represent items in the mind, while these items do not need to partake in computations such as the ones related to intuitive physics. Taken together, our findings and those of Balaban and Ullman (2025) call for future research to further characterize the capacity limit(s) for mental imagery representations per se and downstream computations.

4.4. A shared set of pointers across visual and linguistic inputs?

If we take posterior alpha power to reflect object-based load (i.e., pointer-based load), our results offer new evidence for a long-standing question of whether visual and linguistic inputs engage the same pointer system in the mind. Two opposing hypotheses reside in previous literature on cognitive science, including the philosophy of mind. Some propose that a difference in source suggests separate mental “slots” to represent these objects (Brody, 2020; Brody & Csibra, 2026; Murez et al., 2020), while others hypothesize that a common set of mental slots can be employed to represent objects derived from different sources (Brodbeck et al., 2016; Yu, 2025; Yu & Lau, 2023). As the response profile of posterior alpha power is similar across prior vision studies and our current language study, our results are at least consistent with the hypothesis of a shared pointer system across visual and linguistic inputs. Speculating about evolution, this may suggest that the later-emerging language system “recycled” or was organized to interface with the existing pointer system initially developed to support visual objects.

Our research therefore also serves as another piece of empirical evidence towards the long-lasting debate in the philosophy of mind, as to whether our brain has one vs. many symbolic systems (i.e., Language of Thoughts; Fodor, 1975; Mandelbaum et al., 2022), by focusing on the extent to which the same “mental objects” – the building blocks of symbolic computations in the mind – are truly shared across different input modalities.

5. Limitations and future directions

Last but not least, we would like to note a few limitations of the current study that would serve as a fertile ground for future explorations.

First, while conditions with different object load were matched in the number of features and the size per object, they differed in the spatial extent of space that these objects occupy in mental imagery. Posterior alpha power has also sometimes been observed to be modulated by the spatial extent of attention, beside the number of objects (Diaz et al., 2021; Jones, Diaz, et al., 2024). This suggests that posterior alpha power may be modulated by both the number of items and the spatial extent of attention. While future studies need to disentangle the spatial extent of attention from the number of objects, our current study remains one of the first reports that the spatial extent of attention or the number of objects in mental imagery elicits such a posterior alpha power effect. Furthermore, as posterior alpha activities can be alternatively interpreted as reflecting attentional control processes (Adam et al., 2015; Fukuda, Woodman, & Vogel, 2015), future studies should clarify the explanatory power of this account by additionally evaluating a host of metrics indexing attentional control (Unsworth & Robison, 2017; Zhao et al., 2026).

Second, future studies may benefit from behavioral measurements to characterize trial- and participant-level differences in performance, as well as to characterize the potential correspondence between performance and posterior alpha power. We did, however, ask the Maryland participants at the end of the study to report the maximum number of objects that they could imagine simultaneously across the experimental session. Participants reported a capacity of 5.27 objects (mean; range 3-6). This serves as *prima facie* evidence that the participants were able to perform the task, and were in fact performing the task. Indeed, it is a long-standing challenge to ensure that participants are really performing a mental imagery task. Nevertheless, the fact that we observed a meaningful difference in posterior alpha power across conditions suggests that at least some of the participants were complying with our task instructions. However, participants' reported capacity beyond 3 items appears to be in contradiction with a ~3 item capacity limit. It is plausible that participants were chunking at least some of the items together, while it is also plausible that there were trial-by-trial fluctuations where most trials exhibited a capacity limit of around 3 items. To better capture trial-by-trial fluctuations in mental imagery, future studies could move beyond a single self-report question at the end of the study, and gather trial-by-trial subjective and objective data on the speed, content, and vividness of mental imagery. This includes for example complementary neural metrics and pupillometric responses (Hustá et al., 2019; Laeng & Sulutvedt, 2014). Administering trial-by-trial measures will also enable a more nuanced characterization of individual differences and strategy use (Adam et al., 2015; deBettencourt et al., 2019).

Third, future studies may also benefit from a within subject comparison between this task and visually-derived representations. For

example, whether the pointer-sensitive neural signals are shared across the current mental imagery task and a visual working memory task within participant. This would serve as stronger evidence compared to cross-study comparisons. Multivariate analyses such as cross-decoding or representational similarity analysis (RSA) would also provide important complementary evidence alongside univariate analysis (Jones, Diaz, et al., 2024; Jones, Thyer, et al., 2024; Suplica et al., 2025; Thyer et al., 2022; Yu & Lau, 2026). Furthermore, beyond EEG, neuroimaging methods such as MEG may offer evidence as of where in the brain these representations are shared (and where not), beyond whether and when they are shared.

Lastly, higher statistical power will be obtained by a combination of more participants, more trials, and lower impedances for EEG electrodes (Kappenman & Luck, 2010; Ngiam et al., 2021). We hope our current study and open dataset could serve as a basis for power analysis and meta-analysis in future studies.

6. Conclusion

We found that object-based load of mental objects derived from speech (a non-visual source) modulates left posterior alpha power, in a mental imagery task while participants' eyes were closed. This pattern resembles what has been observed for visually-derived mental object load, where these mental objects are derived from direct visual input in recent history (e.g., in visual working memory delay-match-to-sample tasks). This suggests that object-based load for mental objects from different sources may be accounted for by a shared neural resource. Taken together, our results are consistent with the hypothesis that mental objects tax a shared set of pointers in the mind, irrespective of their source modality.

CRedit authorship contribution statement

Xinchi Yu: Writing – review & editing, Writing – original draft, Visualization, Methodology, Investigation, Formal analysis, Data curation, Conceptualization. **Utku Türk:** Writing – review & editing, Investigation. **Allison Dods:** Writing – review & editing, Investigation. **Xing Tian:** Writing – review & editing, Supervision, Resources. **Ellen Lau:** Writing – review & editing, Writing – original draft, Supervision, Resources, Methodology, Funding acquisition, Conceptualization.

Ethics declaration

Written informed consent to take part in the study and to publish the article has been obtained from all participants or their legal representatives. The privacy rights of participants have been observed.

This study was performed in compliance with relevant laws, regulatory frameworks and guidelines where the research took place. This study was approved by the IRBs of New York University Shanghai and University of Maryland College Park. (Approval No. New York University Shanghai (Protocol #: 001-2015-The cortical basis of speech and language); University of Maryland (Protocol #: 331807))

Ethics statement

The procedures are approved by the IRBs of New York University Shanghai and University of Maryland College Park.

Artificial intelligence (AI) declaration

Some of the supplementary analysis script snippets were created with the help of AI but all carefully reviewed by the first author (X. Y.).

Funding information

The study was supported by NSF #1749407 (to E.L.).

Declaration of competing interest

The authors declare no conflict of interests.

Acknowledgements

We would like to thank Yifan Zeng, William Idsardi and Sebastián Mancha for discussions.

Appendix A. Supplementary data

Supplementary data to this article can be found online at <https://doi.org/10.1016/j.jneuroling.2026.101383>.

Data availability

Data and analysis scripts are openly shared on OSF (<https://osf.io/k4nvd/>).

References

- Adam, K. C. S., Mance, I., Fukuda, K., & Vogel, E. K. (2015). The contribution of attentional lapses to individual differences in visual working memory capacity. *Journal of Cognitive Neuroscience*, 27(8), 1601–1616. https://doi.org/10.1162/jocn_a.00811
- Awh, E., Barton, B., & Vogel, E. K. (2007). Visual working memory represents a fixed number of items regardless of complexity. *Psychological Science*, 18(7), 622–628. <https://doi.org/10.1111/j.1467-9280.2007.01949.x>
- Awh, E., & Vogel, E. K. (2025). Working memory needs pointers. *Trends in Cognitive Sciences*, 29(3), 230–241. <https://doi.org/10.1016/j.tics.2024.12.006>
- Balaban, H. (2026). Working memory's pointer system is governed by physical objecthood, not spatiotemporal information. *Psychonomic Bulletin & Review*, 33(5), 143. <https://doi.org/10.3758/s13423-026-02912-9>
- Balaban, H., Drew, T., & Luria, R. (2019). Neural evidence for an object-based pointer system underlying working memory. *Cortex*, 119, 362–372. <https://doi.org/10.1016/j.cortex.2019.05.008>
- Balaban, H., & Luria, R. (2017). Neural and behavioral evidence for an online resetting process in visual working memory. *Journal of Neuroscience*, 37(5), 1225–1239. <https://doi.org/10.1523/JNEUROSCI.2789-16.2016>
- Balaban, H., & Luria, R. (2025). Visual working memory's pointer system: Past, present, and future. *Neuroscience & Biobehavioral Reviews*, 179, Article 106440. <https://doi.org/10.1016/j.neubiorev.2025.106440>
- Balaban, H., Smith, K. A., Tenenbaum, J. B., & Ullman, T. D. (2024). Electrophysiology reveals that intuitive physics guides visual tracking and working memory. *Open Mind*, 8, 1425–1446. https://doi.org/10.1162/opmi_a.00174
- Balaban, H., & Ullman, T. D. (2025). The capacity limits of moving objects in the imagination. *Nature Communications*, 16(1), 1–11. <https://doi.org/10.1038/s41467-025-61021-8>
- Becker, Y., Eichner, C., Paquette, M., Bock, C., Girard-Buttoz, C., Jäger, C., Gräßle, T., Deschner, T., Zuberbühler, K., Wood, K., Wibbelt, G., Weiskopf, N., Walker, S., Unwin, S., Ulrich, R., Tanga, T., Szentiks, C. A., Southern, L., Stidworthy, M. F., ... Anwender, A. (2025). Long arcuate fascicle in wild and captive chimpanzees as a potential structural precursor of the language network. *Nature Communications*, 16(1), 1–9. <https://doi.org/10.1038/s41467-025-59254-8>
- Becker, Y., Loh, K. K., Coulon, O., & Meguerditchian, A. (2022). The Arcuate Fasciculus and language origins: Disentangling existing conceptions that influence evolutionary accounts. *Neuroscience & Biobehavioral Reviews*, 134, Article 104490. <https://doi.org/10.1016/j.neubiorev.2021.12.013>
- Benedek, M., Schickel, R. J., Jauk, E., Fink, A., & Neubauer, A. C. (2014). Alpha power increases in right parietal cortex reflects focused internal attention. *Neuropsychologia*, 56(1), 393–400. <https://doi.org/10.1016/j.neuropsychologia.2014.02.010>
- Bengson, J. J., & Luck, S. J. (2016). Effects of strategy on visual working memory capacity. *Psychonomic Bulletin & Review*, 23(1), 265–270. <https://doi.org/10.3758/s13423-015-0891-7>
- Berliner, A. (1918). The influence of mental work on the visual memory image. *American Journal of Psychology*, 29(4), 355. <https://doi.org/10.2307/1413963>
- Bi, Y. (2021). Dual coding of knowledge in the human brain. *Trends in Cognitive Sciences*, 25(10), 883–895. <https://doi.org/10.1016/j.tics.2021.07.006>
- Boersma, P. (2001). Praat, a system for doing phonetics by computer. *Glott International*, 5(9/10), 341–345.
- Bowers, D., Blonder, L. X., Feinberg, T., & Heilman, K. M. (1991). Differential impact of right and left hemisphere lesions on facial emotion and object imagery. *Brain*, 114(6), 2593–2609. <https://doi.org/10.1093/brain/114.6.2593>
- Brodbeck, C., Gwilliams, L., & Pykkänen, L. (2016). Language in context: MEG evidence for modality-general and -specific responses to reference resolution. *eNeuro*, 3(6). <https://doi.org/10.1523/ENEURO.0145-16.2016>
- Brody, G. (2020). *Indexing objects in vision and communication*. Central European University.
- Brody, G., & Csibra, G. (2026). Discourse referents in infancy. *Psychological Review*, 133(2), 382–393. <https://doi.org/10.1037/rev0000545>
- Brown, H. D., & Kosslyn, S. M. (1993). Cerebral lateralization. *Current Opinion in Neurobiology*, 3(2), 183–186. [https://doi.org/10.1016/0959-4388\(93\)90208-G](https://doi.org/10.1016/0959-4388(93)90208-G)
- Chen, S., Töllner, T., Müller, H. J., & Conci, M. (2024). ERPs and alpha oscillations track the encoding and maintenance of object-based representations in visual working memory. *Psychophysiology*, 61(7), 1–19. <https://doi.org/10.1111/psyp.14557>
- Cowan, N. (2001). The magical number 4 in short-term memory: A reconsideration of mental storage capacity. *Behavioral and Brain Sciences*, 24(1), 87–114. <https://doi.org/10.1017/S0140525X01003922>
- D'Angiulli, A., & Reeves, A. (2002). Generating visual mental images: Latency and vividness are inversely related. *Memory & Cognition*, 30(8), 1179–1188. <https://doi.org/10.3758/BF03213401>
- de Vries, I. E. J., van Driel, J., & Olivers, C. N. L. (2017). Posterior α EEG dynamics dissociate current from future goals in working memory-guided visual search. *Journal of Neuroscience*, 37(6), 1591–1603. <https://doi.org/10.1523/JNEUROSCI.2945-16.2016>
- deBettencourt, M. T., Keene, P. A., Awh, E., & Vogel, E. K. (2019). Real-time triggering reveals concurrent lapses of attention and working memory. *Nature Human Behaviour*, 3(8), 808–816. <https://doi.org/10.1038/s41562-019-0606-6>
- Dehaene, S., & Cohen, L. (2007). Cultural recycling of cortical maps. *Neuron*, 56(2), 384–398. <https://doi.org/10.1016/j.neuron.2007.10.004>
- Diaz, G. K., Vogel, E. K., & Awh, E. (2021). Perceptual grouping reveals distinct roles for sustained slow wave activity and alpha oscillations in working memory. *Journal of Cognitive Neuroscience*, 33(7), 1354–1364. https://doi.org/10.1162/jocn_a.01719
- Dirani, J., & Pykkänen, L. (2024). MEG evidence that modality-independent conceptual representations contain semantic and visual features. *Journal of Neuroscience*, 44(27), 1–8. <https://doi.org/10.1523/JNEUROSCI.0326-24.2024>
- Ehrlichman, H., & Barrett, J. (1983). Right hemispheric specialization for mental imagery: A review of the evidence. *Brain and Cognition*, 2(1), 55–76. [https://doi.org/10.1016/0278-2626\(83\)90029-5](https://doi.org/10.1016/0278-2626(83)90029-5)
- Etting, S. (2023). Meet the living Primates. In B. Shook, L. Braff, K. Nelson, & K. Aguilera (Eds.), *Explorations: An open invitation to biological anthropology (2nd edition)*. American Anthropological Association.
- Farah, M. J. (1984). The neurological basis of mental imagery: A componential analysis. *Cognition*, 18(1–3), 245–272. [https://doi.org/10.1016/0010-0277\(84\)90026-X](https://doi.org/10.1016/0010-0277(84)90026-X)
- Farah, M. J., Gazzaniga, M. S., Holtzman, J. D., & Kosslyn, S. M. (1985). A left hemisphere basis for visual mental imagery? *Neuropsychologia*, 23(1), 115–118. [https://doi.org/10.1016/0028-3932\(85\)90049-1](https://doi.org/10.1016/0028-3932(85)90049-1)
- Feldmann-Wüstefeld, T. (2021). Neural measures of working memory in a bilateral change detection task. *Psychophysiology*, 58(1), 1–22. <https://doi.org/10.1111/psyp.13683>
- Fodor, J. A. (1975). *The language of thought* (Vol. 5). Harvard University Press.
- Foster, J. J., Sutterer, D. W., Serences, J. T., Vogel, E. K., & Awh, E. (2015). The topography of alpha-band activity tracks the content of spatial working memory. *Journal of Neurophysiology*, 115(1), 168–177. <https://doi.org/10.1152/jn.00860.2015>
- Friederici, A. D. (2017). Evolution of the neural language network. *Psychonomic Bulletin & Review*, 24(1), 41–47. <https://doi.org/10.3758/s13423-016-1090-x>
- Friederici, A. D. (2023). Evolutionary neuroanatomical expansion of Broca's region serving a human-specific function. *Trends in Neurosciences*, 46(10), 786–796. <https://doi.org/10.1016/j.tins.2023.07.004>
- Fukuda, K., Mance, I., & Vogel, E. K. (2015). α power modulation and event-related slow wave provide dissociable correlates of visual working memory. *Journal of Neuroscience*, 35(41), 14009–14016. <https://doi.org/10.1523/JNEUROSCI.5003-14.2015>
- Fukuda, K., & Woodman, G. F. (2017). Visual working memory buffers information retrieved from visual long-term memory. *Proceedings of the National Academy of Sciences of the United States of America*, 114(20), 5306–5311. <https://doi.org/10.1073/pnas.1617874114>

- Fukuda, K., Woodman, G. F., & Vogel, E. K. (2015). Individual differences in visual working memory capacity: Contributions of attentional control to storage. *Mechanisms of Sensory Working Memory: Attention and Performance*, XXV, 105–119. <https://doi.org/10.1016/B978-0-12-801371-7.00009-0>
- Gallardo, G., Eichner, C., Sherwood, C. C., Hopkins, W. D., Anwender, A., & Friederici, A. D. (2023). Morphological evolution of language-relevant brain areas. *PLoS Biology*, 21(9), 1–13. <https://doi.org/10.1371/journal.pbio.3002266>
- Gao, Z., Xu, X., Chen, Z., Yin, J., Shen, M., & Shui, R. (2011). Contralateral delay activity tracks object identity information in visual short term memory. *Brain Research*, 1406, 30–42. <https://doi.org/10.1016/j.brainres.2011.06.049>
- Geller, A. S., Burke, J. F., Sperling, M. R., Sharan, A. D., Litt, B., Baltuch, G. H., Lucas, T. H., & Kahana, M. J. (2014). Eye closure causes widespread low-frequency power increase and focal gamma attenuation in the human electrocorticogram. *Clinical Neurophysiology*, 125(9), 1764–1773. <https://doi.org/10.1016/j.clinph.2014.01.021>
- Givón, T. (2002). The visual information-processing system as an evolutionary precursor of human language. In T. Givón, & B. F. Malle (Eds.), *The evolution of Language out of pre-language* (pp. 3–50). John Benjamins Publishing Company. <https://doi.org/10.1075/tsl.53.03.giv>
- Hauser, M. D., Chomsky, N., & Fitch, W. T. (2002). The faculty of language: What is it, who has it, and how did it evolve? *Science*, 298(5598), 1569–1579. <https://doi.org/10.1126/science.298.5598.1569>
- Hickok, G., & Poeppel, D. (2004). Dorsal and ventral streams: A framework for understanding aspects of the functional anatomy of language. *Cognition*, 92(1–2), 67–99. <https://doi.org/10.1016/j.cognition.2003.10.011>
- Hickok, G., & Poeppel, D. (2007). The cortical organization of speech processing. *Nature Reviews Neuroscience*, 8(5), 393–402. <https://doi.org/10.1038/nrn2113>
- Hockett, C. F. (1960). The origin of speech. *Scientific American*, 203(3), 88–96. <https://doi.org/10.1038/scientificamerican0960-88>
- Hu, Y., & Yu, Q. (2023). Spatiotemporal dynamics of self-generated imagery reveal a reverse cortical hierarchy from cue-induced imagery. *Cell Reports*, 42(10), Article 113242. <https://doi.org/10.1016/j.celrep.2023.113242>
- Hustá, C., Belopolsky, A., Dalmajér, E., & Mathôt, S. (2019). The pupillary light response reflects visual working memory content. *Journal of Experimental Psychology: Human Perception and Performance*, 45(11), 1522–1528. <https://doi.org/10.1037/xhp0000689>
- Ivanova, A. A., Mineroff, Z., Zimmerer, V., Kanwisher, N., Varley, R., & Fedorenko, E. (2021). The language network is recruited but not required for nonverbal event semantics. *Neurobiology of Language*, 2(2), 176–201. https://doi.org/10.1162/nol_a.00030
- Jones, H. M., Diaz, G. K., Ngiam, W. X. Q., & Awh, E. (2024). Electroencephalogram decoding reveals distinct processes for directing spatial attention and encoding into working memory. *Psychological Science*, 35(10), 1108–1138. <https://doi.org/10.1177/09567976241263002>
- Jones, H. M., Thyer, W. S., Suplica, D., & Awh, E. (2024). Cortically disparate visual features evoke content-independent load signals during storage in working memory. *Journal of Neuroscience*, 44(44), Article e0448242024. <https://doi.org/10.1523/JNEUROSCI.0448-24.2024>
- Jouen, A. L., Ellmore, T. M., Madden, C. J., Pallier, C., Dominey, P. F., & Ventre-Dominey, J. (2015). Between the word and image: Characteristics of a common meaning system for language and vision revealed by functional and structural imaging. *NeuroImage*, 106, 72–85. <https://doi.org/10.1016/j.neuroimage.2014.11.024>
- Kappenman, E. S., & Luck, S. J. (2010). The effects of electrode impedance on data quality and statistical significance in ERP recordings. *Psychophysiology*, 47(5), 888–904. <https://doi.org/10.1111/j.1469-8986.2010.01009.x>
- Kim, J. S., Elli, G. V., & Bedny, M. (2019). Knowledge of animal appearance among sighted and blind adults. *Proceedings of the National Academy of Sciences of the United States of America*, 116(23), 12123–12122. <https://doi.org/10.1073/pnas.1900952116>
- Kosslyn, S. M. (1988). Aspects of a cognitive neuroscience of mental imagery. *Science*, 240(4859), 1621–1626. <https://doi.org/10.1126/science.3289115>
- Kosslyn, S. M., & Schwartz, S. P. (1977). A simulation of visual imagery. *Cognitive Science*, 1(3), 265–295. [https://doi.org/10.1016/S0364-0213\(77\)80020-7](https://doi.org/10.1016/S0364-0213(77)80020-7)
- Laeng, B., & Sulutvedt, U. (2014). The eye pupil adjusts to imaginary light. *Psychological Science*, 25(1), 188–197. <https://doi.org/10.1177/0956797613503556>
- Liu, J., Spagna, A., & Bartolomeo, P. (2022). Hemispheric asymmetries in visual mental imagery. *Brain Structure and Function*, 227(2), 697–708. <https://doi.org/10.1007/s00429-021-02277-w>
- Luck, S. J. (2014). *An introduction to the event-related potential technique* (2nd ed.). The MIT Press.
- Luria, R., Balaban, H., Awh, E., & Vogel, E. K. (2016). The contralateral delay activity as a neural measure of visual working memory. *Neuroscience & Biobehavioral Reviews*, 62, 100–108. <https://doi.org/10.1016/j.neubiorev.2016.01.003>
- Malik-Moraleda, S., Ayyash, D., Gallée, J., Affourtit, J., Hoffmann, M., Mineroff, Z., Jouravlev, O., & Fedorenko, E. (2022). An investigation across 45 languages and 12 language families reveals a universal language network. *Nature Neuroscience*, 25(8), 1014–1019. <https://doi.org/10.1038/s41593-022-01114-5>
- Mandelbaum, E., Dunham, Y., Feiman, R., Firestone, C., Green, E. J., Harris, D., Kibbe, M. M., Kurdi, B., Mylopoulos, M., Shepherd, J., Wellwood, A., Porot, N., & Quilty-Dunn, J. (2022). Problems and mysteries of the many languages of thought. *Cognitive Science*, 46(12). <https://doi.org/10.1111/cogs.13225>
- Moore, W. H., & Lux, C. M. (1987). Hemispheric alpha asymmetries for auditorily presented words of high and low visual imagery during tasks of recognition and recall of aphasic and normal subjects. *Aphasiology*, 1(6), 475–490. <https://doi.org/10.1080/02687038708248874>
- Morillon, B., Lehongre, K., Frackowiak, R. S. J., Ducours, A., Kleinschmidt, A., Poeppel, D., & Giraud, A. L. (2010). Neurophysiological origin of human brain asymmetry for speech and language. *Proceedings of the National Academy of Sciences of the United States of America*, 107(43), 18688–18693. <https://doi.org/10.1073/pnas.1007189107>
- Murez, M., Smortchkova, J., & Strickland, B. (2020). The mental files theory of singular thought: A psychological perspective. In *Singular thought and mental files* (pp. 107–142).
- Naughtin, C. K., Mattingley, J. B., & Dux, P. E. (2016). Distributed and overlapping neural substrates for object individuation and identification in visual short-term memory. *Cerebral Cortex*, 26(2), 566–575. <https://doi.org/10.1093/cercor/bhu212>
- Ngiam, W. X. Q. (2024). Mapping visual working memory models to a theoretical framework. *Psychonomic Bulletin & Review*, 31(2), 442–459. <https://doi.org/10.3758/s13423-023-02356-5>
- Ngiam, W. X. Q., Adam, K. C. S., Quirk, C., Vogel, E. K., & Awh, E. (2021). Estimating the statistical power to detect set-size effects in contralateral delay activity. *Psychophysiology*, 58(5), 1–10. <https://doi.org/10.1111/psyp.13791>
- Park, H.-B. (2025). Contextual pointers and sensory codes in working memory: Toward a hierarchical framework. *PsyArxiv*. https://doi.org/10.31234/osf.io/3vzmf_v2
- Pion-Tonachini, L., Kreutz-Delgado, K., & Makeig, S. (2019). ICLabel: An automated electroencephalographic independent component classifier, dataset, and website. *NeuroImage*, 198, 181–197. <https://doi.org/10.1016/j.neuroimage.2019.05.026>
- Popham, S. F., Huth, A. G., Bilenko, N. Y., Deniz, F., Gao, J. S., Nunez-Elizalde, A. O., & Gallant, J. L. (2021). Visual and linguistic semantic representations are aligned at the border of human visual cortex. *Nature Neuroscience*, 24(11), 1628–1636. <https://doi.org/10.1038/s41593-021-00921-6>
- Spagna, A., Hajhajate, D., Liu, J., & Bartolomeo, P. (2021). Visual mental imagery engages the left fusiform gyrus, but not the early visual cortex: A meta-analysis of neuroimaging evidence. *Neuroscience & Biobehavioral Reviews*, 122, 201–217. <https://doi.org/10.1016/j.neubiorev.2020.12.029>
- Styrkowiec, P. P., Ngiam, W. X. Q., Epstein, W., Gneezy, R., Jones, H. M., Awh, E., & Vogel, E. K. (2025). Item-based parsing of dynamic scenes in a combined attentional tracking and working memory task. *Journal of Cognitive Neuroscience*. <https://doi.org/10.1162/jocn.a.96>
- Suplica, D., Jones, H. M., Diaz, G. K., Veillette, J. P., Nusbaum, H. C., & Awh, E. (2025). Neural evidence for modality-independent storage in working memory. *Current Biology*, 1–11. <https://doi.org/10.1016/j.cub.2025.08.007>
- Swan, G., & Wyble, B. (2014). The binding pool: A model of shared neural resources for distinct items in visual working memory. *Attention, Perception, & Psychophysics*, 76(7), 2136–2157. <https://doi.org/10.3758/s13414-014-0633-3>
- Thyer, W., Adam, K. C. S., Diaz, G. K., Velázquez Sánchez, I. N., Vogel, E. K., & Awh, E. (2022). Storage in visual working memory recruits a content-independent pointer system. *Psychological Science*, 33(10), 1680–1694. <https://doi.org/10.1177/09567976221090923>
- Unsworth, N., & Robison, M. K. (2017). The importance of arousal for variation in working memory capacity and attention control: A latent variable pupillometry study. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 43(12), 1962–1987. <https://doi.org/10.1037/xlm0000421>
- Wei, N., Song, J., Zhang, H., & Zhou, T. (2024). Unravelling the object-based nature of visual working memory: Insight from pointers. *Memory & Cognition*. <https://doi.org/10.3758/s13421-024-01643-3>

- Wöstmann, M., Schmitt, L.-M., & Obleser, J. (2020). Does closing the eyes enhance auditory attention? Eye closure increases attentional alpha-power modulation but not listening performance. *Journal of Cognitive Neuroscience*, 32(2), 212–225. <https://doi.org/10.1162/jocn.a.01403>
- Xu, Y. (2009). Distinctive neural mechanisms supporting visual object individuation and identification. *Journal of Cognitive Neuroscience*, 21(3), 511–518. <https://doi.org/10.1162/jocn.2008.21024>
- Xu, Y., & Chun, M. M. (2006). Dissociable neural mechanisms supporting visual short-term memory for objects. *Nature*, 440(7080), 91–95. <https://doi.org/10.1038/nature04262>
- Yu, X. (2025). Mapping the neural taxonomy of mental objects in moment-to-moment cognition. *Journal of Cognitive Neuroscience*, 37(11), 2093–2107. <https://doi.org/10.1162/jocn.a.02348>
- Yu, X., & Lau, E. (2023). The binding problem 2.0: Beyond perceptual features. *Cognitive Science*, 47(2). <https://doi.org/10.1111/cogs.13244>
- Yu, X., & Lau, E. (2025). A finite set of content-free pointers in visual working memory: Magnetoencephalography (MEG) evidence. *NeuroReport*, 36(3), 153–160. <https://doi.org/10.1097/WNR.0000000000002132>
- Yu, X., & Lau, E. (2026). Complementary univariate and multivariate signals in the parietal cortex for visual working memory pointers and features. *Cerebral Cortex*, 36(9). <https://doi.org/10.1093/cercor/bhag123>
- Yu, X., Li, J., Zhu, H., Tian, X., & Lau, E. (2024). Electrophysiological hallmarks for event relations and event roles in working memory. *Frontiers in Neuroscience*, 17, 1–22. <https://doi.org/10.3389/fnins.2023.1282869>
- Yu, X., Thakurdesai, S. P., & Xie, W. (2025). Associating everything with everything else, all at once: Semantic associations facilitate visual working memory formation for real-world objects. *Journal of Experimental Psychology: Human Perception and Performance*, 51(10), 1361–1373. <https://doi.org/10.1037/xhp0001347>
- Zeng, Y., Sauseng, P., & Alamia, A. (2024). Alpha traveling waves during working memory: Disentangling Bottom-Up gating and top-down gain control. *Journal of Neuroscience*, 44(50), Article e0532242024. <https://doi.org/10.1523/JNEUROSCI.0532-24.2024>
- Zhang, W., & Luck, S. J. (2008). Discrete fixed-resolution representations in visual working memory. *Nature*, 453(7192), 233–235. <https://doi.org/10.1038/nature06860>
- Zhao, C., Corriveau, A., Ke, J., Vogel, E. K., & Rosenberg, M. D. (2026). Sustained attention is more closely related to long-term memory than to attentional control. *Journal of Cognitive Neuroscience*, 38(7), 1355–1371. <https://doi.org/10.1162/JOCN.a.2488>