

ABSTRACT

Title of Dissertation: MAPPING THE NEURAL TAXONOMY OF
 MENTAL OBJECTS IN
 MOMENT-TO-MOMENT COGNITION

Xinchi Yu
Doctor of Philosophy, 2025

Dissertation directed by: Associate Professor Ellen Lau
 Department of Linguistics

We mentally represent all kinds of objects across a variety of tasks and source modalities (i.e., mental objects). One intriguing theory is that in our moment-to-moment processing in working memory, objects are represented by content-free, reassignable pointers (or indexicals). Do all mental objects share the same set of pointers in working memory? If not, where should we draw the lines between different kinds of pointers? This dissertation's goal is two-fold. One goal is theoretical: I propose a novel research program precisely centering around the above questions – aiming at unraveling the neural taxonomy of mental objects, by testing the extent to which the neural markers for pointers generalize across different paradigms, task goals, source modalities, and beyond. I review this research program in Chapter 1. The other goal is empirical: I aim to provide a set of initial results within this research program. These investigations make up Chapters 2 to 6, where I

offer initial evidence suggesting a shared set of pointers regardless of input modalities (vision and language), and a shared set of pointers representing both objects and relations between them. I also offer novel MEG markers for pointers in visual working memory tasks that would facilitate future research on this topic: a right posterior sensor-level response that gets reflected in responses in inferior parietal regions. The last but one chapter (Chapter 7) marks methodological development important for this endeavor in a new variety of MEG (OPM-MEG), laying the foundation for future discoveries of relevant neural markers and better characterization of existing neural markers with this novel technology. Chapter 8 concludes, summarizing this dissertation and proposing future steps of research.

MAPPING THE NEURAL TAXONOMY OF MENTAL OBJECTS IN
MOMENT-TO-MOMENT COGNITION

by

Xinchi Yu

Dissertation submitted to the Faculty of the Graduate School of the
University of Maryland, College Park, in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
2025

Advisory Committee:

Ellen Lau, Associate Professor, Chair

Jeffrey Lidz, Professor

Alexander Williams, Associate Professor

Weizhen Xie, Assistant Professor

Yi Ting Huang, Professor, Dean's Representative

© Copyright by
Xinchi Yu
2025

Acknowledgements

I feel extremely lucky and grateful to have spent the past five years in perhaps one of the best places in the world, and I really could not imagine a better graduate school experience.

I would like to thank Ellen for being the role model showing that doing interesting research and thinking out of the box is still possible in today's academia. I really benefited a lot from her ideas, handouts, seminars and our discussions. I also learned a lot from her scientific bravery and curiosity. Her pedagogical style is also amazing: oftentimes I feel that she already has the answer in mind, but she is always patient enough to wait for me to say that answer out aloud. I always feel that doing PhD is like being an apprentice in art: artistic taste is always subjective, but you would want yourself to be an apprentice with someone whose art taste you admire and trust -- which was fulfilled during my PhD journey. I really cannot imagine a better advisor.

I also benefited a lot from my other committee members. As everyone knows, Jeff and Yi Ting always ask the best questions, and Alexander knows everything such that he could make everything make sense. Over the years I had so much fun taking or sitting in their classes and interacting with them. I also felt lucky that Zane joined UMD, and I have learned a lot from him since, even including how to organize folders for data and analysis. There is just so much to learn from every one of these amazing people.

This dissertation also benefited from kind help from Tilmann Sander, Ernst Pöppel, Dongxue Zhang, Thomas Carey, Utku Türk, Allison Dods, Jialu Li, Hao Zhu, and Xing Tian. I would also like to thank the financial support from the Dean's Research Initiative of UMD

BSOS. I also want to say thank you to everyone that makes UMD a nice place, including the Department of Linguistics and the NACS Program. Of course, nothing would have worked without Kim Kwok, Pam Komarek, and Claire Morse.

I feel so lucky to have hung out with the best peers in the world, including Clara Cuonzo, Sebastián Mancha, and Malhaar Shah.

I also want to thank William Idsardi, Alon Hafri, Tonia Bleam, Adam Fishbein, Bob Slevc, Colin Phillips, Eva Ruhnau, Marco Chia-Ho Lai, Yifan Zeng, Xuanyu Wang, Jiaqiu Sun, Elizabeth Swanson, Andrey Laun, Rosa Eun-Kyoung Lee, Luisa Seguin, Imane Bou-Saboun, Jack Yuanfan Ying, Leslie Famularo, Joselyn Rodriguez, Ciaran Stone, Alex Krauska, London Dixon, Carmen Tang, Matthew Ganquan Shi, and Norbert Hornstein; of course, this list goes infinitely long.

Special thanks go to Ernst Pöppel and Yan Bao, without whom I would not have been in academia from the first place. I am really indebted to Ernst, who is always so kind, hardworking, resilient, and at the same time saying the coolest stuff in the world. Thank you also to all researchers around the world that took me seriously along the way; that meant a lot to me. Thank you for taking your time to reply to my email and/or meet with me and/or engage with my work.

Thank you to all my friends not mentioned above. You all truly made my life different, and I could not imagine a life without you all.

Thank you to my parents: I can't imagine myself being remotely as good as you are -- including on being parents. Thank you.

Lastly, I am deeply thankful that I grew up in Dongbei, which I still believe is a magical place.

Table of Contents

Acknowledgements.....	ii
Table of Contents.....	iv
List of Tables.....	vii
List of Figures.....	viii
List of Abbreviations.....	x
1. Chapter 1.....	1
1.1. Introduction.....	1
1.2. Mental objects: delineating the research space.....	3
1.3. Criteria for the neural signatures for pointers in moment-to-moment cognition.....	4
1.4. Candidate neural signatures for pointers in moment-to-moment visual cognition.....	6
1.5. Distinguishing pointers from arguments.....	12
1.6. How to test if the same set of pointers is shared across scenarios?	16
1.7. Object representations beyond the pointer system.....	20
1.8. Assigning features to pointers.....	21
1.9. Outline of this dissertation.....	22
2. Chapter 2.....	24
2.1. Introduction.....	24
2.2. Results.....	29
2.2.1. Experiment 2.1.....	30
2.2.2. Experiment 2.2.....	32
2.2.3. Experiment 2.3.....	33
2.2.4. GLME analysis for Experiments 2.2 and 2.3.....	35
2.3. Discussion.....	35
2.3.1. Implications for shared and distinct representations for biological vs. non- biological objects	36
2.3.2. Connections with prior literature on feature binding.....	37
2.3.3. Limitations and future directions.....	38
2.4. Materials and methods.....	39
2.4.1. Experiment 2.1.....	39
2.4.2. Experiments 2.2 and 2.3.....	42
3. Chapter 3.....	47
3.1. Introduction.....	48
3.1.1. The neural object indexical theory.....	48
3.1.2. What's different about events? Event relations and event roles	53
3.2. Results.....	65
3.2.1. Experiment 3.1.....	65
3.2.2. Experiment 3.2.....	69
3.3. Discussion.....	72
3.3.1. Representing event relations in working memory	72
3.3.2. Representing event roles in working memory	75
3.3.3. Sustained activity vs. activity-silent mechanisms.....	78
3.3.4. Event type and social interactions.....	79
3.3.5. Conclusions.....	80
3.4. Materials and methods.....	81

3.4.1.	Experiment 3.1	81
3.4.2.	Experiment 3.2	90
4.	Chapter 4	96
4.1.	Introduction	96
4.2.	Experiment 4.1: Materials and Methods	99
4.2.1.	Participants	100
4.2.2.	Stimuli	100
4.2.3.	Procedure	101
4.2.4.	EEG recording	102
4.2.5.	EEG data processing	103
4.2.6.	Statistical analyses	104
4.3.	Results	105
4.4.	Discussion	106
4.4.1.	Left posterior alpha power reflects object-based mental object load	106
4.4.2.	The capacity of mental imagery	108
4.4.3.	Is “mental object” one or multiple natural kinds?	109
4.5.	Conclusion	110
4.6.	Future directions	112
5.	Chapter 5	114
5.1.	Introduction	114
5.2.	Experiments 5.1 and 5.2: Methods	117
5.2.1.	Participants	117
5.2.2.	Procedure	117
5.2.3.	Behavioral analysis	120
5.2.4.	MEG recording	120
5.2.5.	MEG analysis	120
5.3.	Results	122
5.3.1.	Experiment 5.1	122
5.3.2.	Experiment 5.2	124
5.4.	Discussion	129
5.4.1.	Conclusion	129
5.4.2.	Limitations and future directions	130
6.	Chapter 6	132
6.1.	Introduction	132
6.2.	Experiment 6.1: Methods	134
6.2.1.	Participants	134
6.2.2.	Procedure	134
6.2.3.	Behavioral analyses	136
6.2.4.	MEG recording	136
6.2.5.	MEG analyses	137
6.3.	Results	139
6.3.1.	Behavioral results	139
6.3.2.	Sensor-level results	140
6.3.3.	Source-level analysis	141
6.4.	Discussion and conclusion	143
7.	Chapter 7	146

7.1.	Introduction.....	147
7.2.	Experiment 7.1: Method	151
7.2.1.	Participants.....	151
7.2.2.	Stimuli and procedure	151
7.2.3.	Behavioral data analysis	154
7.2.4.	OPM Setup.....	154
7.2.5.	OPM data analysis	155
7.3.	Results.....	159
7.3.1.	Behavioral results.....	159
7.3.2.	MEG results	159
7.4.	Conclusion	166
7.5.	Future directions	167
7.5.1.	On applying higher high-pass filters.....	167
7.5.2.	Other pipelines	168
7.5.3.	Exploring OPM-MEG recording with a limited number of sensors.....	168
8.	Chapter 8.....	170
8.1.	Conclusion	170
8.2.	Limitations	172
8.3.	Discussion	173
8.3.1.	Why do we need pointers?	173
8.3.2.	Future directions	175
	Bibliography	179

List of Tables

Table 1.1 Candidate neural markers for pointers in the vision science literature.....	10
Table 3.1 Example of linguistic expressions presented in Experiment 3.1.	84
Table 7.1 Example sentential stimuli for Experiment 7.1.	152

List of Figures

Figure 1.1 Illustration of the proposed research program.....	17
Figure 2.1 Illustration of a trial in Experiments 2.1-2.3.	30
Figure 2.2 The results for our Experiment 2.1.....	31
Figure 2.3 Illustration of the difference between an object-change trial and a location-change trial.	32
Figure 2.4 The results for our Experiment 2.2.....	33
Figure 2.5 The results for our Experiment 2.3.....	34
Figure 3.1 An illustration of the representations proposed by neural object indexical theory.	51
Figure 3.2 Examples of situations with and without “event relations”, as well as examples of different “event types”.....	55
Figure 3.3 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.	59
Figure 3.4 A simplified schematic illustration of our experimental design, highlighting our adapted pinging paradigm.....	63
Figure 3.5 Illustration of the electrodes of our two ROIs in Experiment 3.1; in Experiment 3.2 the NSW ROI included Cz, CPz and Pz because of a different layout.....	64
Figure 3.6 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 Figure b) for Experiment 3.1.....	67
Figure 3.7 Illustration of the results for epochs time-locked to ping onset in the NSW ROI and the CPP ROI for Experiment 1.	68
Figure 3.8 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 Figure b) for Experiment 3.2.....	71
Figure 3.9 Illustration of the results for epochs time-locked to ping onset (-200:1800 ms) in the NSW ROI and the CPP ROI for Experiment 3.2.....	72
Figure 3.10 (A) Illustration of one trial in the event condition in Experiment 3.1; (B) illustration of one trial in the coordination condition in Experiment 3.1.....	86
Figure 3.11 (A) Illustration of one trial in the event condition in Experiment 3.2; (B) illustration of one trial in the coordination condition in Experiment 3.2.....	93
Figure 3.12 Examples of the event pictures and coordination pictures used in the current experiment.....	95
Figure 4.1 Left posterior alpha power inversely predicts speech-induced mental imagery load.	99
Figure 4.2 Posterior alpha power in the left and right ROIs, illustrated across the full time period of the trial from the imagine-one phase to the end of the imaging-more phase.	111
Figure 4.3 Posterior alpha power during the imagine-more phase and scalp distribution for the English and Chinese participants respectively.....	112
Figure 5.1 Experimental procedure and behavioral results (Cowan’s K) across Experiments 5.1 and 5.2.....	119
Figure 5.2 A summary of the MEG results in Experiment 5.1.	124
Figure 5.3 A summary of the MEG results in Experiment 5.2.	127
Figure 5.4 MEG responses during the VWM period (400:1200 ms) for the colored squares in the right and left ROIs, pooled across all 40 participants in Experiments 5.1 and 5.2.....	129

Figure 6.1 Experimental hypotheses (A), design (B) and procedure (C) of Experiment 6.1.	133
Figure 6.2 ERF responses at the RPA ROI for the two load conditions.....	141
Figure 6.3 Responses during VWM formation (200-400 ms) and maintenance (400-1200 ms) in three a priori ROIs: inferior IPS, superior IPS and LOC.....	143
Figure 7.1 Illustration of a trial of our sentence processing study (Experiment 7.1).	154
Figure 7.2 Power spectrum density (PSD) of raw OPM-MEG data for each of the 11 participants.	158
Figure 7.3 Grand average ERFs across all participants (blue line: verb; orange line: noun).	162
Figure 7.4 Median ERFs for noun and verb epochs in one participant (blue line: verb; orange line: noun).	163
Figure 7.5 Cross-trial variance for each pipeline among verb trials (left) and noun trials (right), averaged across all participants.	164
Figure 7.6 RMS (root mean square) across all electrodes, based on the ERFs averaged across verbs and nouns.....	165
Figure 7.7 Comparison of N400 response (300-400 ms) amplitude at sensor F7 across verbs and nouns.....	166

List of Abbreviations

EEG: electroencephalography

MEG: magnetoencephalography

ERP: event-related response

SQUID: superconducting quantum interference device

OPM: optically-pumped magnetometer

1. Chapter 1¹

1.1. Introduction

In our moment-to-moment cognitive processing, we often need to keep one or more “objects” in mind (i.e., mental objects)². These mental objects can be derived from a variety of tasks and source modalities, ranging across direct visual-pictorial input, speech or sign languages (Brodbeck et al., 2016), long-term memory (Fukuda & Woodman, 2017), and even spontaneous thought. For example, the representation of a cat could be set up in the mind based on actually seeing a cat, or being told about a cat via spoken or sign languages, or self-generated based on one’s memory about a specific cat which is not relevant to any perceptual trace in recent history. Do we rely on a shared neural mechanism/resource to represent all kinds of mental objects in working memory, or are some mental objects represented qualitatively differently from other ones depending on e.g., task or source modality?

This question remains largely unexplored, although having been asked in the more theoretical side of the cognitive science literature (Brody, 2020; Brody & Csibra, 2025; Murez et al., 2020; Yu & Lau, 2023). Encouragingly, in recent years, one type of mental object has been extensively studied empirically, namely, when these objects are derived from direct visual input in recent history (i.e., with in the same trial or experimental session). For example, in a trial of a

¹ This chapter contains content published as Yu, X. (2025). Mapping the Neural Taxonomy of Mental Objects in Moment-to-moment Cognition. *Journal of Cognitive Neuroscience*.

² Clearly, “mental objects” is only one of the many formats in all mental representations. For example, when we want to integrate a number of path vectors that we have walked across so that we can go back to our starting point in the shortest path, each of the path vectors should be represented in some format in the mind so that we can add them up (Gallistel & Cramer, 1996), but this format is unlikely to be the same as the data structure for objects.

typical visual working memory (VWM) study, participants first see a memorandum consisting of one or more objects. After a short delay, they need to respond to a probe. In order to respond to whether the probe matches with the memorandum, one needs to hold the memorandum objects in mind across the delay. In this case, mental objects are derived from a direct visual source in recent history, different from e.g., deriving mental objects through hearing speech, or from long-term memory. A large body of work has argued that an object is mentally represented by a content-free pointer (or indexical, or token) acting as a mental stand-in for this object (Awh & Vogel, 2025; Balaban et al., 2019; Ngiam, 2024; Swan & Wyble, 2014; Thyer et al., 2022; Wei et al., 2024; Yu & Lau, 2023). Critically, pointers are represented separately from features in neural terms, so that the pointers are content-free (Awh & Vogel, 2025; Thyer et al., 2022; Y. Xu & Chun, 2006; Yu & Lau, 2025). Moreover, recent neuroimaging work in this field has offered us a promising set of neural markers indexing the number of pointers in the context of representations derived from direct vision (see Section **“Candidate neural signatures for pointers in moment-to-moment visual cognition”**). By combining the construct of “pointers” and these neural markers, the field of cognitive (neuro)science is now in the unprecedented position of being able to provide answers to a concrete version of the question raised in the first paragraph: Do we use the same set of pointers to represent mental objects for different paradigms, task goals, and source modalities? If not, where shall we draw the lines between different natural kinds of “mental objects”? In other words, what is the neural taxonomy of mental objects in moment-to-moment cognition?

In the following text, we will summarize candidate neural markers for pointers in the field of vision science, and propose a novel research program aimed at charting the taxonomic

boundaries between different kinds of mental objects, by testing how the neural markers generalize across paradigms, task goals and source modalities, etc.

1.2. Mental objects: delineating the research space

Concepts in cognitive science often (inevitably) start from some kind of shared intuition, which will and should be further refined by experimental forces. What do our intuitions suggest about what phenomena fall within the category of “mental objects” in working memory? If we think about it (and look at prior research), mental objects may need to be represented in working memory in a variety of tasks over a variety of modalities ranging from vision, audition to language (and more). Alongside, pointer-like representations have been proposed across many tasks and modalities under the name of e.g., pointers, tokens, variables or indexicals.

For example, in a multiple object tracking (MOT) task where multiple circles move randomly on the screen, participants need to respond to which of the circles are the ones that were highlighted before they moved. Interestingly, we are able to reliably track ~ 4 of these circles despite the fact that their locations are ever-changing (Storm & Pylyshyn, 1988), even against mild featural changes (K. Zhou et al., 2010) and temporary visual occlusion (Scholl & Pylyshyn, 1999). To account for such a feat, prior work has proposed that each object is tracked by a content-free pointer (Balaban et al., 2019; Pylyshyn, 2001; Scholl, 2009). Visual working memory (VWM) has also been argued to be supported by content-free tokens (Ngiam, 2024; Swan & Wyble, 2014) or pointers (Awh & Vogel, 2025; Hakim et al., 2019; Thyer et al., 2022; Yu & Lau, 2025). In the case of speech, auditory words or phonemes have been considered as (auditory) objects (Griffiths & Warren, 2004). Similarly, visual words have been thought of as

(visual) objects (Sigurdardottir et al., 2021). Here, “auditory” and “visual” are in parentheses because they may (or may not) rely on a shared set of pointers, and unraveling this kind of question is precisely the goal of our current research program. When we are comprehending a linguistic discourse involving two (or more) main characters e.g., a redheaded boy and an aquatic panda (which as the story develops turns into a flying chihuahua), we also need to track the two characters with separate indexicals in the semantic-conceptual space, which are also hypothesized to be content-free (Brody, 2020; Heim, 1982; Kamp, 1981; Yu & Lau, 2023). Pointers have also been proposed to be involved in reasoning for visual scenes for instance via intuitive physics (Balaban et al., 2024), as well as in the action programs following these reasoning processes for interacting with the world (Ballard et al., 1997). Variables have also been proposed to underlie more complex reasoning and inferences beyond the visual system (Halford et al., 2007).

What is unknown is whether the same set of pointers is shared across these various tasks (including paradigms and task goals) and modalities, and if not, across which tasks and modalities they are shared (and across which they are not). Luckily, there are potential neural markers that we can utilize to answer this question. Before introducing these neural markers, let’s first outline the core criteria for these neural markers, per theory.

1.3. Criteria for the neural signatures for pointers in moment-to-moment cognition

What are the properties that pointers theoretically need to assume? One way to think about these pointers is that they are in fact the “variables” proposed in previous cognitive theories arguing that cognition can be understood as a logic-like or computer-program-like

symbolic system. Namely, these theories include the Language of Thought (LoT) theory (Fodor, 1975; Kazanina & Poeppel, 2023; Lobina & García-Albea, 2017; Mandelbaum et al., 2022; Quilty-Dunn et al., 2023; Wellwood & Hunter, 2023) and the mental programming theory (Tian et al., 2024; Zylberberg et al., 2011). In many logic systems (e.g., predicate logic, first-order logic), there are variables as opposed to predicates, e.g., $RED(x)$ is a formalization of a red object, where x is a variable, and RED is a predicate. “Predicate” roughly corresponds to the construct of “feature” in e.g., vision science³. In many high-level programming languages (e.g., Python, C++), variables like x or y can be assigned values and enter computations. Of note, (at least some) logic systems have been found to be isomorphic to (at least some) computer programming languages (Howard, 1980), therefore LoT and mental programming can be seen as belonging to the same *family* of theories. What is common for the variables in these theories is that they are **content-free** and **reassignable**. Therefore, if we are to understand pointers following these proposals, pointers are content-free and reassignable per theory.

Being content-free (i.e., being independent from predicates) enables a pointer to bind novel feature combinations to it (where these feature combinations are not already stored in long-term memory), and being reassignable enables a pointer to be updated (Drew et al., 2013; B. Park et al., 2020) or reused (Balaban et al., 2018; Balaban & Luria, 2017; Kazanina & Poeppel, 2023). For example, if you assign the value 5 to a variable x in a computer program, the variable x per se doesn’t become 5; rather, it just points to the memory address where the value 5 is represented. In subsequent operations, the variable x can be updated (when e.g., incrementally

³ Note that the features may include relatively categorical features (e.g., being a cat), or relatively continuous features (e.g., represented by a “coordinate” in a continuous color space), or even iconic (“image-like”) representations (Block, 2023; Quilty-Dunn, 2020). Pointers may point to all these kinds of features.

calculating the sum of $5+3+6+8\dots$, where the current sum is incrementally assigned to x) or even reused for other purposes (e.g., x could be assigned a string instead of an integer in a later part of this program for other purposes). From an engineering perspective, content-independence and reassignability are important for solving many problems efficiently. If a system does not have content-independent variables, it would have trouble in e.g., representing multiple 5s simultaneously in an arithmetic task. In other words, although this system may represent types (“5ness”), it is hard for it to represent multiple tokens of the same type (i.e., multiple instances of 5) and compute with these tokens (e.g., when calculating $5+5$). If a system does not have updatable memory registers, incrementally calculating the sum of $5+3+6+8$ would require 4 memory registers to store 5, the sum of $5+3$, the sum of $5+3+6$ and the sum of $5+3+6+8$ separately, in contrast to updating the same variable with different values. Holding the intermediate sums is not always useful and is more storage-consuming (and likely more energy-consuming) than when not. If a system does not have reusable memory registers, the storage space taken up by memory registers that are not useful anymore would not be able to be released.

1.4. Candidate neural signatures for pointers in moment-to-moment visual cognition

The neural signature(s) for pointers, accordingly, need to be **content-free** and **reflect reassignability** (see below). In this section we first outline examples where each of the two properties is met in order to illustrate what we mean by these properties, and then survey potential neural signatures that could meet these properties based on the existing literature. Current neural evidence mainly come from the field of vision science, especially studies with the delay-match-to-sample paradigm (to study VWM) and the MOT task, while the domain of our

research program extends beyond one specific research field and specific paradigms. By examining whether a certain neural marker generalizes across paradigms, task goals and source modalities (and beyond), we would be able to carve out the neural taxonomy of mental objects (see Section “**How to test if the same set of pointers is shared across scenarios?**”).

The neural signatures for pointers should be content-free. Pointers are hypothesized to be independent from predicates, i.e., content-free (Awh & Vogel, 2025; Thyer et al., 2022; Y. Xu & Chun, 2006; Yu & Lau, 2023). Thus the neural signature should only be sensitive to the number of objects represented, but not the type or the number of features bound to each object. This can be achieved by examining e.g., whether a classifier on neural responses trained to tell apart the number of colored squares could also tell apart the number of oriented bars (Thyer et al., 2022), or whether a univariate neural pattern is only modulated by the number of objects but not the number or type of features represented (Naughtin et al., 2016; K. E. Wilson et al., 2012; Y. Xu & Chun, 2006; Yu & Lau, 2025). For example, with EEG, Thyer and colleagues (2022) trained and cross-tested multivariate classifiers across three conditions: memorizing colors, memorizing bar orientations, and memorizing both. Above-chance decoding ability across the VWM delay period suggested that this neural signature, mvLoad (see Table 1.1), is only sensitive to the number of objects represented, but not the type or the number of features bound to each object: that is, mvLoad is content-free. Moreover, this neural marker also generalizes to visually groupable vs. ungroupable (relatively speaking) notched discs, where the extent of spatial attention was controlled for.

The neural signatures for pointers should reflect reassignability. If pointers are reassignable, i.e., “recyclable”, the neural representation for the same pointer would be able to represent object A at time 1, but object B at another time 2 (where object A $\neq$ object B; time 1 $\neq$

time 2). Behavioral evidence suggests that more abrupt featural changes in an object is more likely to trigger the representation of new objecthood after change, compared to less abrupt featural changes (Bonatti et al., 2002; Kibbe & Leslie, 2019; K. Zhou et al., 2010). The EEG marker of CDA (contralateral delay activity), modulated by both the number of pointers and predicates (K. E. Wilson et al., 2012), temporarily drops upon an abrupt featural change of an object (compared to virtually little drop for less abrupt changes) yet quickly recovers back to a similar amplitude for less abrupt changes (B. Park et al., 2020). This suggests first discharging (or more technically, dereferencing) a pointer x_1 (standing in for the object prior the change) upon featural change, and subsequently employing a pointer x_2 to represent the object after change. The fact that similar neural responses hold prior to the change and after recovery suggests that pointer x_1 is pointer x_2 , i.e., that the same pointer was “reset” (Balaban et al., 2018, 2023; Balaban & Luria, 2017) or “remapped” (Kazanina & Poeppel, 2023; van Bree, 2024) or reassigned across the featural change. This reassignment process accompanied a subjective objecthood change yet the objects before and after are neurally represented with the same pointer. Indeed, methods with higher spatial-temporal resolution (e.g., MEG, intracranial recordings) are needed to ensure that pointer x_1 and pointer x_2 elicit the same neural responses, given EEG’s limited spatial resolution; in this case we would have better evidence that a certain cohort of neurons is truly *reassignable*, beyond that a certain neural signature *reflects reassignability* (as for the current CDA example), which is a weaker requirement. Besides, in this particular study it was unclear whether the CDA dynamics could be accounted for by the dynamics of predicate representation (instead of pointer representation), which remains to be clarified in future studies.

Existing research has revealed neural signatures that met one of the two criteria for pointers, mostly, if not all, in vision science. It has been clear that for some signatures, both criteria are not met, yet for some others, it is unclear whether both criteria could be met and thus this invites future studies. Even if they do not meet both criteria, this does not mean that we cannot use them: For example, despite the fact that CDA is sensitive to both the number of pointers and predicates (i.e., it's not content-free; Wilson et al., 2012), it can be used to index the number of pointers taxed across conditions as long as the number of predicates are well-controlled across these conditions (Balaban & Luria, 2017). Here we survey the major relevant neural markers in vision science up to now in Table 1.1, laying ground for future examination of these neural signatures and discovery of novel neural signatures. This survey focused on neural markers corresponding to VWM maintenance, and therefore is not intended as a comprehensive survey on *all* neural markers in moment-to-moment processing, but rather as a starting point for the search of pointer-tracking neural signals that reflect content-free and reassignable representations. These neural signatures are all sensitive to item-based load manipulations, but differ in the known extent to which they are content-free, as well as the known extent to which they reflect reassignability.

Following this research program, we may be able to find a neural signature that meets both criteria. Such a neural signature would, on the one hand, contribute as evidence for the validity of this specific version of pointer theory (where they need to be content-free and reassignable, following LoT and mental programming theories), and on the other hand, be used to study how much the pointers are shared across different paradigms, task goals, and source modalities (cf. Nau et al., 2024); see Section **“How to test if the same set of pointers is shared across scenarios?”**), as well as how we use these pointers in the case of grouping or event

representation, and how we compute with these pointers in e.g., reasoning and inference (see Section “**Further questions**”). If such neural signatures are extremely hard to find across *some* tasks and modalities, we may conclude (with caution) that these corresponding neural system(s) do not operate with pointers (see e.g., logic systems without variables; Akhlaghpour, 2022; Piantadosi, 2021), or in fact in these cases pointers are not implemented in the format of neuronal firings but rather lies in latent activity-silent synaptic connectivities (Kamiński & Rutishauser, 2020; Lewis-Peacock et al., 2012; Stokes, 2015), or even intracellularly in the format of e.g., RNA sequences (Akhlaghpour, 2022; Gallistel & King, 2009; Gershman, 2023). Additionally, neural signals for a shared set of pointers may be masked by differential interactions between these pointers and different types of predicates, or differential downstream processes indirectly related to these pointers. This highlights the necessity for complementary evidence from a variety of neurophysiological or neuroimaging methods, with the hope of disentangling direct reflections of the pointers themselves from these potentially correlated signals.

Table 1.1 Candidate neural markers for pointers in the vision science literature.

Note that this is by no means a comprehensive list (e.g., Lundqvist et al., 2023; Singer, 1993). Apart from the neural markers in this table that mainly reflects VWM (visual working memory) retention, other candidate neural markers for pointers may also exist for visual perception (or VWM formation) (Mazza & Caramazza, 2015; Wurm et al., 2021). ?: yet-known; ERP: event-related potential; ERF: event-related field.

Neural marker [neuroimaging method in square brackets]	Brief overview	Content-free?	Reflect reassignability?
CDA (contralateral delay activity) [EEG]	In each trial, the participants were first instructed to only attend and memorize the items on the left or right side of the screen, then a memorandum appears, with the same number of items on the left and right	No? (under debate) While some studies suggest that CDA can be modulated by both the number of pointers and	Yes (Balaban et al., 2018, 2019, 2024; Balaban & Luria, 2017; Drew et al., 2013)

	sides of the screen. After a short delay period, they are presented with a probe and instructed to respond to whether the item(s) in the probe are the same as in the memorandum (only for the memorized side). CDA is calculated by subtracting contralateral from ipsilateral ERPs (contralateral/ipsilateral to the memorized side) of posterior-occipital electrodes across the delay period (Luria et al., 2016; Perez & Vogel, 2012; Vogel & Machizawa, 2004). CDA can also be calculated for MOT tasks with similar procedures (Balaban et al., 2024; Drew et al., 2012).	the number of predicates (Poncet et al., 2016; K. E. Wilson et al., 2012), some other studies argue that CDA can only be modulated by the number of pointers (Luria & Vogel, 2011; Woodman & Vogel, 2008). This tension remains to be fully resolved in future work.	
NSW (negative slow wave) [EEG]	While the classic CDA paradigm is a “lateralized” paradigm, one could also present all to-be-memorized items around the center of the screen (i.e., a wholefield paradigm). NSW can be calculated in both lateralized and wholefield paradigms by averaging the ERPs of posterior-occipital electrodes across the delay period (Diaz et al., 2021; Feldmann-Wüstefeld, 2021; Fukuda et al., 2015; Klaver, Talsma, et al., 1999).	?	?
Posterior alpha power [EEG]	Posterior alpha power can be calculated in both lateralized and wholefield paradigms by averaging the amplitude of alpha power in posterior-occipital electrodes across the delay period (Diaz et al., 2021; Fukuda et al., 2015; Fukuda & Woodman, 2017).	?	?
Alpha traveling wave power [EEG]	Using the anterior-posterior locations on the scalp as a proxy for anterior-posterior neural sources (note however, that the anteriority on the scalp for EEG and the actual anteriority of underlying neural resources do not always align; Lau et al., 2016), one could calculate the power of alpha traveling wave	?	?

	(forward or backward) across the delay period (Zeng et al., 2024). This metric has only been tested with lateralized presentation paradigms.		
RPA (right posterior activity) [MEG]	In both lateralized and wholefield paradigms, RPA is the averaged ERF at right posterior-occipital channels across the delay period (Mitchell & Cusack, 2011; Robitaille et al., 2009, 2010; Yu & Lau, 2025).	Yes? [more evidence is needed] (Yu & Lau, 2025)	?
Posterior parietal responses [fMRI]	Posterior parietal responses is calculated by averaging BOLD signals in posterior parietal voxels across the BOLD time course approximately corresponding to the delay period (Naughtin et al., 2016; Y. Xu & Chun, 2006, 2007). This metric has only been tested with wholefield paradigms.	Yes (Naughtin et al., 2016; Y. Xu & Chun, 2006)	?
mvLoad [EEG]	In both lateralized and wholefield paradigms, a machine-learning multivariate classifier can be trained to tell the number of items for one condition or some trials, and one can test if this classifier can predict the number of items for another condition or the other trials, across the delay period (Adam et al., 2020; Jones, Diaz, et al., 2024; Jones, Thyer, et al., 2024; Thyer et al., 2022).	Yes (Jones, Diaz, et al., 2024; Jones, Thyer, et al., 2024; Thyer et al., 2022)	?

1.5. Distinguishing pointers from arguments

This section deals with a potential confusion between pointers and arguments. Readers for whom the distinction is already clear can skip this section.

In our current paper, reassignability (including the property of being able to be updated and reset) can be understood through the example of computer programming languages, where one can update a pointer (i.e., variable) x_1 , e.g.,

```
x1 = 1
x1 = x1 + 4
```

or reset its value to e.g., 5 directly, or even to a qualitatively different data type, e.g.,

```
x1 = 1
x1 = 5
```

or

```
x1 = 1
x1 = "Hello world!"
```

Notably, these properties are distinct from the primary concerns of the classic literature on role-filler independence (Y. Fan et al., 2021; Hafri et al., 2024; Quilty-Dunn et al., 2023). For example, in the role-filler independence literature, it is believed that we have some already-existing structures, e.g., HITTING(AGENT = __, PATIENT = __), where AGENT is the role of the initiator of an action, while PATIENT is the role of the receiver of an action (Dowty, 1991; Papeo et al., 2024). Different fillers could be flexibly “plugged into” this representation, for example, if you plug in [JOHN, MARY], you will get the representation for JOHN HIT MARY, while if you plug in [SAM, LI], you will get the representation for SAM HIT LI. In these cases, what gets plugged into these structures (JOHN, MARY, SAM, LI) are mental objects -- what would be represented by pointers (O’Reilly et al., 2022; Yu & Lau, 2023). Therefore, this literature discusses the independence between structures with these placeholders (often referred to as “arguments” in the realms of syntax, semantics (Alexander Williams, 2015) and programming languages) and the variables/pointers that get “plugged into” these structures,

which seems to be different from the flexibility of predicates that can be bound to a pointer following our current sense of reassignability.

Notably, how “plugging in” neurally works remains elusive and requires further research. What is also of note is that predicates like COP or THIEF can also be plugged into somewhat similar structures even though they don’t name a specific entity. For example, in cases of general knowledge in semantic memory like COP ARREST THIEF, COP represents a type (i.e., category) instead of a specific token (i.e., instance) of COP. At the same time, one would have no problem representing that, in an unusual situation, a specific cop might get arrested by a specific thief (in this case, the specific cop and the specific thief are mental objects), and this would presumably minimally affect the general semantic knowledge about COP and THIEF types. Another issue is that the type-token distinction extends to relations like HITTING as well: One could represent a type HITTING or a specific HITTING token (Ji & Scholl, 2024). What is the cognitive/neural distinction between types and tokens? One theory proposes that tokens are represented by linking pointers with a type (Bowman & Wyble, 2007; Kanwisher, 1991; Yu & Lau, 2023), while another non-mutually-exclusive theory proposes that types and tokens are in a generative relation: types generate tokens (Prasada, 2016; Prasada & Hall, 2019). Many more theories may exist, which is a fertile ground for research.

Another aspect to highlight is that, although we can store a vast amount of structures and functions with many *arguments* in long-term memory (functions e.g., DIVIDE(DIVISOR = __, DIVIDEND = __) that turn inputs into outputs following a relatively fixed processing routine), the number of *pointers that can be held and computed with at the moment* is extremely capacity-limited for humans. The number of pointers that can be held at the moment can be estimated by the number of distinct items that can be held in simple delay-match-to-sample tasks with

relatively meaningless items as stimuli (e.g., colored squares), where little consultation towards long-term structures and functions would occur. Work in this line has uncovered a surprisingly limited number (≤ 4) of variables that can be held at the moment for visual and auditory objects (Cowan, 2001): load-dependent neural responses would typically plateau beyond this limit (Alunni-Menichini et al., 2014; Vogel & Machizawa, 2004; Y. Xu & Chun, 2006; Yu & Lau, 2025), which has been corroborated by behavioral evidence (Awh et al., 2007). Given that human long-term memory has a rather high capacity (Brady et al., 2008), this bottleneck may be precisely the one that “forces” our moment-by-moment processing to be (partly) serial. Since the capacity of holding and computing with pointers in moment-by-moment processing is limited, one is forced to compute some processes serially instead of at once. For example, consider a function that receives an input string, e.g., *abc*, and outputs another string deterministically in a different order: $\text{REORDER}(\text{ITEM1} = a, \text{ITEM2} = b, \text{ITEM3} = c) = bca$. If a system can only hold at most 2 variables (i.e., pointers) at a moment, it simply cannot run this function at all because running this function would require one to represent three variables “at once”. This system can have a two-argument $\text{SWAP}()$ function that might do some work in this line: $\text{SWAP}(\text{ITEM1} = a, \text{ITEM2} = b) = ba$. For simplicity’s sake, here we assume that $\text{SWAP}()$ can work well with only two variables (a more complex scenario is discussed in the “Further questions” section). In this case, one can either “hack” the system by treating *bc* as a chunk and execute $\text{SWAP}()$ once (which will no longer work if the segment to be chunked gets much longer), or one has to apply $\text{SWAP}()$ multiple times serially to the input string *abc*, in order to get the desired output *bca*. Note that what may complicate the situation even more is that another variable may (or may not) be needed to hold the whole input string *abc* or the whole output string *bca*, apart from the parts *a*, *b* and *c*. Although our limited moment-to-moment capacity for

pointers appears to be sufficient for many tasks in daily life (Halford et al., 1998; so that we rarely realize the existence of this bottleneck), it does explain individual differences in e.g., real-life decision making (Xie et al., 2020) or fluid intelligence (Conway et al., 2002).

1.6. How to test if the same set of pointers is shared across scenarios?

Once we establish a neural signature that is content-free and reflects reassignability, we can use this signature as a hallmark to study whether the same set of pointers is shared across different cognitive-neural systems, paradigms, task goals, and source modalities (Figure 1.1). To name an example of this kind of research question, while some hypothesize that the same set of pointers is recruited for visually-derived and language-derived objects (Brodbeck et al., 2016; Yu & Lau, 2023), others hypothesize that they recruit separate sets of pointers (Brody & Csibra, 2025; Murez et al., 2020). Namely, when we are comprehending linguistic discourses such as stories, we may need to track different numbers of discourse referents in moment-to-moment comprehension. For example, during comprehension of the sentence “John likes singing, and he also likes dancing”, we only need to track one discourse referent representing John, while for the sentence “John likes singing and Mary likes dancing”, we need to track both John and Mary (Figure 1.1B). While some hypothesize that these discourse referents are represented by the same set of pointers that are recruited in for example VWM tasks, others argue that these discourse referents recruit a separate set of pointers.

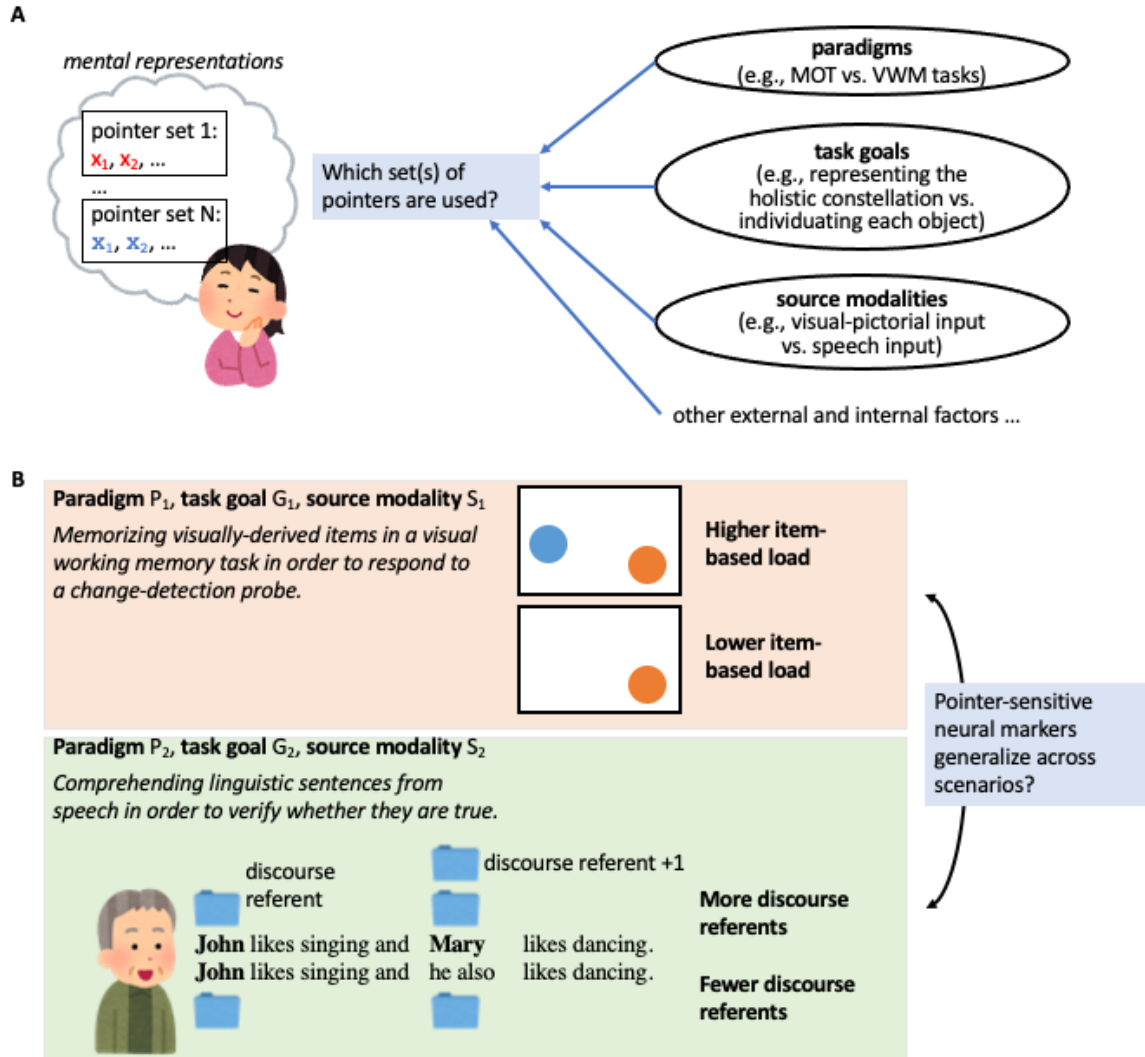


Figure 1.1 Illustration of the proposed research program.

By testing how the neural markers for pointers generalize across different paradigms, task goals, source modalities, etc., we will be able to infer the taxonomy of different pointer systems (panel A). Furthermore, here we illustrate a concrete example of such approach in panel B.

Generalizable pointer-sensitive neural markers across the two scenarios suggest a shared set of pointers, while a failure in generalization may suggest (with caution) separate sets of pointers. The latter argument could be strengthened by double dissociative evidence, where a pointer-sensitive marker is sensitive to object-based load in Scenario A but not B, and another pointer-sensitive marker is sensitive to object-based load in Scenario B but not A. Following this research program, we will be able to not only unravel how many sets of pointers there are, but also characterize how the recruitment of these pointer systems is shaped by external and internal factors. Note that here we leave open the possibility that e.g., an object in one's visual field may elicit more than one pointer, belonging to different pointer systems.

How to test if the same set of pointers is shared across different scenarios? For univariate neural signatures (e.g., CDA, NSW), we can directly examine if a certain neural signature is modulated in a similar way by the number of objects to be represented, across experimental conditions. The neural signatures examined should be content-free and reflect reassignability (i.e., tracking the number of pointers). The neural signatures could still be examined if they reflect reassignability but track both object-based and feature-based load, as long as feature-based load (and other confounding factors, if any) is experimentally controlled for across load conditions. As the quality of known neural signatures remains to be fully tested (Table 1.1), here we provide illustrative examples for how one can test if the same set of pointers is shared across different scenarios. For instance, while visually-derived object load induces a suppression in posterior alpha power (Chen et al., 2024; Fukuda et al., 2015), phonological unit load appears to induce an increase in posterior alpha power (Proskovec et al., 2019), suggesting that these two types of mental objects belong to different natural kinds. As another example, tracking more discourse referents in linguistic discourse may drive a more negative NSW response (Cruz Heredia et al., 2021), similar to memorizing more items in VWM tasks (Fukuda et al., 2015); this could suggest a shared set of pointers for visually-derived mental objects and discourse referents. More evidence is needed in this particular case as discourse referent load does not always drive consistent ERP responses in EEG (Lau, 2018). For multivariate neural signatures, in the case of mvLoad (see Table 1.1), a neural classifier of WM (working memory) load can be trained from one experimental condition, and we can examine if it could decode WM load for another experimental condition (Jones, Diaz, et al., 2024; Jones, Thyer, et al., 2024; Thyer et al., 2022). Representational similarity analysis (RSA) could also be applied to “regress out” the contribution

of predicates (Jones, Diaz, et al., 2024; Jones, Thyer, et al., 2024). For example, Jones, Thyer and colleagues (2024) examined the neural responses for holding colored moving dot clouds during VWM maintenance, manipulating object-based load and the feature to attend to (i.e., color or whether the motion was coherent). To parse out predicates from pointers, apart from a dissimilarity matrix across conditions based on object-based load (pointer-based load), they also created dissimilarity matrices for predicates (e.g., load similarity based on color features and motion coherence features respectively) as well as for other potentially confounding factors (e.g., pupil size). All these dissimilarity matrices were taken into account when evaluating their similarities with the dissimilarity matrix for time-window-by-time-window EEG data. The pointer-based load dissimilarity matrix remained predictive of EEG response patterns across the VWM delay in spite of other dissimilarity matrices, suggesting robust pointer-based neural responses that is indifferent to the task (here, which feature to attend to), and that this shared response was unlikely driven by the other confounding factors considered in the analysis as well. Indirect evidence for whether the same set of pointers are shared across scenarios may also be obtained from behavioral experiments, testing whether retaining two types of objects in WM interfere with each other (Fougnie & Marois, 2011; Morey & Cowan, 2004, 2005; Yu & Lau, 2024).

The scenarios to be tested can span a variety of paradigms, task goals and source modalities, for example, tracking objecthood across saccadic eye movements (Cavanagh et al., 2010; Zhu et al., 2020), tracking objects when they are visually available (Balaban et al., 2023), tracking objects when they are visually unavailable in the case of e.g., temporary occlusion (Balaban et al., 2024), attending to an undetermined object based on exogenous visual cues (e.g., pointing; Brody & Csibra, 2024; Pätzold & Liszkowski, 2019), attending to an undetermined

object based on exogenous linguistic cues such as English demonstratives “this” or “that”, holding visually-derived objects in WM that interacts little with long-term memory (Luria et al., 2016; Vogel & Machizawa, 2004), holding visually-derived objects in WM retrieved from long-term memory (Fukuda & Woodman, 2017), holding non-visually-derived (e.g., auditory, haptic) objects in WM that interacts little with long-term memory, holding non-visually-derived objects in WM retrieved from long-term memory, holding objects in visual imagery (Balaban & Ullman, 2024; Hu & Yu, 2023), comprehending linguistic discourses involving individuals (Brody, 2020; Yu & Lau, 2023), and reasoning about multiple individuals in spontaneous thought.

Moreover, somewhat orthogonally, it is also less clear whether the same set of pointers is recruited for perception and working memory stages, even within the same task environment and for the same modality. This question is particularly interesting given the ongoing debate on the extent to which perceptual and working memory representations are shared (Block, 2023; Fukuda et al., 2022; Teng & Kravitz, 2019). This could also be studied by examining whether the same univariate signals generalize across these temporal stages, and for multivariate neural responses, through decoding or RSA across temporal stages: e.g., training a classifier based on data during the perceptual phase, and test decodability on data during the working memory phase (Carlson et al., 2011; Teichmann et al., 2022; Y. Xu, 2021).

1.7. Object representations beyond the pointer system

Before proceeding with the dissertation, we should make it even more explicit that these pointers are unlikely the only format to represent objects in the world mentally (see also Footnote 2 in this chapter). When one is presented with a visual scene with multiple objects, one may be

representing each of the objects with a pointer, while one may also represent the entire scene in an iconic format (Block, 2023), or represent a certain magnitude feature (e.g., average or total size, location) extracted from the entire scene (Whitney & Yamanashi Leib, 2018). Object representations in long-term memory may have yet a different format. These other formats may exist simultaneously with the pointer representations, participating in different computations within different neuro-cognitive systems.

Furthermore, pointers are also unlikely the only vehicle of symbolic cognition. Cognition can be understood symbolically with mental variables (Quilty-Dunn et al., 2023), and pointers for object representation can be understood as one specific type of mental variables. Different types of mental variables may have slightly different properties, for example, the extent to which they are content-free and reassignable.

1.8. Assigning features to pointers

While pointers serve as binding resource for features, and provide an intuitive account for objecthood after moderate feature change, several relevant questions remain. For example, how do we know at all that the moving pixels in front of us belong to the same object, and how is the mapping between a certain pointer and certain features determined? In other words, how does the neural system “know” that a specific pointer should point to a specific array of features?

As the current investigation primarily focuses on the representation after these questions are resolved neurally (i.e., we focus on the end-result of the abovementioned computations),

answering these questions is beyond the scope of the dissertation⁴. However, these questions are clearly important and intriguing. One thing we can say is that the visual cues determining these abovementioned computations are dissociable from the pointers per se: whether or not pointers are recruited, how many pointers are recruited, as well as how these pointers map onto features are all *outcomes* of the abovementioned computations, thus these computations per se can be dissociated from pointers per se. The abovementioned computations likely rely on visual cues and knowledge such as spatiotemporal coherence (or other Gestalt principles) within an object (Spelke, 1990), as well as whether a certain feature is “essential” or incidental (Macnamara, 1999; for example, a color change can happen within the same chameleon, while seeing an object of a different species clearly suggests separate objecthood). At a gross level, these computations require two inter-related processes: the identification of featural changes, and weighing these featural changes so as to modulate feature-pointer correspondence. It is also very likely that this process involves both relatively “lower-level” Gestalt principles and “higher-level” rich conceptual knowledge about objects and their relations (Yu, Thakurdesai, et al., 2025). Exactly how these processes work cognitively and neurally remains to be fully elucidated by future studies.

1.9. Outline of this dissertation

In this dissertation, we take the initial explorative steps for this unfolding research program. As some prior studies suggest that even for visual inputs, biological and non-biological visual

⁴ That said, effects in early time windows in EEG/MEG studies may reflect online processes relevant to these computations.

objects do not share the same set of pointers, in Chapter 2 we report a line of behavioral experiments demonstrating the opposite, and provide possible explanations to reconcile these divergent prior studies. The general logic for Chapters 3 and 4 is similar: testing whether two types of mental objects are represented by the same set of pointers by examining whether they modulate a shared neural marker hallmarking pointers. Chapter 3 employs the EEG marker of the negative slow wave (NSW) – an evoked and sustained ERP response – to study whether relations between objects are represented by similar pointers as objects per se. Due to the technical limitations of evoked sustained response in tasks with longer epochs, in Chapter 4 we employed the EEG marker of posterior alpha power to study whether visually-induced objects and linguistically-induced ones share the same set of pointers. Chapters 3 and 4 offers initial evidence suggesting a shared set of pointers across visually-derived objects and relations, and across visually-derived and linguistically-derived objects in moment-to-moment cognition. In Chapters 5 and 6, with a series of MEG experiments, we developed a novel neural marker hallmarking pointers in visual working memory tasks -- the right posterior activity (RPA). A mirroring effect was observed in the source space in the inferior parietal regions, suggesting that working memory pointers are hosted there. This neural marker can serve our research program in future studies. Chapter 7 marks my methodological development on OPM-MEG, an emerging neuroimaging technology, as an effort to expand the neuroimaging toolkit available for identifying and characterizing candidate neural hallmarks of mental object representation. Lastly, Chapter 8 offers a summary of all the above research, as well as questions for the next steps in future research.

2. Chapter 2⁵

In this chapter, we report behavioral evidence suggesting that biological and non-biological visual objects share the same set of pointers in working memory (short-term memory).

2.1. Introduction

A ball moves behind an occluder. Now a book comes out from behind the same occluder, before returning behind it. Finally the occluder is removed, but there is only one object there! Although perplexing to adults, young preverbal infants seem not be surprised by this sequence of events, as measured by looking time (Kibbe & Leslie, 2011; F. Xu & Carey, 1996), although they will be surprised if *no* objects are there. This is one striking example from a broad base of infant and adult evidence for the idea that a set of contentless visual pointers (or indexicals, indexes), pointing to features, supports human visual working memory (VWM) and visual object tracking (Baillargeon & Carey, 2012; Brody, 2020; Carey & Xu, 2001; Kahneman et al., 1992; Leslie et al., 1998; Pylyshyn, 1989, 2001; Quilty-Dunn et al., 2023; F. Xu, 1999; Yu & Lau, 2023). In this case, the now-standard interpretation of the infants' behavior is that infants represent the ball moving behind the occluder with an object pointer to "ball" conceptual features, and that when the book emerges, they represent this as the "same" object by changing the pointed features to those of "book" (C. Cheng & Kibbe, 2022; Kibbe & Leslie, 2013).

⁵ This chapter contains content published as Yu, X., & Lau, E. (2025). Same set of visual pointers for biological and non-biological objects in working memory. *Visual Cognition*.

The pointers are hypothesized to be content-free and reassignable (Balaban et al., 2019, 2023; Thyer et al., 2022; Yu, 2025), i.e., that they are independent from the representation of features (which bind features together but do not represent features in themselves), and that they can be “remapped” to bind different features to the same object, or even represent different objects across time. Indeed, the existence of contentless visual pointers in working memory has also been supported by neural evidence in adults (Balaban et al., 2019; Naughtin et al., 2016; Thyer et al., 2022; Y. Xu, 2009; Y. Xu & Chun, 2006; Yu & Lau, 2025; for discussion see Yu & Lau, 2023). For example, Xu and Chun identified subregions in the posterior parietal cortex demonstrating load-dependent, capacity-limited responses during VWM retention with fMRI (functional magnetic resonance imaging); critically, this representation was agnostic to the number of features. This observation was subsequently corroborated by MEG (magnetoencephalography) evidence of a qualitatively similar response over the right posterior cortex (Yu & Lau, 2025). Complementary to the above univariate neural responses, recent work has explored multivariate responses tracking the number of VWM pointers as well (Jones, Diaz, et al., 2024; Jones, Thyer, et al., 2024; Thyer et al., 2022). Moreover, the number of pointers is limited to around 3, as suggested by numerous infant and adult studies using VWM or object tracking tasks (Awh et al., 2007; Cowan, 2001; Feigenson et al., 2002; Luck & Vogel, 1997, 2013; Y. Xu & Chun, 2009). Apart from the limited capacity for pointers, the representation of features is also capacity-limited, as VWM is best modeled with a limited capacity for features *and* a limited set of pointers (Ngiam, 2024; Swan & Wyble, 2014; van den Berg et al., 2014; Xie & Zhang, 2017b)⁶.

⁶ This isn't to say that this pointer system is the only format of VWM; evidence also suggests a complementary system for numerosity estimation and summary statistics (Feigenson, 2011; Knowlton & Gomes, 2022; Kutter et al., 2023; Odic & Starr, 2018).

An intuitive hypothesis would be that these ~ 3 visual pointers support the representation of all kinds of objects, irrespective of e.g., their animacy. Somewhat surprisingly, some recent studies on infants and adults seem to offer evidence against this “same set” hypothesis, prompting the proposal that biological objects recruit a separate set of visual pointers other than those for non-biological objects (the separate set hypothesis; Y. Liu et al., 2019; Murez & Smortchkova, 2014). Relevant infant and adult literatures are often not integrated; therefore here we first provide an overview of these two lines of evidence. We will show that the existing results can be accounted for by either hypothesis, and we then go on to introduce our current experiments designed to resolve the ambiguity in the adult studies.

Infant studies. In Xu and Carey’s (1996) experiments, infants showed insensitivity to changes in the object behind an occluder, just as adults asked to visually track an object in motion are not disrupted by some featural changes (K. Zhou et al., 2010). However, the stimuli used in these infant studies were mostly inanimate. Subsequent studies observed that infants in a similar age range (or earlier) were surprised when the feature change is between biological and non-biological objects (Bonatti et al., 2002; Decarli et al., 2020; Kibbe & Leslie, 2019; Surian & Caldi, 2010). For example, Kibbe and Leslie observed a longer looking time when an inanimate ball changed into a doll head (or vice versa), compared to when the feature change did not change animacy status. This type of observation led Murez and Smortchkova (2014) to propose that biological objects recruit a separate set of object pointers (i.e., object files) from non-biological objects.

However, such evidence is not necessarily inconsistent with the same set hypothesis; one can simply assume that some kinds of object changes prompt a “reset” of visual pointers while others do not, and that these pointer resetting parameters are subject to change across the course

of development. For example, in adult multiple object tracking studies, it has been found that some feature changes do not affect behavioral accuracy (suggesting the same pointer represents the object before and after the feature change), while other feature changes do hurt accuracy (Lyu et al., 2015; Zhao et al., 2020; K. Zhou et al., 2010). The process of pointer resetting is associated with a disruption in CDA (contralateral delay activity) in EEG (electroencephalography) which tracks the number of visual pointers in VWM (Balaban et al., 2023, 2024). In sum, infant evidence, although previously argued to support the separate set hypothesis, is in fact equally consistent with the same set hypothesis.

Adult studies. A different way to evaluate these hypotheses in adults is to examine whether the working memory capacity for non-biological objects is shared with the capacity for biological objects, as predicted by the same set hypothesis. A number of studies using the delay-match-to-sample paradigm have observed that holding more biological motions (BMs) in working memory does not appear to interfere with the capacity limit for non-biological objects in working memory, which might seem to support the separate set hypothesis. We note, however, that several of these studies presented the sample objects sequentially at the same location (Y. Liu et al., 2019; Wood, 2007, 2011, Experiment 4-5). Continuity of location is a strong cue for the objective continuation of the same object across time, and thus against pointer resetting (Bapat et al., 2017; Golomb et al., 2014). Therefore, sequential, same-location presentation may encourage participants to represent the input as feature change within a single object pointer, rather than as multiple distinct objects. In keeping with this, fMRI work has shown that responses associated with pointer maintenance only minimally increase with set size in sequential, same-location presentation, compared to simultaneous different-location presentation (Harrison et al., 2010; Y. Xu & Chun, 2006).

Perhaps the strongest current evidence for the separate-set hypothesis comes from two experiments by Y. Liu et al. (2019) which used simultaneous different-location presentation (Exp. 2 and 3). In these two experiments, Liu and colleagues found that memorizing more BMs does not affect the capacity for color-shape bindings and vice versa. This result is consistent with the hypothesis that biological objects recruit a separate set of pointers from non-biological objects. However, another property of the experiment design provides an alternative explanation. In these experiments, the “foil” memory probes always used an object that had not appeared anywhere in the sample display (“object-change trials”) rather than using an “old” object that had been in the sample display but that was probed in the wrong location (“location-change” trials). This leaves open the possibility that participants could succeed at rejecting the foil trials without binding the BM features to object pointers at all, but simply by noting whether the probe’s BM feature had been detected recently (cf. Oberauer, 2017). In other words, in order to correctly respond to object-change trials, one does not need to represent BMs as bound objects, but just as a conceptual feature. And as the point-light displays (BM) and color-shape bindings recruit very different feature spaces, confusion between the features themselves would be unlikely.

Fortunately, the use of location-change trials provides a straightforward way of probing the pointers for binding. If the probe is an old object (present in the sample) but at a wrong location previously occupied by another object (“location-change trials” hereafter), one cannot respond to these probes correctly by the conceptual features only; rather, correct response for location-change probes relies on the representation of conceptual feature-location *bindings*. The same set hypothesis predicts an interference effect when examining these location-change trials. In the current study we put this to the test. In Experiment 2.1, we conceptually replicated the results of Y. Liu et al. (2019) with another set of inanimate bound objects, complex shapes. In

Experiment 2.2 and Experiment 2.3 (a replication), we introduced location-change trials to provide the crucial comparison. If biological and non-biological object representation relies on the same set of pointers, memory interference effects should be observed in these trials; the separate set hypothesis again predicts no interference.

2.2. Results

In Experiment 2.1 (Figure 2.1), we aimed to replicate previous results on biological and non-biological object maintenance using another set of bound inanimate objects: complex (multi-featured) shapes. Here we used the same object-change probes as previous experiments, but rather than asking participants to “judge whether the probe had appeared in the memory array” (Y. Liu et al., 2019), we asked them to judge whether the probe matched the object that had been presented at that location, in order to encourage the representation of bound objects rather than free features.

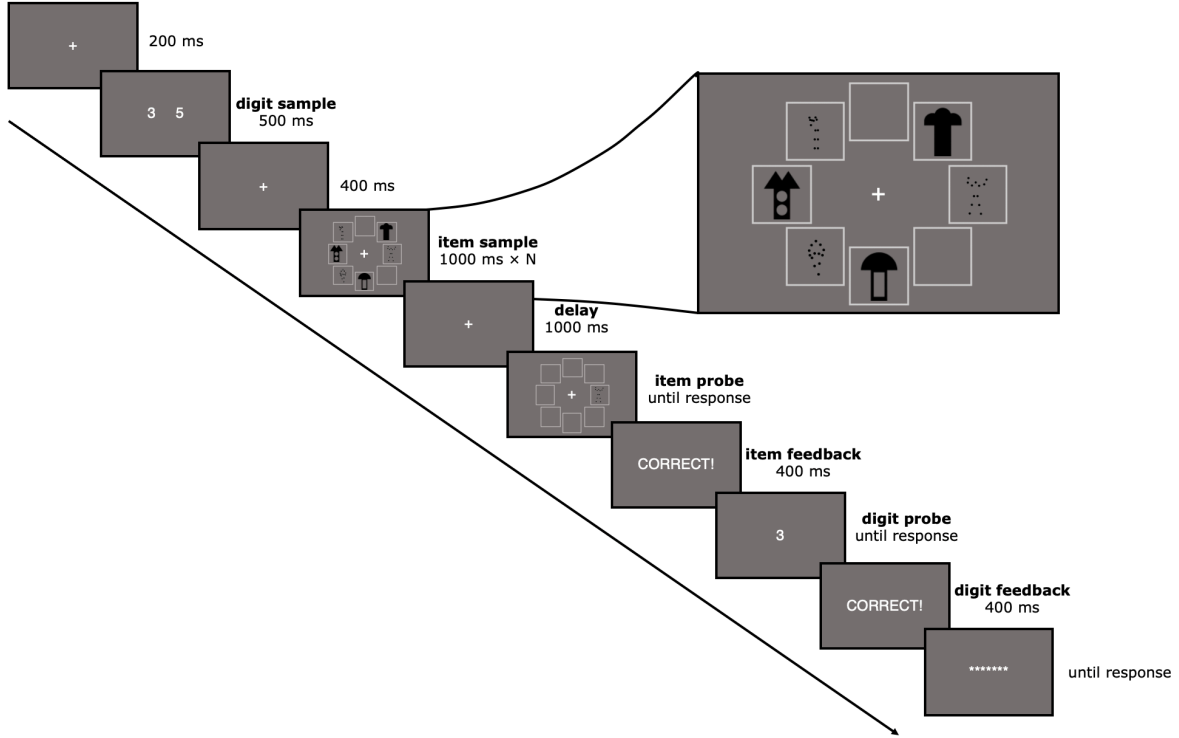


Figure 2.1 Illustration of a trial in Experiments 2.1-2.3.

2.2.1. Experiment 2.1

Subjects reached ceiling at the digit task (mean accuracy = 98%). In terms of Cowan's K, we observed a significant difference in K_{ACT} across the two conditions, $W = 0$, $z = -4.3$, $p < 0.001$ (K_{ACT} for 3OBJ+1ACT: $M \pm SE = 0.83 \pm 0.04$; K_{ACT} for 3OBJ+3ACT: $M \pm SE = 1.75 \pm 0.10$). The difference in K_{OBJ} across the two conditions was only marginally significant, $W = 184.5$, $z = 1.8$, $p = 0.06$ (K_{OBJ} for 3OBJ+1ACT: $M \pm SE = 1.39 \pm 0.16$; K_{OBJ} for 3OBJ+3ACT: $M \pm SE = 1.16 \pm 0.17$). See Figure 2.2.

In Experiment 2.1 we replicated Y. Liu et al.'s (2019) observation of a non-interference effect, i.e., an increase in the Cowan's K for BMs held in working memory does not significantly discount the Cowan's K for complex shapes. However, this result may be driven by the fact that

in both our Experiment 2.1 and Y. Liu et al. (2019), the probe was always a new object (an object that was not in the sample). Although multi-featured complex shapes (our Experiment 2.1) and color-shape bindings (Y. Liu et al., 2019) would be represented as bound objects, meaningful BMs can be represented as a conceptual *feature* only (action type, e.g., JUMPING, WAVING), without recruiting the pointers to bind multiple features together. Therefore, in Experiment 2.2 and Experiment 2.3, we added trials in which the probe is an old object presented at a wrong location previously occupied by another object (Figure 2.3). These trials cannot be correctly responded to by just holding the conceptual feature, enabling us to examine the Cowan's K for conceptual feature-location *bindings*. Note that the marginally significant effect for K_{OBJ} may reflect a true effect but with a small effect size. This is in line with our argument that biological and non-biological objects share the same set of VWM pointers.

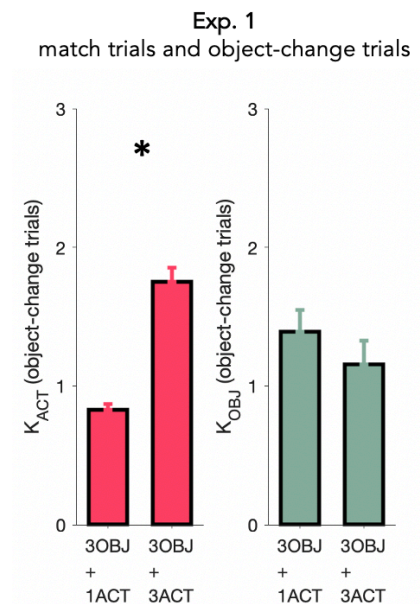


Figure 2.2 The results for our Experiment 2.1.

K_{ACT} is the Cowan's K for BMs, and K_{OBJ} is the Cowan's K for static shapes. 3OBJ+1ACT is the condition where 3 static shapes and 1 BM were in the sample, and 3OBJ+3ACT is the condition

where 3 static shapes and 3 BMs were in the sample. *: $p \leq 0.05$. Error bars stand for standard error.

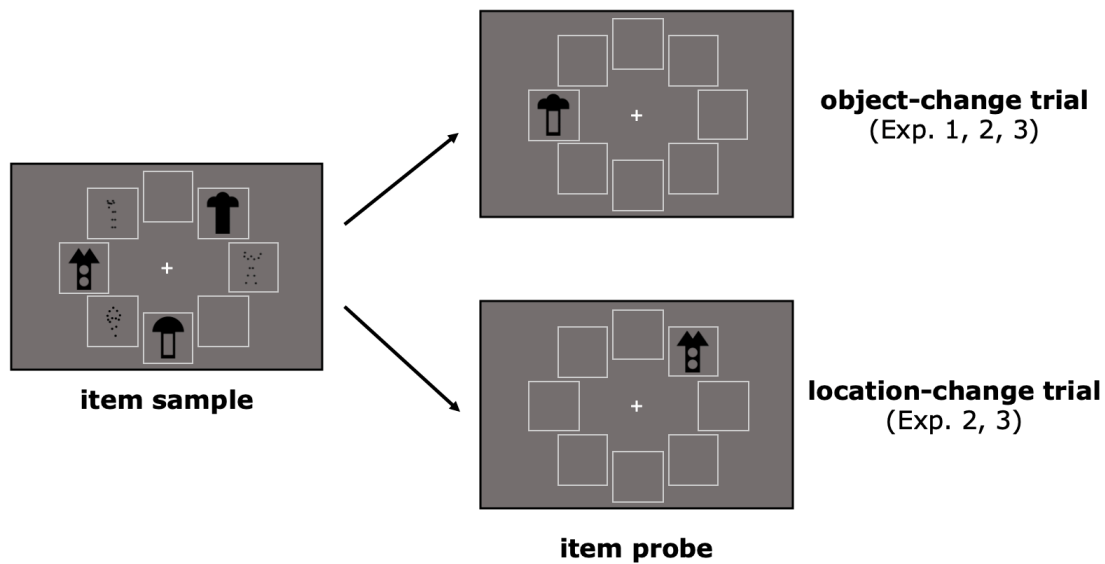


Figure 2.3 Illustration of the difference between an object-change trial and a location-change trial.

In an object-change trial, the probe was a new object that didn't appear in the sample; in a location-change trials, the probe as an old object at a wrong location previously occupied by another sample object.

Experiments 2.2 and 2.3 were very similar, both including the crucial location-change trials; Experiment 2.3 was essentially a replication of Experiment 2.2. The only difference between these experiments was the number of trials allocated to each condition, as described below.

2.2.2. Experiment 2.2

Subjects reached ceiling at the digit task (mean accuracy = 99%). For the match trials and the object-change trials, there was a significant difference between the K_{ACT} across the two

conditions, $W = 18$, $z = -3.8$, $p < 0.001$ (K_{ACT} for 3OBJ+1ACT: $M \pm SE = 0.79 \pm 0.05$; K_{ACT} for 3OBJ+3ACT: $M \pm SE = 1.37 \pm 0.12$). Meanwhile, there wasn't a significant difference between the K_{OBJ} across the two conditions, $W = 169$, $z = 1.4$, $p = 0.17$ (K_{OBJ} for 3OBJ+1ACT: $M \pm SE = 1.28 \pm 0.11$; K_{OBJ} for 3OBJ+3ACT: $M \pm SE = 1.08 \pm 0.15$). See Figure 2.4.

For the match trials and the location-change trials, there was a significant difference between the K_{ACT} across the two conditions, $W = 3$, $z = -4.2$, $p < 0.001$ (K_{ACT} for 3OBJ+1ACT: $M \pm SE = 0.91 \pm 0.04$; K_{ACT} for 3OBJ+3ACT: $M \pm SE = 1.93 \pm 0.13$). There was also a significant difference between the K_{OBJ} across the two conditions, $W = 206$, $z = 2.6$, $p = 0.01$ (K_{OBJ} for 3OBJ+1ACT: $M \pm SE = 1.90 \pm 0.10$; K_{OBJ} for 3OBJ+3ACT: $M \pm SE = 1.61 \pm 0.14$). See Figure 2.4.

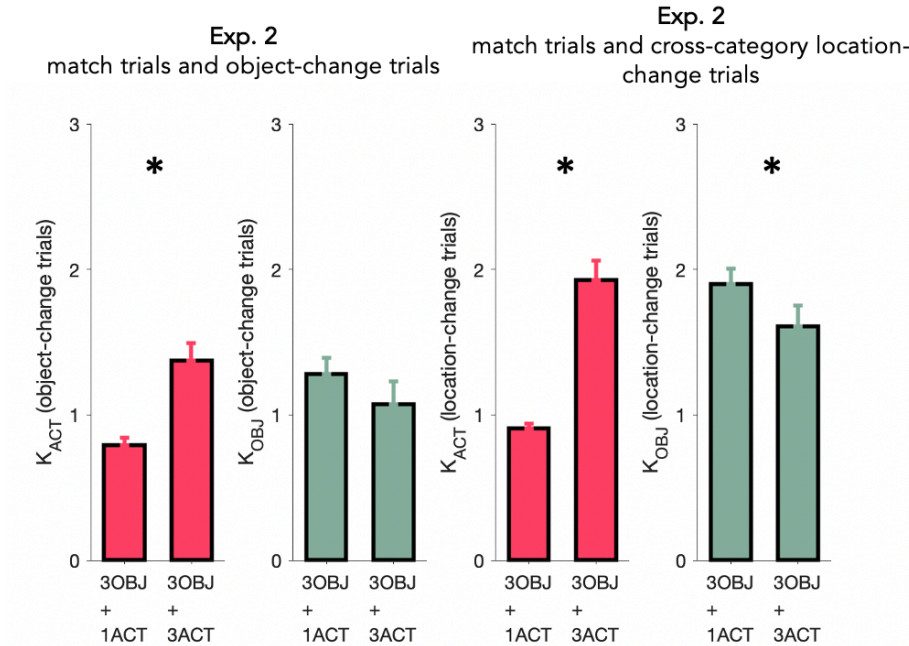


Figure 2.4 The results for our Experiment 2.2.

*: $p \leq 0.05$. Error bars stand for standard error.

2.2.3. Experiment 2.3

Subjects reached ceiling at the digit task (mean accuracy = 98%). For the match trials and the object-change trials, there was a significant difference in K_{ACT} across the two conditions, $W = 21.5$, $z = -3.7$, $p < 0.001$ (K_{ACT} for 3OBJ+1ACT: $M \pm SE = 0.80 \pm 0.03$; K_{ACT} for 3OBJ+3ACT: $M \pm SE = 1.33 \pm 0.12$). Meanwhile, there was no significant difference in K_{OBJ} across the two conditions, $W = 155.5$, $z = 1.4$, $p = 0.17$ (K_{OBJ} for 3OBJ+1ACT: $M \pm SE = 1.07 \pm 0.12$; K_{OBJ} for 3OBJ+3ACT: $M \pm SE = 0.91 \pm 0.12$). See Figure 2.5.

For match trials and location-change trials, there was a significant difference in K_{ACT} across the two conditions, $W = 9$, $z = 4.0$, $p < 0.001$ (K_{ACT} for 3OBJ+1ACT: $M \pm SE = 0.86 \pm 0.03$; K_{ACT} for 3OBJ+3ACT: $M \pm SE = 1.72 \pm 0.12$). There was also a significant difference in K_{OBJ} across the two conditions, $W = 202$, $z = 1.9$, $p = 0.05$ (K_{OBJ} for 3OBJ+1ACT: $M \pm SE = 1.50 \pm 0.16$; K_{OBJ} for 3OBJ+3ACT: $M \pm SE = 1.25 \pm 0.14$). See Figure 2.5.

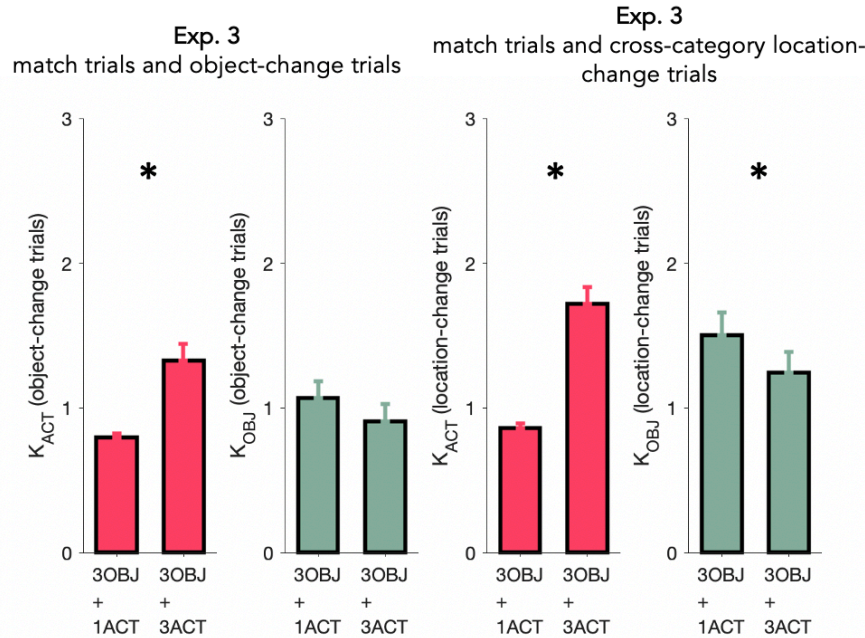


Figure 2.5 The results for our Experiment 2.3.

*: $p \leq 0.05$. Error bars stand for standard error.

2.2.4. GLME analysis for Experiments 2.2 and 2.3

In order to implement a direct comparison between the significant and non-significant effects observed in Experiments 2.2 and 2.3 for K_{OBJ} , we then implemented a generalized linear mixed-effect (GLME) model (cf. Yu, Thakurdesai, et al., 2025) for trial-by-trial results from Experiments 2.2 and 2.3, testing whether there was an interaction effect between the trial type and BM load on accuracy. Indeed, this interaction effect was statistically significant ($\beta = -0.33$, $SE = 0.13$, $Z = -2.52$, $p = 0.012$). The main effect of trial type was significant ($\beta = 1.10$, $SE = 0.13$, $Z = 8.22$, $p < 0.001$), and the main effect of the number of BMs was not ($\beta = -0.01$, $SE = 0.05$, $Z = -0.21$, $p = 0.83$). Based on the interaction effect, we separately analyzed object-change trials and location-change trials. We found that a main effect of BM load only manifested in location-change trials ($\beta = -0.38$, $SE = 0.13$, $Z = -3.13$, $p = 0.004$), but not in object-change trials ($\beta = -0.01$, $SE = 0.05$, $Z = -0.22$, $p = 0.83$). The overall accuracies for complex shapes across all conditions were 75.4% (1 BM, object-change trials), 75.1% (3 BM, object-change trials), 92.0% (1 BM, location-change trials), and 85.7% (3 BM, location-change trials). This complementary analysis strengthened the case that the interference across BMs and complex shapes was more pronounced for location-change trials than object-change trials.

2.3. Discussion

Across three experiments using a new set of complex shapes, we replicated the pattern observed previously: in object-change trials, the number of BMs to be remembered does not affect memory capacity for inanimate objects. However, under these conditions the task can be solved without using the limited set of pointers to bind BMs, by instead just coding the motions as

unbound features of the scene (e.g., JUMPING, WAVING). Crucially, when we used location-change trials in Experiments 2.2 and 2.3 that preclude this solution by requiring feature-location *bindings*, we now observed an interference effect between the number of BMs and inanimate objects, exactly as the same set hypothesis would predict. A task-based explanation of the previous results can also help explain why some previous behavioral studies have observed interference between biological and non-biological objects, if variations in task parameters encourage binding vs. feature-based coding to different degrees (C. Li et al., 2023; X. Lu et al., 2019; Shen et al., 2014). Apart from task parameters (e.g., types of probes, and the proportion of trials with different types of probe), other factors including task instructions and individual preferences may also affect the extent to which binding is encouraged (Cao & Deouell, 2023; Qian et al., 2019; Vergauwe & Cowan, 2015); the contributions of these factors and their interactions remain to be fully established in the future. In sum, the current results provide novel support for the hypothesis that biological and non-biological objects recruit the same set of visual pointers in working memory. This allows a simpler VWM architecture in which a single set of pointers can be used to bind and track visual objects regardless of their content.

2.3.1. Implications for shared and distinct representations for biological vs. non-biological objects

Our observation that a shared set of VWM pointers support the representation for both biological and non-biological objects is in interesting contrast with a line of research suggesting that biological and non-biological objects recruit rather different neural substrates. For example, it has recently been proposed that biological stimuli recruit a separate visual pathway (Küçük et al., 2024; Pitcher & Ungerleider, 2021), compared to non-biological ones. Our findings is not at

odds with this line of research, since the *features* for biological and non-biological objects could indeed be represented by very different neural bases, given that these features appear to occupy rather different sub-spaces in the feature space, across visual and more conceptual features (Capitani et al., 2003; Carlson et al., 2013; Grootswagers et al., 2019; Hebart et al., 2020; Long et al., 2017; Spriet et al., 2022; R. Wang et al., 2022). In this context, the shared pointers allow for flexible binding of very distinct features to the same object: For example, if an object resembling a plane is approaching, and at a certain point as it comes closer you realize that it is actually Superman (Kahneman et al., 1992), one would be able to bind INANIMATE and ANIMATE sequentially to the same pointer without the need to treat the later-recognized animate object as a brand new object.

2.3.2. Connections with prior literature on feature binding

Our current study is consistent with prior observations suggesting that dual-task cost for working memory performance is more pronounced when feature binding instead of only the representation of features was probed (Wheeler & Treisman, 2002). For example, holding visual and auditory features simultaneously in VWM induces little interference with each other, while a more pronounced interference is observed when probing the capacity for bindings across visual/auditory features and location (Fougnie et al., 2015; Fougnie & Marois, 2011). Taken together, although the representation of features may be largely domain-specific and thus take up non-overlapping working memory resources (Fougnie et al., 2015; B. Wang et al., 2017), a domain-general binding resource (i.e., pointers) may bind features from a variety of representational systems. Whether or not we only have one set of pointers, and what each set of

pointers covers (in terms of the types of features) remains to be fully charted (Yu, 2025; Yu & Lau, 2023). Our current study represents one attempt in this line.

2.3.3. Limitations and future directions

One limitation of our current study is that, for location-change trials, the interference effect across BMs and complex shapes may be explained by a limited VWM capacity for *locations*, instead of pointers. This in fact touches on two competing conceptions of the nature of VWM pointers. One conception is that VWM pointers are just locations (Treisman, 1998). The other conception is that VWM pointers are truly content-free, such that location is not qualitatively different from other features such as color and orientation, in the sense that they are all represented independently from pointers (Swan & Wyble, 2014; Wei et al., 2024; Yu, 2025).

While some studies have attempted to dissociate the contributions of location load and object load in VWM on neural responses (Jones, Diaz, et al., 2024; Thyer et al., 2022; Y. Xu & Chun, 2006), this endeavor remains challenging given that location is perhaps one of the strongest cues towards objecthood. For example, flashing two objects sequentially at the same location (Jones et al., 2024a) could be mentally interpreted as one single object changing its features, or as one object replacing another. In the real world, an object at a certain location can indeed change its other features while remaining the same object (e.g., the same chameleon changing its color), or one object can take the place of another at the same location (e.g., the chameleon goes away and a parrot took its place). Therefore, whether two or one pointers are recruited when two objects are presented sequentially, likely depends on e.g., task parameters,

including the type and extent of feature change (B. Park et al., 2020; K. Zhou et al., 2010), as well as individual differences.

2.4. Materials and methods

2.4.1. Experiment 2.1

2.4.1.1. Participants

24 subjects (17 female; age 18-30, $M \pm SD = 21.5 \pm 3.7$) participated in Experiment 2.1, matching the sample size in Y. Liu et al. (2019). All subjects reported having normal or corrected-to-normal vision. Informed consent was obtained from all participants and they received monetary reimbursement for their participation. Procedures were approved by the UMCP IRB Office.

2.4.1.2. Stimuli

Following Y. Liu et al. (2019), 7 point-light BMs from Vanrie & Verfaillie (2004) were used (waving, walking, spading, painting, jumping, cycling, chopping); the points were in black. For inanimate bound objects, 9 black complex shapes similar to those in Y. Xu & Chun (2006) were used.

2.4.1.3. Procedure

Each subject saw 168 trials in the main experiment. In 80 trials, the item sample consisted of 3 static shapes and 1 BM (3OBJ+1ACT). In another 80 trials, the item sample consisted of 3 static

shapes and 3 BMs (3OBJ+1ACT). Of the 80 trials in each condition, 20 trials were probed with a static shape at the same location as in the sample (match), 20 trials were probed with a static shape not in the item sample (non-match), 20 trials were probed with a BM at the same location as in the sample (match), and 20 trials were probed with a BM not in the item sample (non-match). Non-match shape probes replaced the location of another static shape, and non-match biological probes replaced the location of another BM. An additional 8 filler trials consisted of 3 static shapes and 2 BMs. Each subject went through at least 12 practice trials before the experiment.

The experiment was run in MATLAB 2020b with Psychtoolbox 3 (Brainard, 1997; Kleiner et al., 2007) using customized script on a 13-inch MacBook Pro. Like Y. Liu et al. (2019), we included a digit task to occupy participants' phonological loop. At trial onset, a fixation cross was centrally presented center for 200 ms on a gray background. Then, two digits randomly sampled from 1 to 9 were centrally presented for 500 ms. Subjects were instructed to repeat the two digits aloud until a line of asterisks indicated the end of a trial. After another 400 ms fixation screen, the item sample appeared. Eight light grey squares (each $1^\circ \times 1^\circ$) were equally distributed at 0° , 45° , 90° , ..., 315° around an invisible circle with a radius of 1.5° centered on the screen. The squares were employed to separate the objects and discourage visual grouping. The static shapes and BMs were presented at the center of the squares, within a visual angle of $0.9^\circ \times 0.9^\circ$. Following Y. Liu et al. (2019) the duration of item sample was $1000 \text{ ms} \times N$ (the number of items) in order to ensure enough encoding time. Thus for the 3OBJ+1ACT condition, the item sample was presented for 4000 ms; for the 3OBJ+3ACT condition, the item sample was presented for 6000 ms. The BMs were presented on a loop with 12 frames across 1s. In a few rare trials ($<0.5\%$ of all experimental trials), a technical glitch resulted in an inaccurate duration

of the item sample; these trials were excluded from further analysis. After a 1000 ms delay, a probe item (also within the range of $0.9^\circ \times 0.9^\circ$) appeared. Subjects were asked to respond whether the probe item matched with the item presented before at the same location, by pressing m (match) or n (non-match) using their right hand. Feedback on their response (correct/incorrect) was presented for 400 ms. Then, a digit probe was presented at the center of the screen and participants responded whether it matched with one of the two digits presented before (50% possibility), pressing m (match) or n (non-match) using their right hand. Response feedback was then presented for 400 ms. A line of asterisks ended the trial; subjects pressed the space bar to proceed to the next trial. Note that neither the response to the item probe nor the digit probe was time-constrained, different from Y. Liu et al. (2019). This is because the formula for Cowan's K does not take into account no-response trials (Cowan, 2001; Rouder et al., 2011). See Figure 2.1 for a schematic trial illustration.

2.4.1.4. Data analysis

Cowan's K. Memory task performance was quantified with Cowan's K (Cowan, 2001; Rouder et al., 2011), an indicator of the number of items held in working memory. The accuracy for the digit task was calculated as the proportion of correct responses among the rest of the 160 trials. Then for each subject, four Cowan's Ks were calculated: the Cowan's K for static shapes (K_{OBJ}) in the 3OBJ+1ACT condition, the Cowan's K for BMs (K_{ACT}) in the 3OBJ+1ACT condition, the K_{OBJ} in the 3OBJ+3ACT condition, and the K_{ACT} in the 3OBJ+3ACT condition. The Cowan's K formula for single probe experiments is $K = N \times (H - F)$, where N is the number of sample items, H is the hit rate for detecting the change in the probe on non-match trials, and F is the false alarm rate (erroneously "hallucinating" a change in the probe on match trials). For example, K_{OBJ} for

the 3OBJ+1ACT condition was calculated using shape-probe trials which had 3 static shapes in the sample. Thus, $N = 3$, H is the proportion of “non-match” responses among non-match trials, and F is the proportion of “non-match” responses in match trials. K_{ACT} for the 3OBJ+1ACT condition was calculated using motion-probe trials which had 1 BM in the sample. Thus, $N = 1$, H is the proportion of “non-match” responses among non-match trials, and F is the proportion of “non-match” responses in match trials. The other two K s were similarly calculated. Planned two-tailed paired-sample t tests were administered between the K_{OBJ} across the two conditions (i.e., 3OBJ+1ACT vs. 3OBJ+3ACT) and the K_{ACT} across the two conditions. Wilcoxon signed-rank tests were run in JASP 0.16.1 (JASP Team, 2022); a non-parametric test was employed because one distribution only marginally passed the Shapiro-Wilk normality test in this experiment ($p = 0.1$).

2.4.2. Experiments 2.2 and 2.3

2.4.2.1. Participants

Another 24 subjects (18 female; age 18-29, $M \pm SD = 20.9 \pm 3.1$) participated in Experiment 2.2; yet another 24 subjects (14 female, age 18-27, $M \pm SD = 20.2 \pm 2.4$) participated in Experiment 2.3. The sample size of Experiment 2.3 was sufficient to detect the K_{OBJ} difference for location-change trials observed in Experiment 2.2 with $> 80\%$ power ($\alpha = 0.05$; G*Power 3.1, Faul et al., 2007). All subjects reported having normal or corrected-to-normal vision. Informed consent was obtained from all participants and they received monetary reimbursement or course credit for participation. Procedures were approved by the UMCP IRB Office.

2.4.2.2. Stimuli

Same as Experiment 2.1.

2.4.2.3. Procedure

In addition to the 168 trials in Experiment 2.1, we added 80 non-match trials where the probe had appeared in the sample but in a different location (40 trials in each sample condition; half shape probes and half motion probes). The only other change from Experiment 1 was that this time our instructions explicitly requested subjects to use the digit task to help them avoid verbally labelling and rehearsing the items, in order to further reduce the contribution of the articulatory loop.

The only difference between Experiment 2.2 and 2.3 is the allocation of different types of trials within the 80 non-match trials. In Experiment 2.2, 40 non-match probes were old biological action, and 40 were old complex shape, presented at an old but wrong location. Because we selected that wrong location randomly, and because in the 1ACT+3OBJ condition there were more object locations than action locations, the number of cross-category location-change trials to calculate K_{ACT} varied between 2 and 10 for the 1ACT+3OBJ condition; for three subjects there were only a few (2 or 3) cross-category location-change trials with which to calculate K_{ACT} for the 1ACT+3OBJ condition⁷. Therefore, in Experiment 2.3 we replicated Experiment 2.2 but this time fixed the number of different types of trials based on the expectancy for random swapping. For the 1ACT+3OBJ condition, the number of location-change trials from a complex shape to a biological action was 10, the number of location-change trials from a biological action

⁷ Removing these three subjects did not qualitatively change the results.

to a complex shape was 10, and the number of within-category location-change trials for complex shapes was 20. For the 3ACT+3OBJ condition, the number of location-change trials from a complex shape to a biological action condition was 12, the number of location-change trials from a biological action to a complex shape was 12, the number of within-category location-change trials for complex shapes was 8, and the number of within-category location-change trials for BMs was 8.

2.4.2.4. Data analysis: Experiment 2.2

Cowan's K. As in Experiment 2.1, we excluded from analysis a small number of trials with inaccurate timing (<0.5% experimental trials across the entire dataset). For Cowan's K, we first calculated K_{ACT} and K_{OCT} with the match trials and the object-change trials as in Experiment 2.1. Then, we calculated K_{ACT} and K_{OCT} with the match trials and the newly-added location-change trials. The absolute value of Cowan's K can be affected by the similarity between the objects before and after change (Awh et al., 2007; Brady & Alvarez, 2015; Broers & Busch, 2021; Fukuda et al., 2010; Gao et al., 2013); therefore, in order to match for this across Ks, we only used location-change trials where the change is cross-category (i.e., a BM changing into a complex shape or vice versa). K_{OBJ} was calculated with match trials where the probe was a complex shape, and location-change trials where a complex shape changed to a BM (3OBJ+1ACT: 20 trials/subject; 3OBJ+3ACT: on average 11.8 trials/subject). K_{ACT} was calculated with match trials where the probe was a BM, and location-change trials where a BM changed to a complex shape (3OBJ+1ACT: 6.2 trials/subject; 3OBJ+3ACT: 11.7 trials/subject).

GLME (generalized linear mixed-effect) model analysis. Complementarily, we used a GLME model to fit trial-level data to predict participants' trial-by-trial VWM performance

(coded 1 for correct, and 0 for incorrect responses), for data in Experiments 2 and 3 (where both object-change trials and location-change trials were present) collectively. In this model, participants' overall likelihood of recognition success ($P_{success}$) across trials can be directly modeled based on a linear combination of trial-level predictors with a logistic link,

$$\ln\left(\frac{P_{success}}{1 - P_{success}}\right) \sim \beta_0 + \beta_{trial_type} + \beta_{BM_load} + \beta_{trial_type \times BM_load} + 1|(\text{subject: experiment})$$

, where β_{trial_type} captures the variance accounted for by trial type (object-change trial vs. location-change trial), β_{BM_load} captures the variance accounted for by BM load (1 or 3). $\beta_{trial_type \times BM_load}$ stands for the interaction between trial type and BM load. Subject and Experiment (Experiment 2.2 vs. Experiment 2.3) were treated as random effects. This analysis was performed on the same trials that entered the Cowan's K analyses. BM load was standardized before regression. Based on the significant interaction effect between trial type and BM load, we conducted separate regression analyses for data from each trial type (i.e., object-change trials and location change trials). Statistical significance was evaluated using Wald's Z test, and all reported p-values are two-tailed.

2.4.2.5. Data analysis: Experiment 2.3

Cowan's K. Same as Experiment 2.2. As in Experiments 2.1 and 2.2, we excluded from analysis a small number of trials with inaccurate timing (0.5% experimental trials across the entire dataset).

GLME model analysis. Detailed in the previous section.

3. Chapter 3⁸

In this chapter, we report EEG evidence suggesting that relations between objects appear to recruit pointers in working memory similar to objects per se.

In this study, we also employed a new paradigm inspired by the “pinging paradigm” (e.g., Wolff et al., 2017), with the hope of highlighting the representation of the agent and patient respectively so as to disentangle neural responses for AGENTNESS vs. PATIENTNESS. For example, during the maintenance of the event “the elephant punched the lion” in short-term memory, we hope that flashing the word “ELEPHANT” on the screen would highlight the representation of the agent object and the AGENT role, and that flashing the word “LION” on the screen would highlight the representation of the patient object and the PATIENT role. Comparing across trials where the agent vs. patient object is pinged, any systematic difference would reflect differential representations of the AGENT role (i.e., AGENTNESS) vs. the PATIENT role (PATIENTNESS), as each animal can be agent or patient across trials. In the current chapter, we only analyzed ERP differences based on two relatively known ERP effects (CPP and NSW). Although these analyses did not reveal significant differences across AGENTNESS vs. PATIENTNESS, this does not exclude a potential neural difference that may be captured by other univariate ERP effects or multivariate analyses. This is clearly an intriguing direction and requires further examination.

⁸ This chapter contains content published as 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, 1282869. Jialu Li, Hao Zhu and Xing Tian contributed to data collection. We would like to thank Matti Laine for sharing the stimuli from Hultén et al. (2014).

3.1. 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.

3.1.1. The neural object indexical theory

3.1.1.1. 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 (Brody, 2020; Carey, 2009; Green & Quilty-Dunn, 2021; Kahneman et al., 1992; Pylyshyn, 1989, 2001; Quilty-Dunn et al., 2023; Thyer et al., 2022; Xie & Zhang, 2017a; Y. Xu & Chun, 2007; Zhu et al., 2020; for a review see Yu & Lau, 2023). The number of discrete object indexicals is limited, classically suggested to have a limit of around 4 (Awh et al., 2007; Cowan, 2001; Luck & Vogel, 1997, 2013; Ngiam, Foster, et al., 2023; Y. Xu & Chun, 2009; W. Zhang & Luck, 2008). 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 & 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 (Balaban et al., 2019; Y. Xu & Chun, 2007, 2009). 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 (Knops et al., 2014; Matsuyoshi et al., 2010; Mitchell & Cusack, 2008; Robitaille et al., 2010; Song & Jiang, 2006; Todd & Marois, 2004; Y. Xu & Chun, 2006) 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 (Naughtin et al., 2016; Y. Xu & Chun, 2006); rather, these indexicals connect to features represented in other regions, as illustrated in Figure 3.1⁹ (Naughtin et al., 2016; Y. Xu, 2007). This theory is also consistent with the proposal that the posterior parietal cortex is crucial to working memory in a more general sense (Bettencourt & Xu, 2016; Jia et al., 2021; for a review see Y. 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, 2024). 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 &

⁹ 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 & 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 3.3.1).

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 (Hedayati et al., 2022; Matthey et al., 2015; Schneegans & Bays, 2017; Swan & Wyble, 2014; see Schneegans & 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.

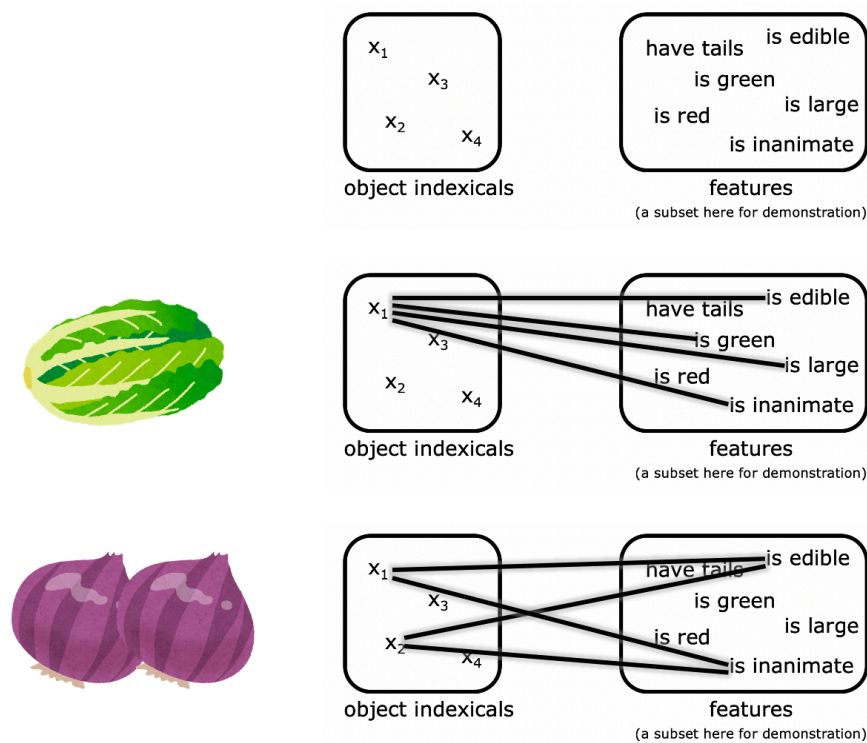


Figure 3.1 An illustration of the representations proposed by neural object indexical theory.

The object indexicals are hosted in some subregions of the posterior parietal cortex, while the features are represented in other regions. The pictures of the Napa cabbage and onions were adapted from IRASUTOYA (<https://www.irasutoya.com/>), which is free for non-commercial use.

3.1.1.2. 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 & Machizawa, 2004). In CDA studies, multiple visual objects are presented on both sides 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., Cai et al., 2022; Gao et al., 2013; Green & Quilty-Dunn, 2021; Hakim et al., 2019).

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 (Diaz et al., 2021; Fukuda et al., 2015; Klaver, Smid, et al., 1999; Ruchkin et al., 1992; for a review see Feldmann-Wüstefeld, 2021)¹⁰. 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 (Diaz et al., 2021; Feldmann-Wüstefeld, 2021; Fukuda et al., 2015; Klaver, Smid, et al., 1999; Mecklinger & Pfeifer, 1996; Ruchkin et al., 1992), and reaches a plateau beyond 3-4 items (Feldmann-Wüstefeld, 2021; Fukuda et al., 2015; L. Zhou & Thomas, 2015), just like the CDA¹¹. Therefore, here we use the magnitude of the NSW as an index of working memory load related to the object indexical system.

3.1.2. 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

¹⁰ 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., Y. Li & Noguchi, 2022; Perez & Vogel, 2012). 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.

¹¹ 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; Bidelman et al., 2021; Lefebvre & Jolicoeur, 2016; Simal & Jolicoeur, 2020).

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. Fitch, 2020; Jackendoff, 1987; Mahon & Kemmerer, 2020; Wurm & Caramazza, 2019). 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”.

3.1.2.1. The neural representation of event relations

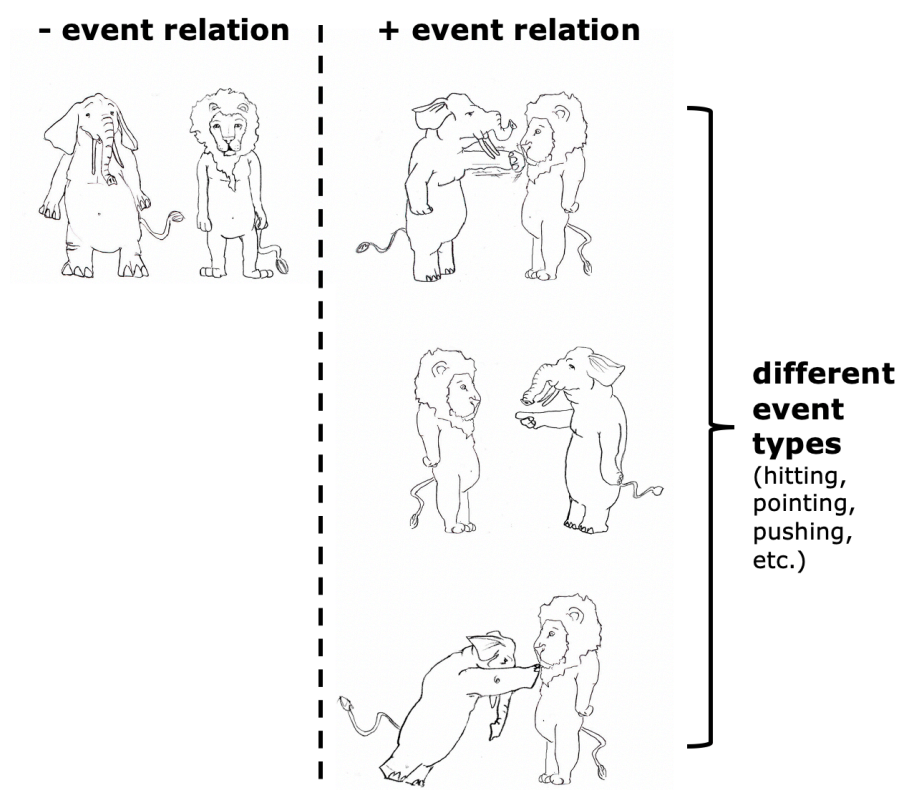


Figure 3.2 Examples of situations with and without “event relations”, as well as examples of different “event types”.

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 3.2). 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 3.2; 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 (de la Rosa et al., 2015; Dobel, Gumnior, et al., 2007; Hafri et al., 2013), 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. (2019), 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 Gaston (2020) and Matchin et al. (2019) 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 & Hummel, 2014; Shen et al., 2021). The results of Clevenger & 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 coordinations¹². In their study, pictures of events and coordinations were presented for 1500 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 1s) time-locked to picture onset in the posterior parietal cortex. Although this difference did not survive statistical significance in their analysis, this could be due to the small cohort of subjects (N=10).

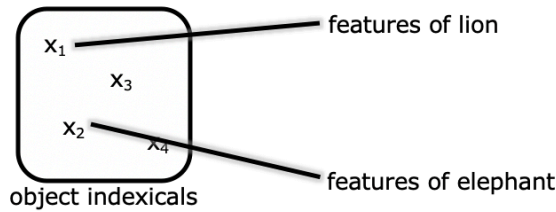
How should we expect event representation in working memory to manifest differently from simply representing a pair of objects? Through the lens of the object indexical system, we think that there are at least two major possibilities (see Figure 3.3). One possibility (Figure 3.3a) is that the event relation may be represented with a separate object indexical from the participants (cf. Clevenger & 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

¹² 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; Isik et al., 2017; Matchin, Liao, et al., 2019; Quadflieg et al., 2015), 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.

HITTING and other features of hitting¹³. In this case, there will be an NSW-like difference between the maintenance of events and coordinations. Another possibility (Figure 3.3b) 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 & 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., Eichenbaum & Cohen, 2014; Konkel & Cohen, 2009). 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 3.1.2.2 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.

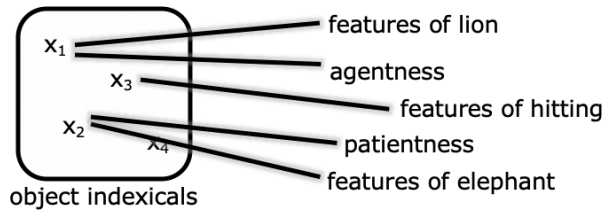
¹³Note that, just as the posterior parietal indexicals for objects are not thought to represent the features of the objects themselves but to connect to those features represented elsewhere, according to this hypothesis the indexicals for event relations connect to relevant event features and knowledge elsewhere.

the lion and the elephant: 2 object indexicals



the lion hit the elephant

Possibility (a): 3 object indexicals



Possibility (b): 2 object indexicals

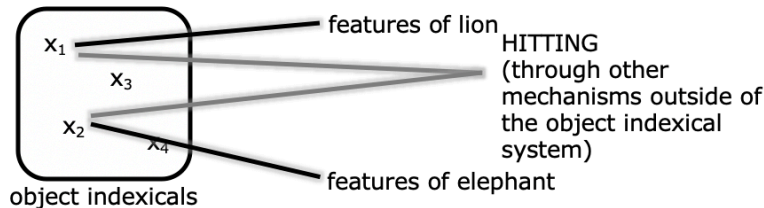


Figure 3.3 An illustration of the two possibilities of how event relations are represented.

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.

3.1.2.2. 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 Rissman & Majid, 2019; Strickland, 2017; Ünal et al., 2019; V. A. D. 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, Gumnior, et al., 2007; Hafri et al., 2013). In an EEG study by Cohn et al. (2017), subjects 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 & Greene, 2015, 2020; Lalisie & Smolensky, 2021; J. Wang et al., 2016). 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 & 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 & Chang, 2022). Peng et al. (2017), 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 (Duncan et al., 2023; Y. Fan & Luo, 2023; Fornaciai & Park, 2020; Huang et al., 2021; ten Oever et al., 2020; van der Werf et al., 2021; Wolff et al., 2015, 2017, 2019, 2020). 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 3.4. 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.

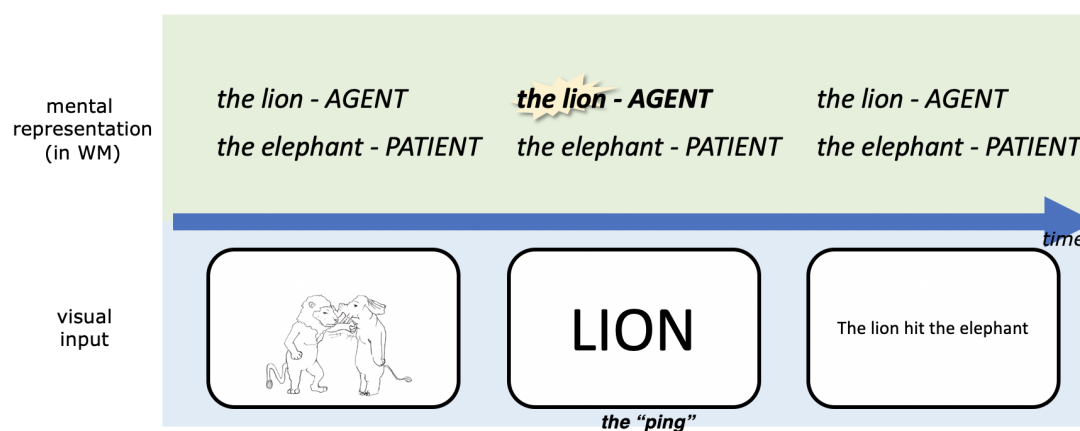


Figure 3.4 A simplified schematic illustration of our experimental design, highlighting our adapted pinging paradigm.

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.

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”¹⁴. 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; Dobel, Diesendruck, et al., 2007; Maass & Russo, 2003; but see L. J. P. Altmann et al., 2006; Barrett et al., 2002; Dobel et al., 2014; Geminiani et al., 1995; Kazandjian et al., 2011 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 (Buetti & Walsh, 2009; Sheahan et al., 2020; Summerfield et al., 2020; Walsh, 2003; 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, Luyckx et al., 2019; O’Connell et al., 2012; Spitzer et al., 2017; Yeung & Sanfey, 2004; for an overview see O’Connell & 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 (Luyckx et al., 2019; Spitzer et al., 2017), 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.

¹⁴ Other possibilities will be discussed in Section 3.3.2.

3.1.2.3. 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 3.1; N=19) and one in English in the US (Experiment 3.2; 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 3.1) and Cz, CPz, Pz (Experiment 3.2) as the CPP ROI for event role maintenance (see Figure 3.5).

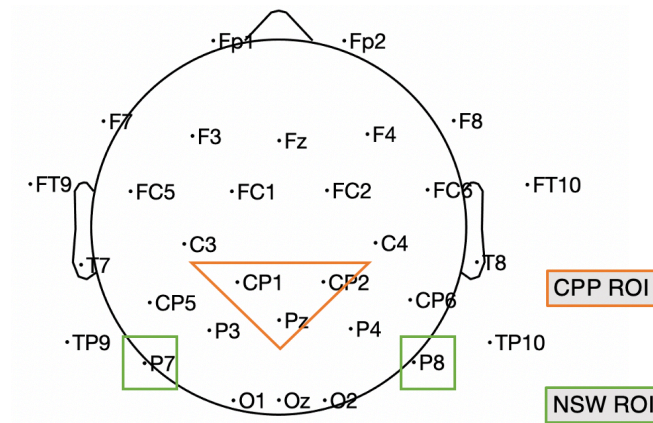


Figure 3.5 Illustration of the electrodes of our two ROIs in Experiment 3.1; in Experiment 3.2 the NSW ROI included Cz, CPz and Pz because of a different layout.

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.

3.2. Results

3.2.1. Experiment 3.1

3.2.1.1. 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 1241 ± 288 ms when the agent was pinged, 1230 ± 275 ms when the patient was pinged, 1236 ± 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 & Cusack, 2011; H.-B. 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 3.3).

3.2.1.2. EEG results

ERP responses to event pictures and coordination pictures are illustrated in Figure 3.6. 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 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 & 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 = 9999) 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 2469 ms after picture onset ($p=1\times 10^{-4}$, Figure 3.6). 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.

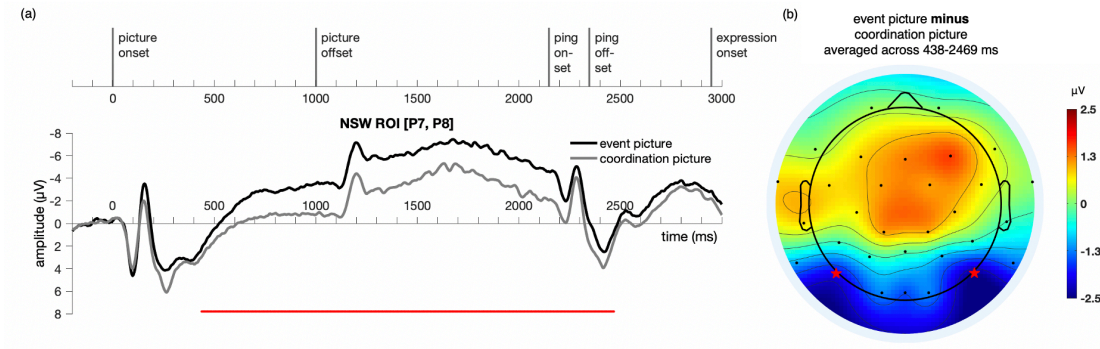


Figure 3.6 Illustration of the results for epochs time-locked to picture onset in the NSW ROI (electrodes P7 and P8, as marked by the stars in Figure b) for Experiment 3.1.

(a) The significant cluster (438-2469 ms) is marked in red. (b) The scalp distribution average across the 438-2469 ms time window. NSW: negative slow wave; ROI: region of interest.

ERP responses to the delay-period agent ping and the patient ping in the event condition are illustrated in Figure 3.6. Cluster-based permutation test (Maris & 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 agent ping and patient ping were identified in the CPP ROI, nor in the NSW ROI (Figure 3.7).

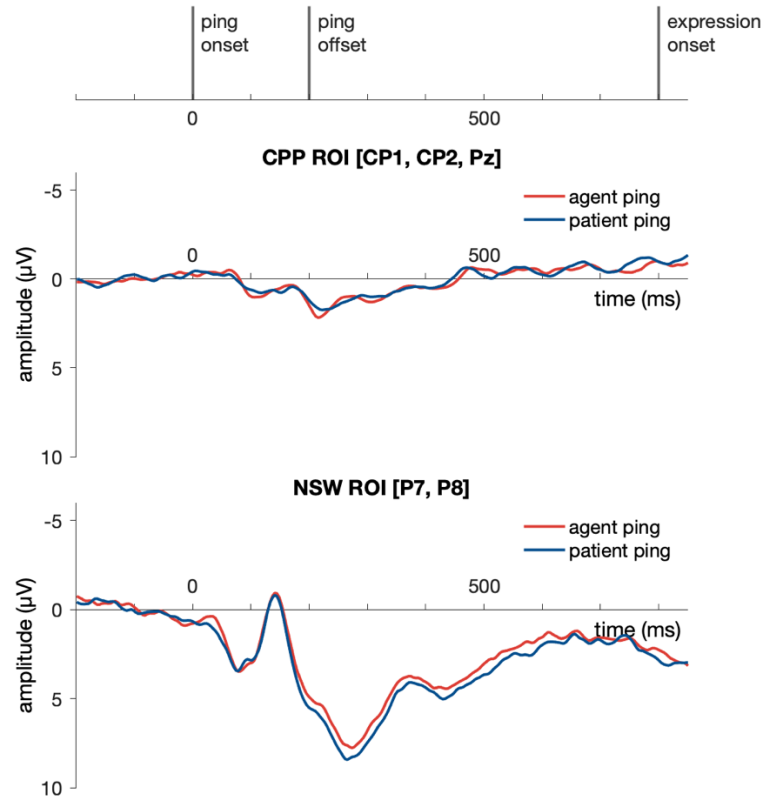


Figure 3.7 Illustration of the results for epochs time-locked to ping onset in the NSW ROI and the CPP ROI for Experiment 1.

NSW: negative slow wave; CPP: centro-parietal positivity; ROI: region of interest.

3.2.1.3. Rationale for Experiment 3.2

In Experiment 3.1, 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 EEG data analysis for more detail). In Experiment 3.2, we aimed to replicate the effect in a different population (Experiment 3.1 was run in Shanghai with stimuli in Chinese and Experiment 3.2 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 3.1 effects suggested that the presentation of the ping may have acted to reduce or eliminate the sustained negativity, we also took Experiment 3.2 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 3.1 (1000 ms), as an additional measure to ensure that the sustained effects we observed did indeed reflect delay-period activity rather than continued visual processing.

3.2.2. Experiment 3.2

3.2.2.1. 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 1617 ± 345 ms when the agent was pinged, 1616 ± 357 ms when the patient was pinged, 1617 ± 345 ms for all event trials and 1149 ± 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 3.1.

3.2.2.2. EEG results

Similar to Experiment 3.1, a cluster-based permutation test (Maris & 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 1109 ms after picture onset ($p=0.006$, Figure 3.8). As in Experiment 3.1, 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 3.2 than Experiment 3.1, perhaps because of the shorter picture duration for encoding in Experiment 3.2 (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 3.1. 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.

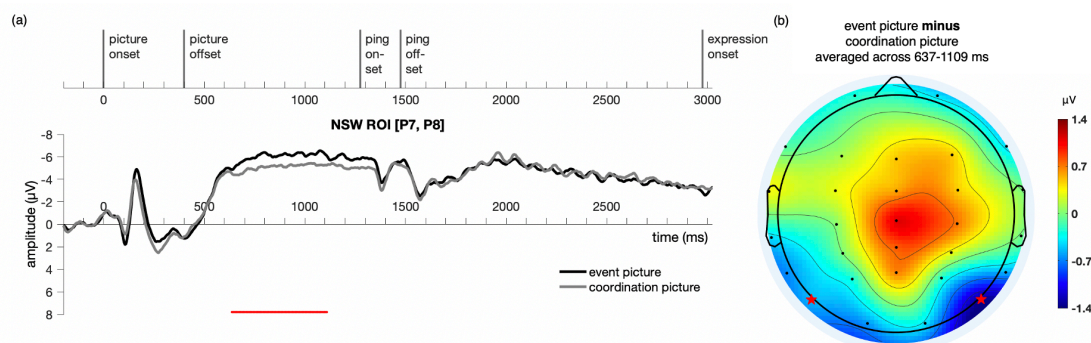


Figure 3.8 Illustration of the results for epochs time-locked to picture onset in the NSW ROI (electrodes P7 and P8, as marked by the stars in Figure b) for Experiment 3.2.

(a) The significant cluster (637-1109 ms) is marked in red. (b) The scalp distribution average across the 637-1109 ms time window. NSW: negative slow wave; ROI: region of interest.

ERP responses to the delay-period agent ping and the patient ping in the event condition are illustrated in Figure 3.9. Similar to Experiment 3.1, a cluster-based permutation test (Maris & 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.

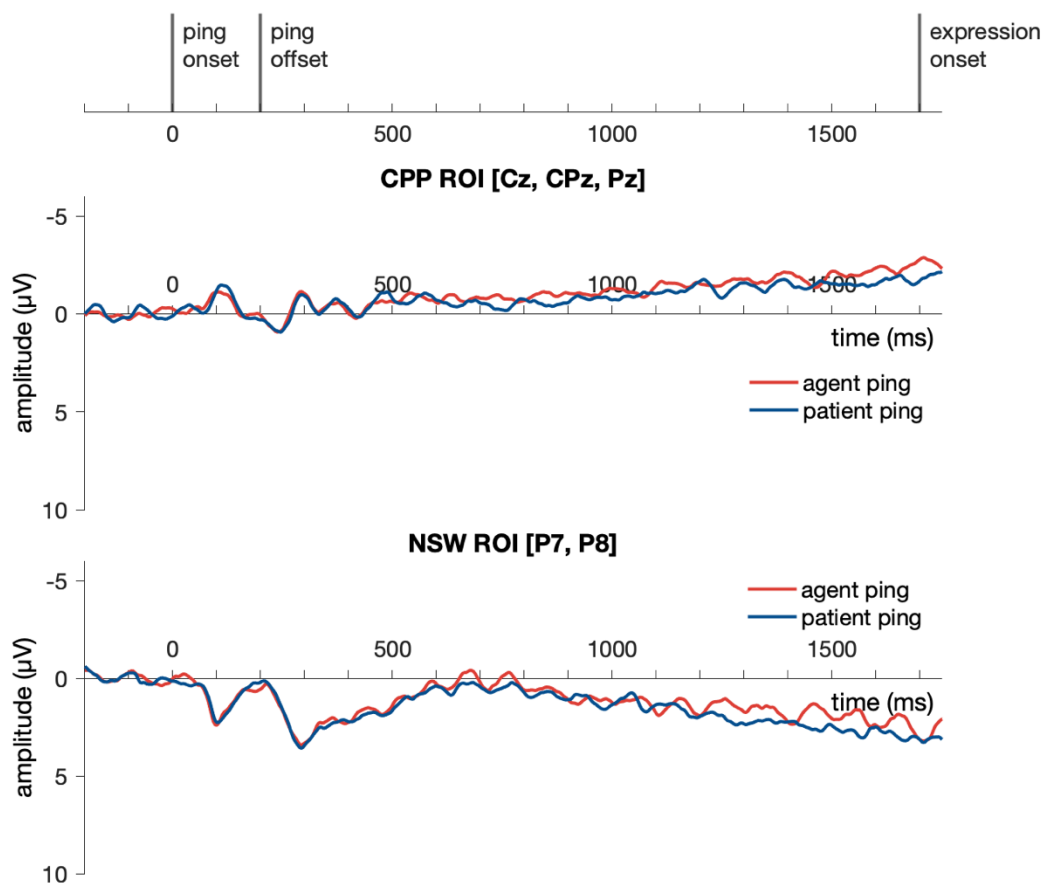


Figure 3.9 Illustration of the results for epochs time-locked to ping onset (-200:1800 ms) in the NSW ROI and the CPP ROI for Experiment 3.2.

NSW: negative slow wave; CPP: centro-parietal positivity; ROI: region of interest.

3.3. 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 3.1 and 3.2, 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 3.1 to Experiment 3.2.

3.3.1. 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 3.1 onset at ~ 450 ms and extended even until ~ 2500 ms; for Experiment 3.2, it onset at ~ 600

ms and extended until ~ 1100 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 towards 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 towards 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 (Becke et al., 2015; Bettencourt & Xu, 2016; Z. Fan et al., 2012; K. Jia et al., 2021; Robitaille et al., 2010) and event representation (Centelles et al., 2011; Matchin, Brodbeck, et al., 2019; Matchin, Liao, et al., 2019; Thompson et al., 2007; J. Wang et al., 2016; Adina Williams et al., 2017). Moreover, the posterior parietal cortex has been observed to be involved in both linguistic and visual processing, suggesting that it is hosting conceptual representations (Baldassano et al., 2017; Ivanova et al., 2021; Jouen et al., 2015; C. Liu et al., 2020; Menenti et al., 2012; Zadbood et al., 2017). 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. If this were the case, it would have important and interesting consequences for our understanding of what pointers represent. If we think of the pointers for objects as deeply intertwined with intuitive physics (Balaban et al., 2024), it would seem unlikely a priori that these pointers should stand in for events, or for other relations between objects or events as a whole, as most people do not have the intuition that events or relations move in space. Therefore, 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 judgement 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 (X. Ding et al., 2017; X. Lu et al., 2022; Paparella & Papeo, 2022; Stahl & Feigenson, 2014). 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 & Luria, 2016; McCants et al., 2020; Rabbitt et al., 2017). 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).

3.3.2. 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 & 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 J. 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 (Hummel & Holyoak, 2003; A. E. Martin & Doumas, 2017; Shastri, 1999, 2002, 2007; Shastri & Ajjanagadde, 1993), yet have only recently been tested by Peng et

al. (2017). 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 (Balestrieri et al., 2021; Chota et al., 2022; J. Jia et al., 2017; Peters & Kriegeskorte, 2021; Pomper & Ansorge, 2021). 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 Alexander 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 Alexander 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, 2006b, 2006a; Reisinger et al., 2015; Rissman & Lupyan, 2021, 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.

3.3.3. 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; X. J. Wang, 2021), at least two lines of research have argued that working memory retention can also sometimes be implemented by “activity-silent mechanisms” (Beukers et al., 2020; Kamiński & Rutishauser, 2020; Lewis-Peacock et al., 2012; Rose et al., 2016; Stokes, 2015) or other mechanisms that do not induce sustained changes in neural metabolism (e.g., Lundqvist et al., 2021; Wutz et al., 2020).

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 & Eimer, 2016; Feldmann-Wüstefeld et al., 2018; J. 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) (Cruz Heredia et al., 2021; Kaan et al., 2000; Lau & Liao, 2018; Lau & Yang, 2023; Lo & Brennan, 2021; McKinnon & Osterhout, 1996; Sprouse et al., 2021; Yano & Koizumi, 2018, 2021; 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 3.1 and 3.2 demonstrated that the working

memory maintenance of event relations *can* be realized by sustained activity. However, as Experiment 3.2 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 & Eimer, 2016; Feldmann-Wüstefeld et al., 2018; J. Zhang et al., 2021), yet resulted in little detriment to behavioral accuracy (Kreither et al., 2022)¹⁵. 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.

3.3.4. 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, Hafri et al., 2017; Walbrin & Koldewyn, 2019; Wurm et al., 2016; Wurm & Caramazza, 2019a; Wurm & Lingnau, 2015) 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

¹⁵ Notably, the reaction times seemed to be longer for Experiment 3.2 compared to Experiment 3.1, despite similar stimulus-onset asynchrony across the picture and the linguistic probe. This is in line with the observation in J. 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.

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 (X. Lu et al., 2016; H. 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 & Caramazza, 2019; Yang et al., 2020).

3.3.5. Conclusions

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.

3.4. Materials and methods

3.4.1. Experiment 3.1

3.4.1.1. 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.

3.4.1.2. Stimuli

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 Figure 3.12.

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 3.1. 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.

Table 3.1 Example of linguistic expressions presented in Experiment 3.1.

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"		

3.4.1.3. 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 1000 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 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 3.1) 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 1000-2000 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 3.1 see Figure 3.10.

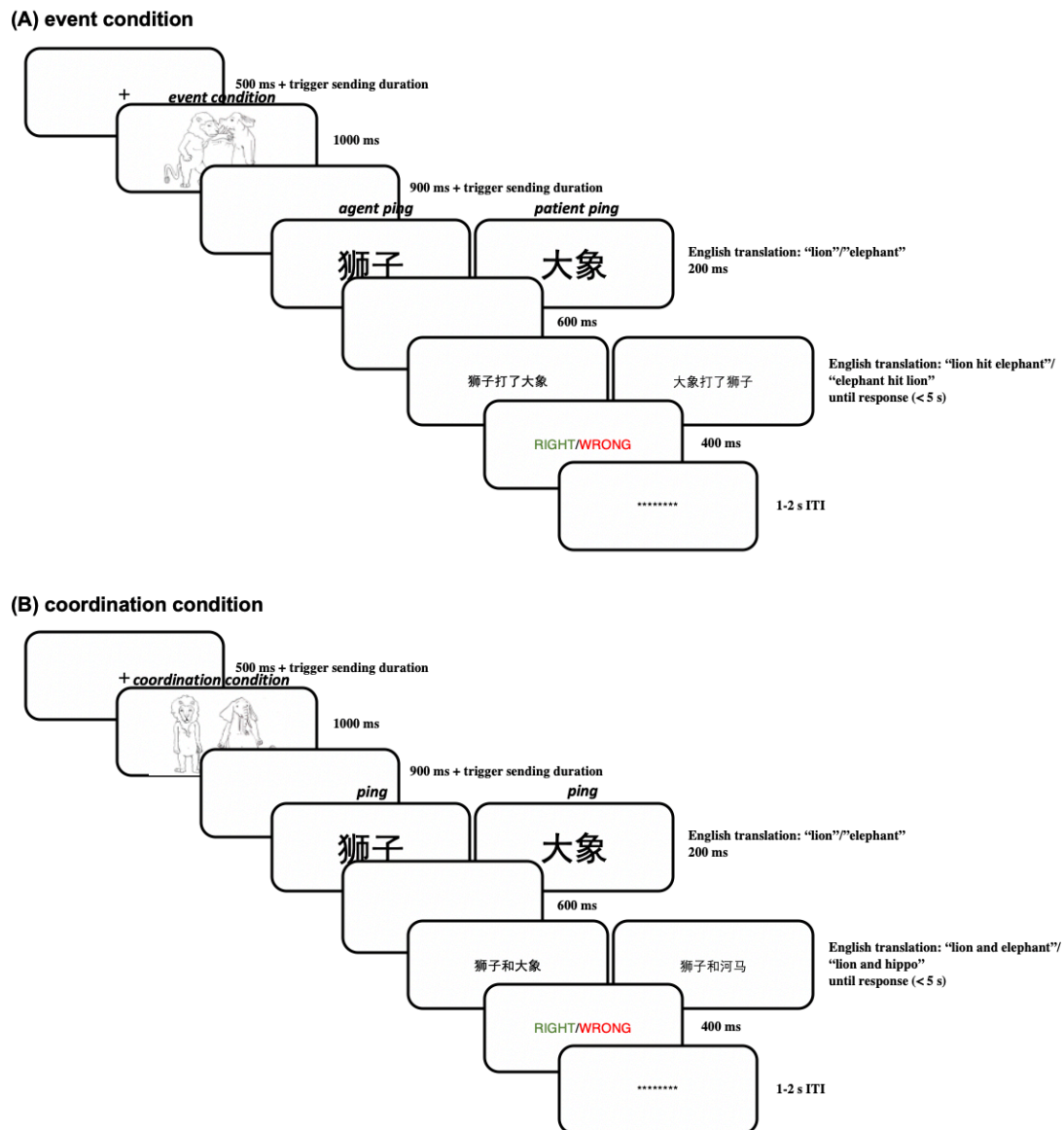


Figure 3.10 (A) Illustration of one trial in the event condition in Experiment 3.1; (B) illustration of one trial in the coordination condition in Experiment 3.1.

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 grey, and the texts were white. The pictures were created in 1600×1200 px, and were presented on the screen as $24^\circ \times 18^\circ$. 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.

3.4.1.4. 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).

3.4.1.5. EEG recording

EEG signal was recorded using a 32-channel active electrode system (Brain Vision actiCHamp; Brain Products) with a 1000 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 k Ω . 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 PyCorder software (<http://www.brainvision.com/pycorder.html>) and filtered online between DC and 200 Hz with a notch filter at 50 Hz.

3.4.1.6. EEG data analysis

EEG data were analyzed using customized MATLAB (version: R2020b) scripts based on EEGLAB v2021.0 (Delorme & Makeig, 2004) and ERPLAB v8.20 (Lopez-Calderon & 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 ICLabel 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 labelled (using the ICLabel 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 to 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 μ 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 μ 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 & 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 = 9999) 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 μ 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 & 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 3.2 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.

3.4.2. Experiment 3.2

3.4.2.1. 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 3.1, detailed in Section 3.4.1.5. 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.

3.4.2.2. Stimuli

The picture stimuli in Experiment 3.2 were the same as Experiment 3.1. The probe linguistic expressions in Experiment 3.2 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 3.2 were English words in capital letters, e.g., “LION”.

3.4.2.3. Procedure

The procedure of Experiment 3.2 is generally similar to Experiment 3.1, 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 1500 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 1000-2000 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 3.2 see Figure 3.10.

The background of the screen was grey, 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.

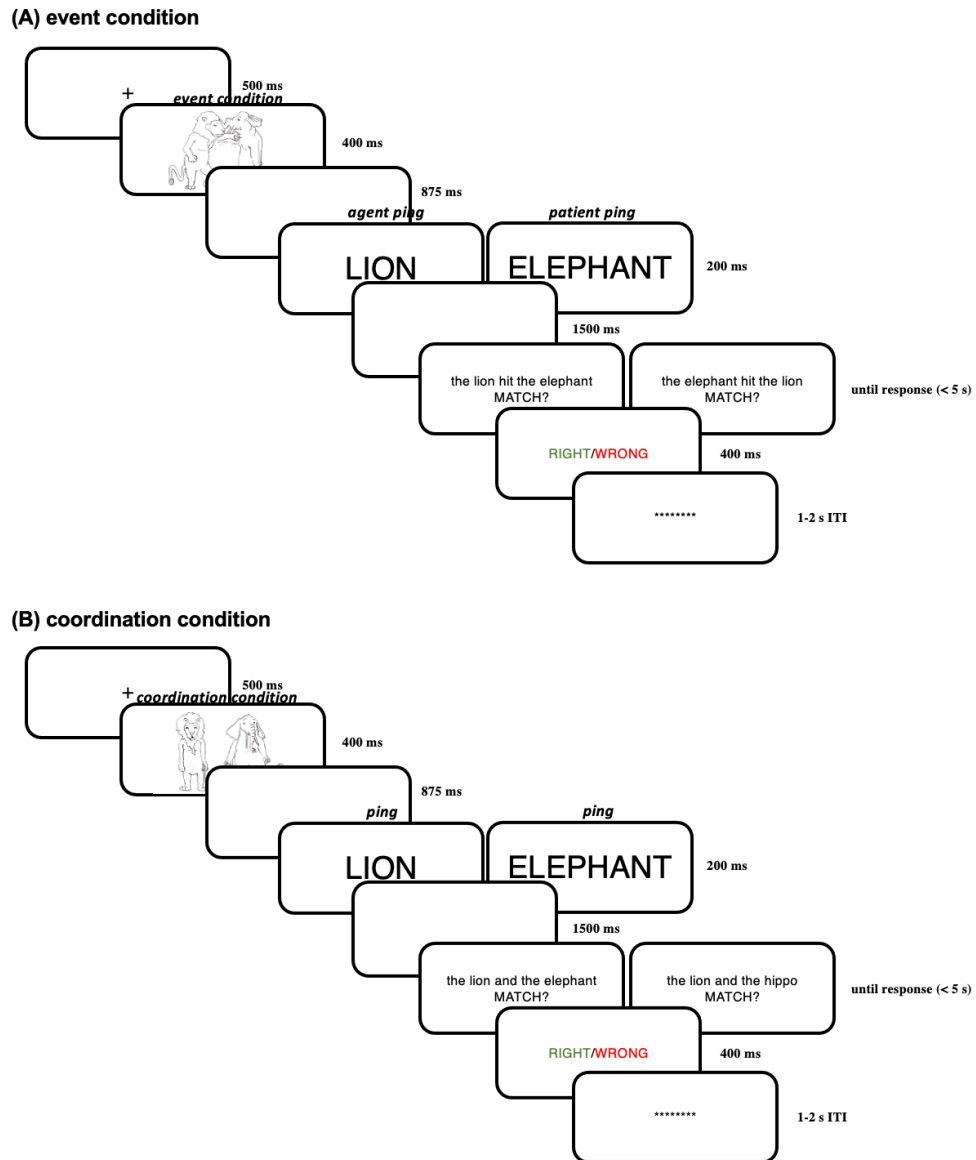


Figure 3.11 (A) Illustration of one trial in the event condition in Experiment 3.2; (B) illustration of one trial in the coordination condition in Experiment 3.2.

3.4.2.4. Behavioral data analysis

Same as Experiment 3.1.

3.4.2.5. 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 k Ω for all scalp electrode sites, and less than 10 k Ω 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.

3.4.2.6. EEG data analysis

The analysis pipeline was the same as Experiment 3.1 (Section 3.4.1.6), except that for Experiment 3.2, 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.

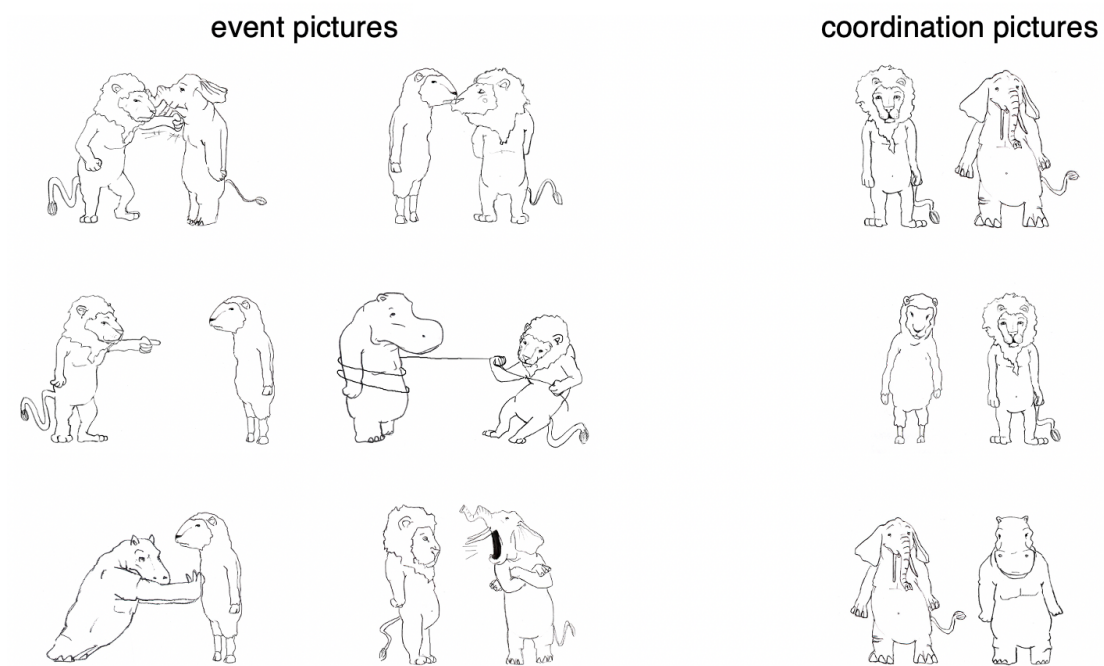


Figure 3.12 Examples of the event pictures and coordination pictures used in the current experiment.

4. Chapter 4

In this chapter, we report EEG evidence suggesting that linguistically-derived mental objects recruit pointers in working memory similar to visually-derived ones.

4.1. Introduction

In our moment-to-moment cognition, we often need to keep one or more “objects” in the mind. These mental objects can be derived from a variety of sources, from direct visual input, to speech, to spontaneous thought (Brodbeck et al., 2016; Fukuda & Woodman, 2017). For example, the representation of a particular cat could be set up in the mind based on actually seeing a cat, or being told about a cat via spoken or sign languages. The most commonly studied scenario is when these objects are derived from direct visual input in recent history (i.e., within the same trial or experimental session). For example, in a trial of a typical visual working memory study (e.g. Luria et al., 2016), participants first see a memorandum consisting of one or more objects. After a short delay, they need to respond to a probe. To respond to whether the probe matches with the memorandum, one needs to hold the memorandum objects in mind across the delay. In this case, mental objects are derived from a direct visual source in recent history (i.e., the presented memorandum).

What about non-visually derived mental objects? 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, 2025; 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). According to the latter hypothesis, holding multiple cat individuals in the mind after viewing them or based on linguistic instructions (e.g., being told that “I saw a brown cat and a yellow cat yesterday in our yard”) would tax shared object-based resources in the mind, which would not be predicted by the former hypothesis. Of note, 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 et al., 2019; Ngiam, 2024; Swan & Wyble, 2014; Thyer et al., 2022; Wei et al., 2024; Yu & Lau, 2025). Critically, in this view, pointers are represented separately from features, corresponding to object-based load and feature-based load respectively. 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; Y. Xu, 2009).

Accompanying this theoretical trend, recent years have seen a surge in interest in exploring neural markers tracking the number of pointers in moment-to-moment processing (Jones, Diaz, et al., 2024; Jones, Thyer, et al., 2024; Thyer et al., 2022; Yu, in press). For example, an evoked response—the negative slow wave (NSW)—appears to track visual working memory load (Feldmann-Wüstefeld, 2021; Fukuda et al., 2015; Yu, Li, et al., 2024). However, a trial in the current mental imagery study necessarily takes multiple seconds, and evoked

sustained responses over longer EEG epochs are easily derailed by much larger yet non-neural drifts (Kappenman & Luck, 2010; Luck, 2014). In the current study, we instead examine posterior alpha power (8-12 Hz). Analysis of the alpha band is unlikely to be affected by slow drifts at multi-second level, and alpha power is also known to track visual working memory load (Diaz et al., 2021; Fukuda et al., 2015; Fukuda & Woodman, 2017). Encouragingly, posterior alpha power tracks object-based load when the number of features and the spatial extent of attention are controlled for (Chen et al., 2024). This allows us to use posterior alpha power as a measure of the number of pointers a participant is currently holding in mind.

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-derived 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 (Figure 4.1A). 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.

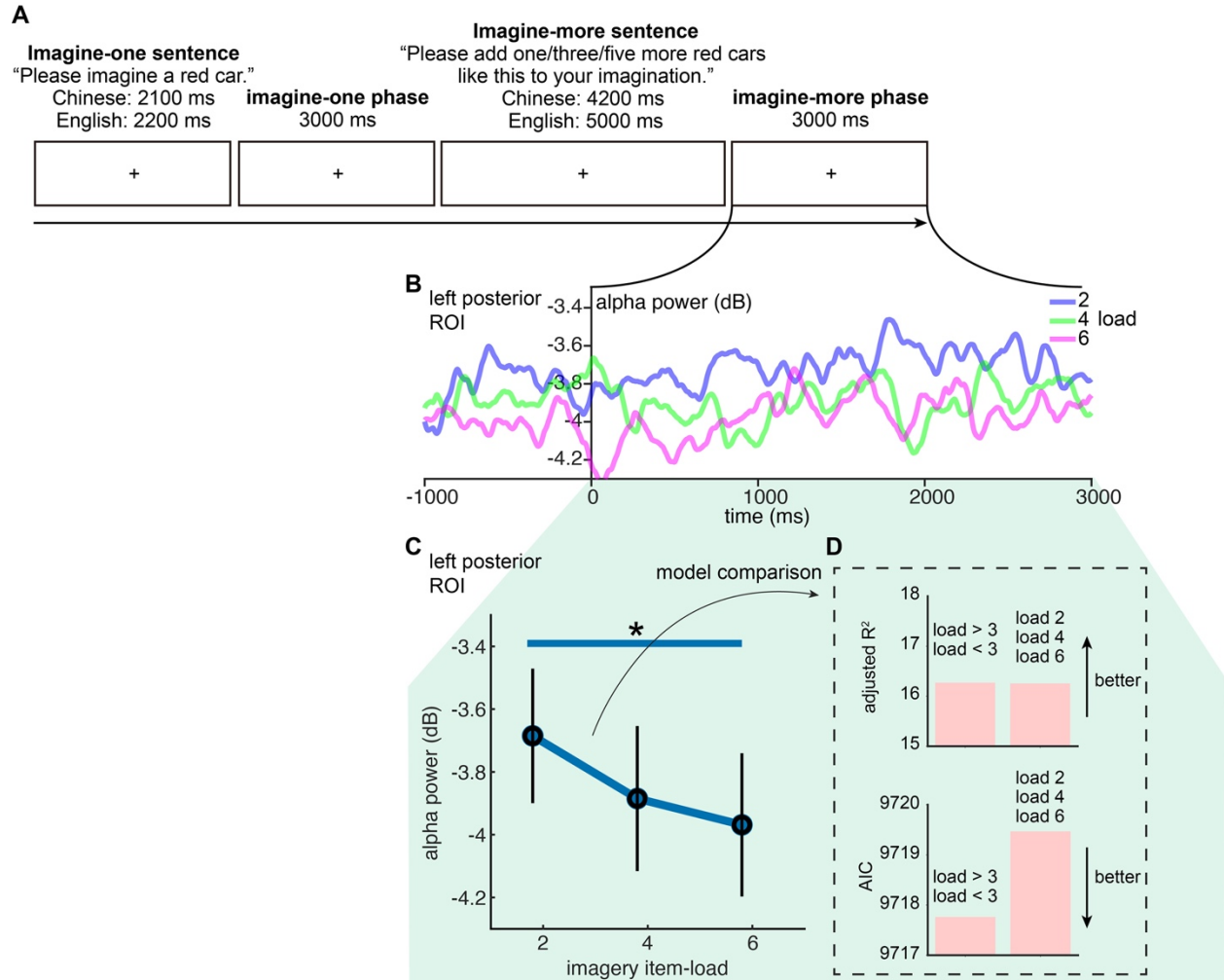


Figure 4.1 Left posterior alpha power inversely predicts speech-induced mental imagery load.

(A) Illustration of the procedure of a trial in our Experiment 4.1. (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 Figure 4.2. (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.

4.2. Experiment 4.1: Materials and Methods

4.2.1. Participants

The entire dataset consists of 52 participants (age 18-29, 31 female), half tested in Shanghai, and half tested in Maryland. 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 et al., 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.

4.2.2. 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: 请想象一 [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).

4.2.3. 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 et al., 2015; Fukuda & Woodman, 2017; Y. 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.

4.2.4. 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 k Ω . 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.

4.2.5. 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. 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 1,763 remaining (81% of all trials; two: 578; four: 597; six: 588) epochs in the Chinese dataset and 1,847 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). 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.

4.2.6. 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: $Posterior\ alpha\ power \sim 1 + load + ROI + load \times ROI + (1|participant:site:trial)$. The response variable was modeled with a normal distribution and linked with an identity function.

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: $Posterior\ alpha\ power \sim 1 + load + (1|participant:site)$. 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 ~ 3 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, Thakurdesai, et al., 2024). All reported p -values are two-tailed.

4.3. 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, $\beta = -0.007$, $SE = 0.015$, $t = -0.46$, $p = 0.64$ (Figure 4.1B, C). This lateralization effect was consistent with the visual impression from scalp distributions (Figure 4.2). Complementary GLME analysis on the left ROI revealed a significant difference between load 2 and 6 ($\beta = -0.075$, $SE = 0.037$, $t = -2.01$, $p = 0.044$), and a numerical trend 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 ($\beta = 0.043$, $SE = 0.022$, $t = 2.01$, $p = 0.044$), but not between load 4 and 6 ($\beta = -0.012$, $SE = 0.021$, $t = -0.54$, $p = 0.58$).

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 (Figure 4.1D). Taken together, modeling load conditions as three levels appeared not to be necessary, compared to the two-load-level model.

4.4. Discussion

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

With a novel imagine-one-then-more paradigm leveraging 3,610 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 et al., 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; J. 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). While our current study was not designed to adjudicate this long-standing debate, we observed that imagery load for objects induced a left-lateralized EEG response. Given the limited spatial resolution of EEG, future studies employing neuroimaging methods with higher spatial resolution (e.g., MEG) will be critical in clarifying the laterality of neural sources for this effect.

4.4.2. 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 and Luck, 2008; Bengson and 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 (2024) 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 (2024) 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 (2024) call for future research to further characterize the capacity limit(s) for mental imagery representations per se and downstream computations.

4.4.3. Is “mental object” one or multiple natural kinds?

We demonstrated that mental objects created by linguistic instructions also drive posterior alpha suppression, similar to mental objects created by direct visual experience in recent past. This suggests that these two types of mental objects may be supported by the same system. Do all “objects” belong to the same ‘natural kind’ in this sense? This is not necessarily the case: for example, although phonological units (e.g., phonemes, syllables) have been seen as “auditory objects” (Griffiths & Warren, 2004), increasing the load of phonological units is accompanied by an increase of posterior alpha power (Proskovec et al., 2019) as opposed to a decrease. This may suggest that phonological units belong to another category of mental objects. Therefore, more studies with a variety of tasks and modalities need to be conducted in order to carve the joints of “mental objects”, i.e., what kind of mental objects belong to the same natural kind, and where to draw the lines among different natural kinds (Yu, 2025). This proposed line of research can thus also serve as an empirical reply to 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 tasks, modalities and neural systems.

4.5. 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 support the hypothesis that mental objects tax a shared set of pointers in the mind, irrespective of their source modality.

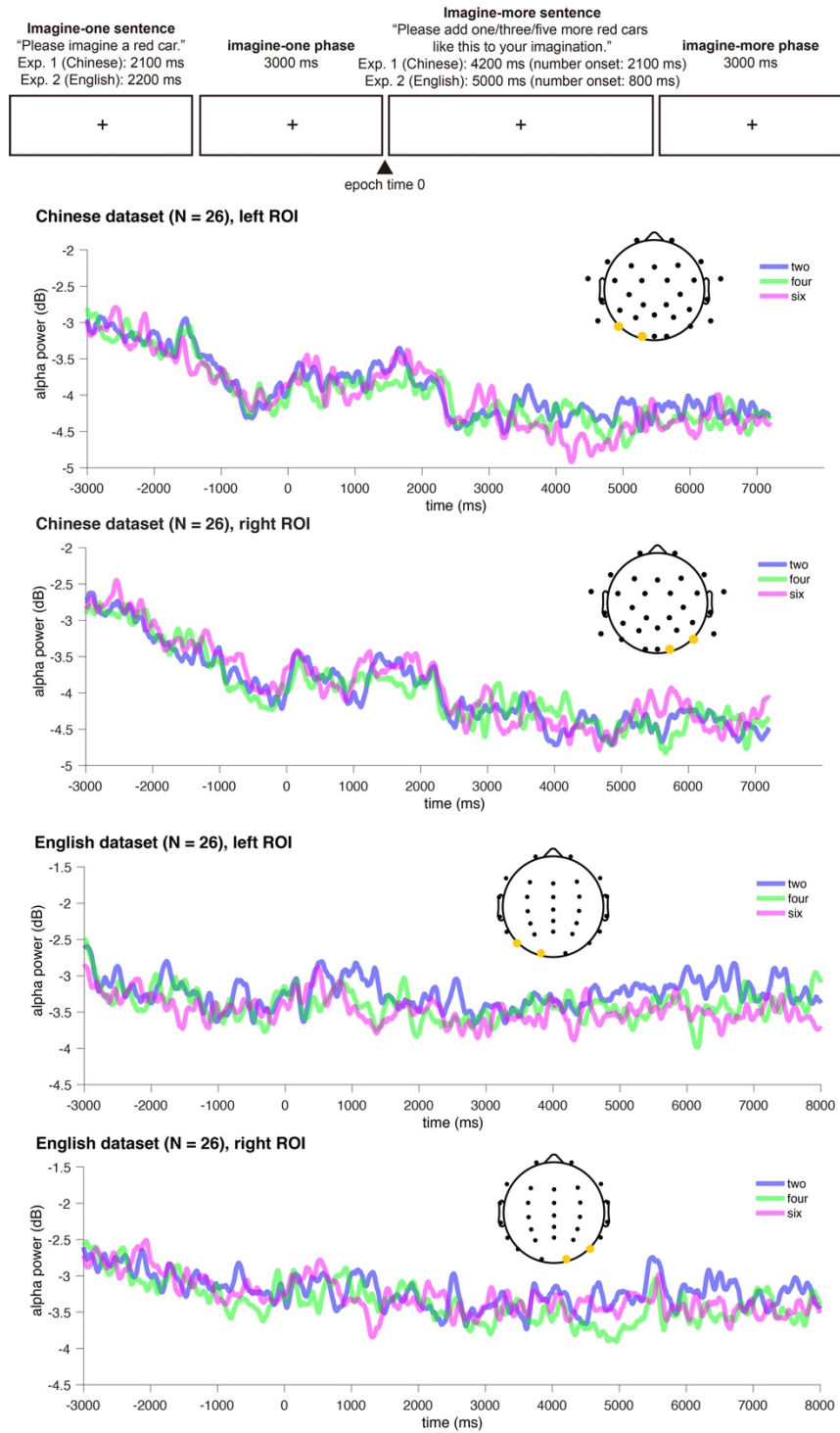


Figure 4.2 Posterior alpha power in the left and right ROIs, illustrated across the full time period of the trial from the imagine-one phase to the end of the imagining-more phase.

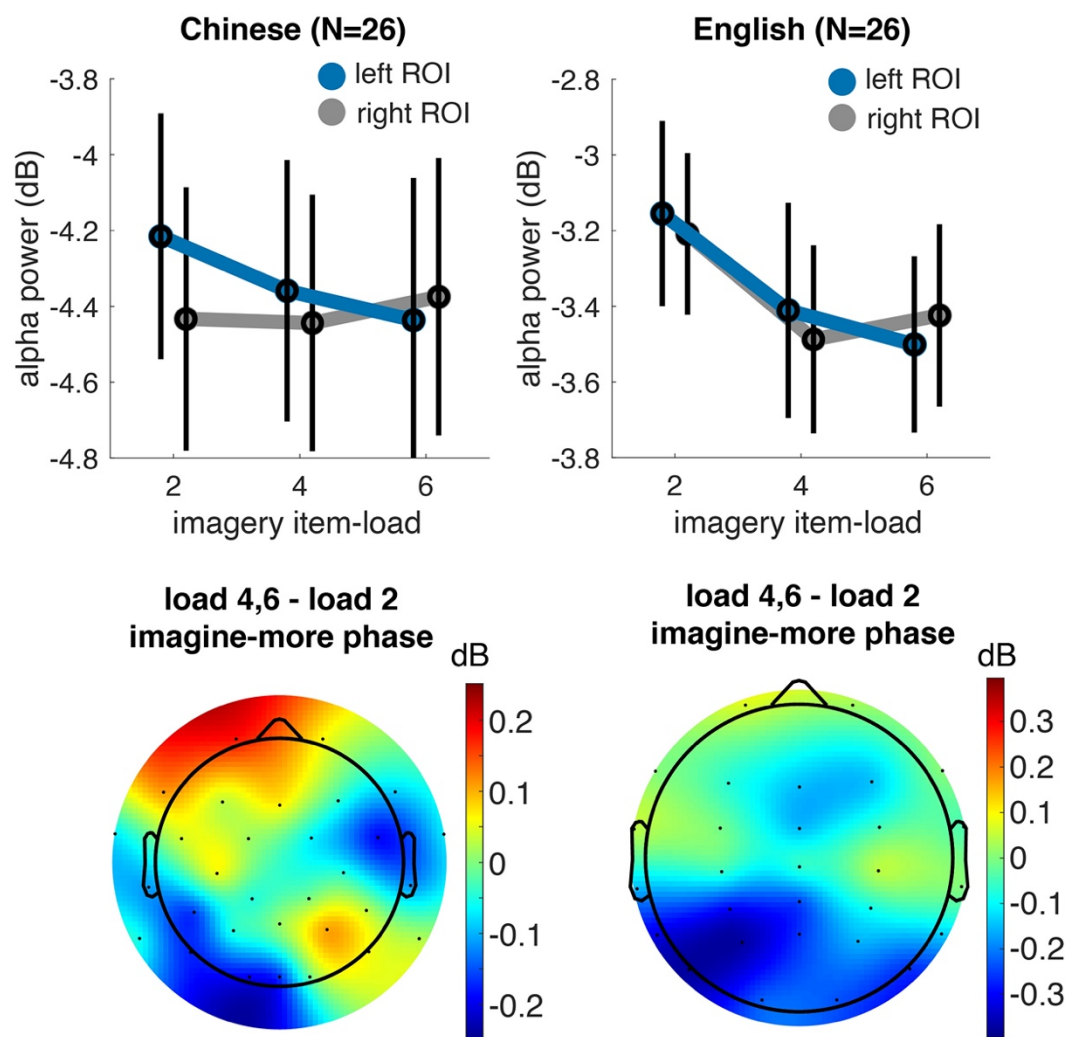


Figure 4.3 Posterior alpha power during the imagine-more phase and scalp distribution for the English and Chinese participants respectively. Error bars represent standard error across participants.

4.6. Future directions

What would be stronger evidence that discourse referents (derived from language) and objects derived from vision share the same set of pointers? In our current case, objects are derived from language, but in language comprehension, we do not seem to always phenomenologically construct visual imagery in the mind, yet we appear to be able to track discourse referents all the time. In other words, visual imagery appears not to be compulsory for representing discourse

referents. In order to test whether the current results extend to cases where little to no phenomenological visual imagery is involved, future studies could utilize sentences that are harder to visualize (e.g., “John likes singing and Mary likes dancing”, Figure 1.1), without explicit instructions to imagine visually, and with individuals that have little to no visual imagery (e.g., aphantasia).

5. Chapter 5¹⁶

In this chapter, we report evidence for a novel neural MEG sensor-level hallmark for pointers in visual working memory tasks: the right posterior activity (RPA).

5.1. Introduction

Visual working memory (VWM, or visual short-term memory, VSTM) for objects is a crucial yet extremely capacity-limited faculty for humans (Bays et al., 2024; Cowan, 2001; Fukuda et al., 2010; Luck & Vogel, 1997; Ma et al., 2014; Y. Xu, 2017; W. Zhang & Luck, 2008). Recent work has attempted to reconcile earlier debates between “object-only” or “feature-only” theories of the capacity limit with theories of VWM architecture that contain elements of both and are thus better able to account for the ever-growing body of experimental findings at the same time (for a review see (Ngiam, 2024)). One such proposal is that VWM is implemented by a capacity-limited, content-free set of pointers (Swan & Wyble, 2014; Thyer et al., 2022; Y. Xu & Chun, 2006; Yu & Lau, 2023): the finite set of pointers (or indexes/indexicals/tokens) act as stand-ins for each individual object in the mind (Balaban et al., 2023, 2024; Pylyshyn, 1989), and bind features together for the corresponding objects. For example, to represent an individual red square, a pointer would be set up in the mind that binds the RED feature and the SQUARE feature together. This is different from traditional conjunctive coding, where a given neuron might only respond to the combination of RED and SQUARE, because in a conjunctive code a

¹⁶ This chapter contains content published as Yu, X., & Lau, E. (2025). A finite set of content-free pointers in visual working memory: magnetoencephalography (MEG) evidence. *NeuroReport*, 36(3), 153-160.

unit always responds to this particular combination, and in that sense both red and square are part of its “content”. In contrast, a content-free pointer has no permanent content, as it can bind different features from moment to moment (B. Park et al., 2020; K. Zhou et al., 2010). In this sense, neither the RED feature nor the SQUARE feature is represented within the pointer itself. This model, therefore, has allowed for separate capacity limits for pointers and visual features (Bowman & Wyble, 2007; Green & Quilty-Dunn, 2021; Hedayati et al., 2022; Swan & Wyble, 2014; B. Wang et al., 2017). Per theory, the representation for pointers in the brain, if any, needs to meet the following three criteria: (1) **load-dependent**: that this representation co-varies with (within-capacity) VWM load; (2) **capacity-limited (i.e., finiteness)**: that this representation reaches some kind of plateau or homeostasis beyond VWM capacity (of ~ 3 objects); (3) **content-free**: the pointers contain only addresses, not the features themselves, and their function is to bind the addressed features together. The current study looks for such neural responses in MEG (magnetoencephalography). Recent EEG (electroencephalography) evidence supports the existence of content-free representations in VWM with machine-learning multivariate decoding (Thyer et al., 2022) or univariate event-related potentials (Balaban et al., 2019, 2023; K. E. Wilson et al., 2012); however, it’s yet unclear whether these neural markers reflect the capacity limit. Furthermore, earlier fMRI (functional magnetic resonance imaging) evidence in fact suggested univariate representations in the brain that meets all three criteria, in regions including the posterior parietal cortex (Naughtin et al., 2016; Y. Xu & Chun, 2006). The limitation of fMRI, however, is that it is hard to tell apart perceptual and VWM processes temporally; methods such as EEG and MEG have superior temporal resolution. Here we offer MEG evidence supporting the existence of VWM pointer representations that meet all three criteria: A right

posterior MEG component is load-dependent, capacity-limited (finite) and content-free at the same time.

Previous MEG studies have identified load-dependent, capacity-limited responses during VWM in left and right posterior channels (Mitchell & Cusack, 2011; Robitaille et al., 2009, 2010), but with a bilateral-presentation paradigm and a lateralized change-detection task¹⁷, with the hope of finding a CDA [contralateral delay activity]-like response as in EEG. For EEG, posterior responses contralateral to the attended side were found to be modulated by VWM load, while ipsilateral responses are minimally affected by VWM load (Arend & Zimmer, 2011). Moreover, the attended side of the memorandum accompanies a clear maxima in contralateral posterior electrodes for EEG during the delay between the memorandum and the probe (Mccollough et al., 2007). However, MEG studies have found that the amplitude of responses in both ipsilateral and contralateral posterior channels can be modulated by VWM load (Robitaille et al., 2009). Moreover, side-sensitive channels in MEG appeared to cluster around the midline, rather than being lateralized (Robitaille et al., 2010). Taken together, it is unclear whether posterior responses in MEG reflect the exact same neural sources as for CDA in EEG (Becke et al., 2015; Robitaille et al., 2009), therefore, a similar lateralized task and contralateral-minus-ipsilateral treatment of data for MEG may not be necessary. In the current study, we employed a similar logic as in the pioneering fMRI study of Xu and Chun (Y. Xu & Chun, 2006). In Experiment 5.1, we tested whether the left and right posterior effects in MEG are still load-

¹⁷ In EEG studies examining the CDA response, separate stimuli are presented on the left and right sides of the screen respectively. The participants are cued beforehand to memorize the memorandum on one of the two sides (this being a lateralized as opposed to a whole-field change-detection task (Adam et al., 2020)). The CDA is computed by subtracting ipsilateral from contralateral ERP (event-related potential) responses during the delay between the memorandum and the probe, and its amplitude is load-dependent and capacity-limited (Luria et al., 2016; Vogel & Machizawa, 2004).

dependent and capacity-limited with a central-presentation paradigm and a whole-field change-detection task (as in (Y. Xu & Chun, 2006)). To preview, we found a reliable load-dependent and capacity-limited effect in right posterior channels. In Experiment 5.2, we showed that this right posterior load-dependent, capacity-limited effect is also content-free. Therefore, this right posterior activity (RPA) meets all three criteria for a neural index of VWM pointers.

5.2. Experiments 5.1 and 5.2: Methods

5.2.1. Participants

For Experiment 5.1, 20 participants entered data analysis (12 female, age 19-41, $M=24$, 3 left-handed); 1 additional participant's data was excluded from analysis due to excessive noise (see MEG analysis). For Experiment 5.2, 20 participants entered data analysis (13 female, age 18-37, $M=21$, 2 left-handed); 1 additional participant's data was excluded from analysis due to excessive noise. All subjects reported having normal or corrected-to-normal vision. Informed consent was obtained from all participants and they received monetary reimbursement for their participation. Procedures were approved by the UMD IRB Office.

5.2.2. Procedure

Experiments 5.1 and 5.2 had the same procedure. Each subject was presented with two blocks of 200 trials each, which differed in the features of the objects: one block in which the objects were colored squares, and another block in which the objects were complex shapes. The order of the

blocks was balanced across subjects. For each block, there were 50 trials for each VWM load condition (1, 2, 4 and 6); within each 50 trials, half of the trials had a match probe and the other half had a change probe.

In each trial, the participants were first presented with a fixation cross with a duration of 500 ms (Experiment 5.1) or 800 ms (Experiment 5.2) against a grey background, followed by a sample screen with a duration of 160 ms (Experiment 5.1) or 200 ms (Experiment 5.2)¹⁸. Then, they were presented with the sample; see Figure 5.1. The sample consisted of 1, 2, 4, 6 different objects: for colored-square trials, they were sampled from 8 colors (black [0,0,0], dark blue [0,0,139], green [0,255,0], light blue [173,216,230], pink [255,192,203], red [255,255,0], white [255,255,255] and yellow [255,255,0]); for complex-shape trials, they were sampled from 9 complex shapes similar to those used in (Y. Xu & Chun, 2006). 8 fixed locations were evenly distributed 4.5° from the center of the screen, and these locations were marked by 8 light grey squares (2.5° × 2.5° each). The objects (colored squares or complex shapes) were presented at the center of the light grey squares, spanning 2.1° × 2.1°. The light grey squares served as a visual cue to encourage the individuation of each object and to discourage grouping. A blank screen delay of 1000 ms followed the sample, after which the probe appeared on the screen. At the probe, only one object was presented at one of the 8 locations. On “match” trials, this object was the same one as the one in the sample for that location; on “change” trials, this probe object was randomly selected from other objects used in this block. Participants pressed one button to indicate “match”, and pressed another button to indicate “change” (buttons balanced across participants), with no time limit for response. After their response, they were shown a feedback

¹⁸ Because of a programming issue, 5 participants in Experiment 5.1 had a fixation cross duration of 800 ms, and 5 participants in Experiment 5.2 had a sample duration of 160 ms.

screen for 600 ms, followed by an inter-trial interval (ITI) of 2500-3500 ms. During the ITI, a line of asterisks was presented on the screen; before the experiment, participants were instructed to blink only during this period, and to avoid blinking during the trials. While the fixation cross was presented on the screen, the participants were instructed to fixate at the center of the screen. They were also instructed against verbal rehearsal. At least 8 practice trials for each block were run before the main experiment for each participant, and a self-paced break was administered every 50 trials. Experiment 5.1 analyses included only the colored square block¹⁹, while Experiment 5.2 analyses included both a colored square block and a complex shape block (order balanced across participants). The experiment was run in MATLAB R2009a with Psychtoolbox 3 (Brainard, 1997; Kleiner et al., 2007).

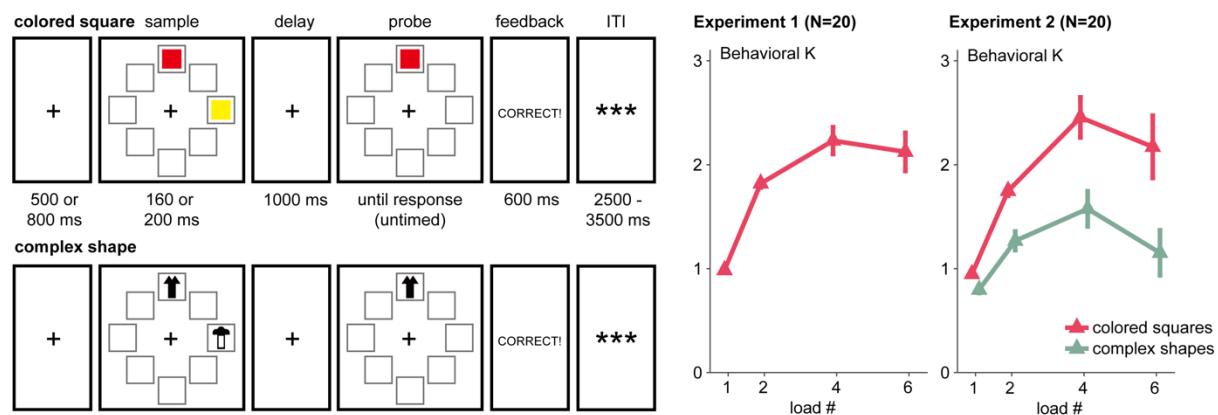


Figure 5.1 Experimental procedure and behavioral results (Cowan's K) across Experiments 5.1 and 5.2.

Error bars stand for standard error. ITI: inter-trial interval. For detailed presentation parameters of the stimuli see Procedure.

¹⁹ We originally included a complex shape block in Experiment 5.1, but due to a timing error we were unable to analyze the data for those trials.

5.2.3. Behavioral analysis

Based on the hit rates and false alarm rates for each condition, we calculated Cowan's K (Cowan, 2001; Rouder et al., 2011), which reflects the number of objects successfully held in VWM. The formula for Cowan's K is as follows: $K=N \times (H-F)$, where N is sample load (i.e., 1, 2, 4 or 6), H is the hit rate for detecting the change in the probe on change trials, and F is the false alarm rate (erroneously "hallucinating" a change in the probe on match trials).

5.2.4. MEG recording

Prior to recording, five head position indicator coils were affixed to each participant's head, and the position of these coils relative to the nasion and tragus, as well as the participant's head-shape, was digitized using a Polhemus 3SPACE FASTRAK system in order to determine the participant's accurate placement in the MEG dewar. During the experimental sessions, participants laid supine in a dimly lit magnetically shielded room (Vacuumschmelze GmbH & Co. KG, Hanau, Germany). Continuous MEG recording was executed using a 160-channel axial gradiometer whole-head system (Kanazawa Institute of Technology, Kanazawa, Japan), and data was sampled at 1000 Hz (60 Hz online notch filter, DC-200 Hz recording bandwidth).

5.2.5. MEG analysis

MEG data was analyzed by customized code in MATLAB R2020b with MNE-Python 1.5.1 (Gramfort et al., 2013). First, noisy and flat channels were identified for each participant with

Maxwell filtering, and these channels were interpolated with the MNE algorithm (no channels in the regions-of-interest described below required interpolation). Then the environmental magnetic interferences were suppressed using temporal signal space separation (tSSS; Taulu & Simola, 2006). A low-pass infinite impulse response (IIR) filter with an upper cutoff frequency of 40 Hz was applied to the MEG data. Ocular and cardiac artifacts were removed using independent component analysis (ICA; fastica algorithm); 2-5 ICA components were removed from each participant. Epochs of -200:1200ms time-locked to the presentation of sample stimuli in each trial were extracted. The 200 ms pre-stimulus interval was used as the baseline interval. Trials with a maximum peak-to-peak signal amplitude higher than 3000 fT were excluded. Participants with less than 40 valid trials per condition were excluded from data analysis (1 for each experiment, see Participants). For the analyzed participants, only 0.7% (Experiment 5.1), 0.1% (Experiment 5.2, colored squares), and 0.4% (Experiment 5.2, complex shapes) of all trials were excluded. For each participant and each condition (i.e., a certain load condition in a certain feature block), we calculated the mean event-related field (ERF) response for each channel.

The first 20 trials in each condition of Experiment 5.1 were used to determine regions-of-interest (ROIs) across both experiments, and as such were excluded from all subsequent analyses. The contrast used for ROI selection was the mean ERF difference between larger (load 4 and 6) vs. smaller (load 1 and 2) conditions (Figure 5.2). In line with previous reports (Mitchell & Cusack, 2011; Robitaille et al., 2009, 2010), we observed two opposing maxima in posterior channels. We took three channels each from the two maxima to construct right and left parietal ROIs (right: MEG054, MEG055, MEG071; left: MEG051, MEG050, MEG076), and computed mean ERFs for each from 400:1200 ms post-stimulus-onset to capture the VWM maintenance (as opposed to perception) period.

Comparisons between conditions were statistically evaluated with repeated-measures ANOVA (using the Greenhouse-Geisser whenever the sphericity assumption was not met), and post hoc comparisons were Bonferroni-corrected. In order to better evaluate whether there was a response plateau beyond the capacity limit, we also administered Bayesian repeated-measures ANOVA with a uniform prior, and the post hoc comparisons were administered with Bayesian paired samples t-test (with a default Cauchy prior). All statistical analyses reported in this paper were conducted with JASP 0.18.1 (JASP Team, 2023) and Bayesian factors (BFs) were interpreted following (van Doorn et al., 2021): A BF higher than 10 or lower than 1/10 is considered “strong”, between 3-10 or 1/10-1/3 is considered “moderate”, and between 1/3-3 is considered “weak”.

5.3. Results

5.3.1. Experiment 5.1

5.3.1.1. Behavioral results

Repeated measures ANOVA revealed a significant main effect of load, $F(1.80, 34.28) = 30.22, p < 0.001$. Post-hoc comparisons (Bonferroni corrected) revealed a significant difference between loads 1 vs. 2 ($t(19) = -5.75, p < 0.001$), 1 vs. 4 ($t(19) = -8.59, p < 0.001$), 1 vs. 6 ($t(19) = -7.84, p < 0.001$) and 2 vs. 4 ($t(19) = -2.84, p = 0.038$). The difference between 2 vs. 6 ($t(19) = -2.10, p = 0.24$) did not reach statistical significance, nor did the difference between 4 vs. 6 ($t(19) = 0.74, p = 1$). Bayesian repeated measures ANOVA revealed strong evidence for the effect of load, $BF_{incl} = 1.66 \times 10^{10}$. There was strong evidence for a difference between 1 vs. 2 ($BF_{10} = 1.09 \times 10^{12}$), 1

vs. 4 ($BF_{10} = 1.73 \times 10^6$), 1 vs. 6 ($BF_{10} = 4.13 \times 10^3$), and 2 vs. 4 ($BF_{10} = 16.11$), and moderate evidence against a difference between 4 vs. 6 ($BF_{10} = 0.28$), suggesting a plateau of Cowan's K beyond a load of ~ 4 . The evidence for a difference between 2 vs. 6 was weak ($BF_{10} = 0.79$).

5.3.1.2. MEG results

The repeated measures ANOVA for MEG responses in the right ROI (Figure 5.2) showed a statistically significant main effect of load, $F(1.98, 37.52) = 4.06$, $p = 0.026$. Post-hoc comparisons (Bonferroni corrected) revealed a significant difference between loads 1 vs. 4 ($t(19) = -2.73$, $p = 0.050$) and 1 vs. 6 ($t(19) = -3.23$, $p = 0.012$). The other comparisons did not reach statistical significance (p 's > 0.5). Bayesian repeated measures ANOVA revealed moderate evidence for the effect of load, $BF_{\text{incl}} = 4.15$. There was moderate evidence for the difference between 1 vs. 6 ($BF_{10} = 3.71$), as well as moderate evidence against a difference between 4 vs. 6 ($BF_{10} = 0.27$). Evidence for other differences was weak ($1/3 < BF_{10} < 3$). We note that our number of trials per load condition (~ 30) here is rather low compared to previous MEG studies of this kind ((Robitaille et al., 2009): 384 trials/condition; (Robitaille et al., 2010): 200 trials/condition; (Mitchell & Cusack, 2011): 120 trials/condition) but with a similar sample size. Below we pool all 40 participants across the two experiments for a higher-powered analysis.

In the left ROI, the main effect of load was not statistically significant, $F(3, 57) = 2.15$, $p = 0.10$. Evidence for the effect of load was also weak, $BF_{\text{incl}} = 0.62$.

Exp. 1 (N=20)

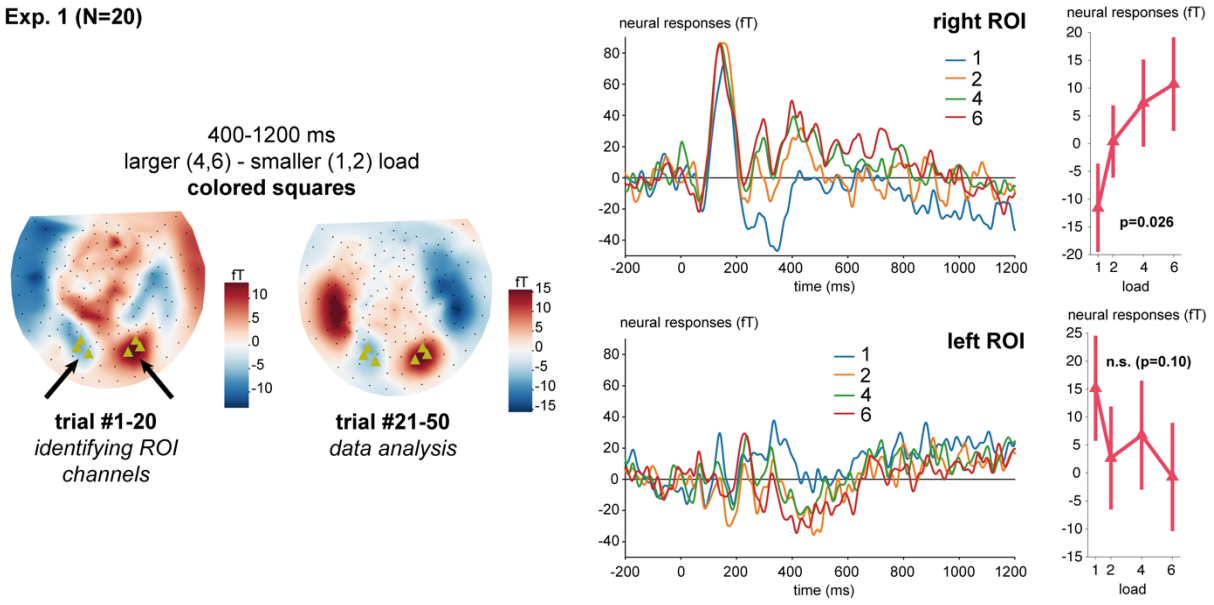


Figure 5.2 A summary of the MEG results in Experiment 5.1.

Error bars stand for standard error. ROI: region of interest. The ROI channels are marked with green triangles.

5.3.2. Experiment 5.2

5.3.2.1. Behavioral results

A repeated measures ANOVA crossing load and feature (colored square vs. complex shape), showed a significant main effect of load on Cowan's K, ($F(1.51, 28.76) = 16.38, p < 0.001$), a main effect of feature ($F(1, 19) = 50.45, p < 0.001$), and a significant interaction ($F(2.13, 40.48) = 7.35, p = 0.002$). Bayesian repeated measures ANOVA also revealed strong evidence for the effect of load ($BF_{\text{incl}} = 2.42 \times 10^7$), feature ($BF_{\text{incl}} = 4.61 \times 10^5$) and the interaction effect ($BF_{\text{incl}} = 715.77$). The main effect of feature reflected overall higher performance for colored squares than complex shapes, in line with previous observations (e.g., Y. Xu & Chun, 2006).

We followed up the interaction with one-way ANOVAs at each level of feature. In colored square trials, the main effect of load was statistically significant, $F(1.55, 29.42) = 18.16$,

$p < 0.001$. Post-hoc comparisons (Bonferroni corrected) revealed significant differences between loads 1 vs. 2 ($t(19) = -3.67, p = 0.003$), 1 vs. 4 ($t(19) = -6.92, p < 0.001$), 1 vs. 6 ($t(19) = -5.62, p < 0.001$) and 2 vs. 4 ($t(19) = -3.25, p = 0.012$). The differences between 2 vs. 6 ($t(19) = -1.95, p = 0.34$) and 4 vs. 6 ($t(19) = 1.30, p = 1$) were not statistically significant. Bayesian repeated measures ANOVA also suggested strong evidence for the effect of load, $BF_{incl} = 1.82 \times 10^6$. There was strong evidence for the difference between 1 vs. 2 ($BF_{10} = 1.36 \times 10^{11}$), 1 vs. 4 ($BF_{10} = 1.08 \times 10^5$), 1 vs. 6 ($BF_{10} = 63.91$) and 2 vs. 4 ($BF_{10} = 87.07$). Evidence for the difference between 2 vs. 6 ($BF_{10} = 0.61$) and 4 vs. 6 ($BF_{10} = 0.44$) was weak.

In complex shape trials, the main effect of load was also significant ($F(1.99, 37.74) = 7.01, p = 0.003$). Post-hoc comparisons (Bonferroni corrected) revealed significant differences between loads 1 vs. 2 ($t(19) = -2.74, p = 0.049$) and 1 vs. 4 ($t(19) = -4.53, p < 0.001$). The other differences were not statistically significant (p 's > 0.09). Bayesian repeated measures ANOVA also suggested strong evidence for the effect of load, $BF_{incl} = 93.00$, and for the difference between 1 vs. 2 ($BF_{10} = 6.03 \times 10^4$) and 1 vs. 4 ($BF_{10} = 136.50$). Evidence for the other differences was weak ($1/3 < BF_{10} < 3$).

In sum, the colored square and complex shape comparisons patterned similarly in showing main effects of load. The interaction between feature and load appeared to be driven only by the slight drop observed in performance for complex shapes at load 6.

5.3.2.2. MEG results

The 2×4 (feature $\times$ load) repeated measures ANOVA for MEG responses in the right ROI (Figure 5.3) showed a significant main effect of load ($F(1.50, 28.47) = 6.31, p = 0.010$). The main effect of feature ($F(1, 19) = 0.086, p = 0.77$) and the interaction effect ($F(3, 57) = 0.90, p = 0.45$)

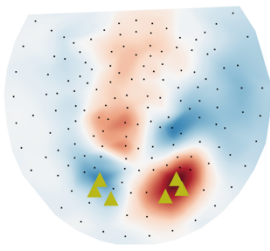
were not significant. Post-hoc comparisons between load conditions (Bonferroni corrected) revealed a significant difference between loads 1 vs. 4 ($t(19) = -3.68, p = 0.003$) and 1 vs. 6 ($t(19) = -3.79, p = 0.002$); the other differences were not significant (p 's ≥ 0.30). Bayesian repeated measures ANOVA revealed moderate evidence against the inclusion of feature ($BF_{\text{incl}} = 0.22$) and the interaction effect ($BF_{\text{incl}} = 0.15$) in the model, and strong evidence for the inclusion of load ($BF_{\text{incl}} = 29.78$) in the model. There was moderate evidence for the difference between 1 vs. 2 ($BF_{10} = 4.80$) and 1 vs. 6 ($BF_{10} = 29.58$) and strong evidence for the difference between 1 vs. 4 ($BF_{10} = 44.75$); there was moderate evidence against the difference between 4 vs. 6 ($BF_{10} = 0.17$), suggesting a plateau beyond ~ 4 . Evidence for the difference between 2 vs. 4 ($BF_{10} = 1.80$) and 2 vs. 6 ($BF_{10} = 2.18$) was weak.

For the left ROI, the main effect of load was significant, ($F(1.97, 37.36) = 4.75, p = 0.015$). The main effect of feature ($F(1, 19) = 1.37, p = 0.26$) and the interaction effect ($F(3, 57) = 0.42, p = 0.74$) were not statistically significant. Post-hoc comparisons (Bonferroni corrected) revealed a significant difference between loads 1 vs. 2 ($t(19) = 2.79, p = 0.043$) and 1 vs. 4 ($t(19) = 3.58, p = 0.004$); the other differences were not statistically significant (p 's > 0.12). Bayesian repeated measures ANOVA revealed moderate evidence against the inclusion of feature ($BF_{\text{incl}} = 0.29$) and the interaction effect ($BF_{\text{incl}} = 0.11$) in the model, and moderate evidence for the inclusion of load ($BF_{\text{incl}} = 4.79$) in the model. There was moderate evidence for the difference between 1 vs. 2 ($BF_{10} = 22.98$) and 1 vs. 4 ($BF_{10} = 9.58$), and moderate evidence against the difference between 2 vs. 4 ($BF_{10} = 0.25$) and 2 vs. 6 ($BF_{10} = 0.19$), suggesting a plateau beyond ~ 2 . Evidence for the difference between 1 vs. 6 ($BF_{10} = 2.36$) and 4 vs. 6 ($BF_{10} = 0.79$) was weak.

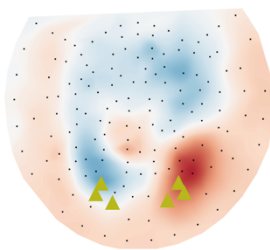
Exp. 2 (N=20)

400-1200 ms
larger (4,6) - smaller (1,2) load

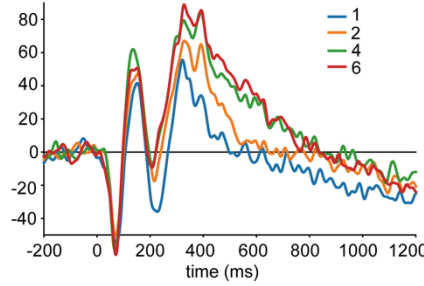
colored squares



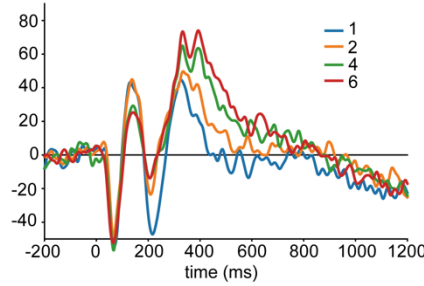
complex shapes



right ROI: colored squares
neural responses (fT)

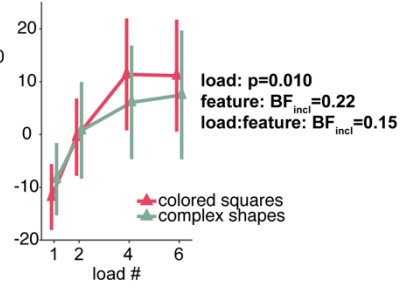


right ROI: complex shapes
neural responses (fT)

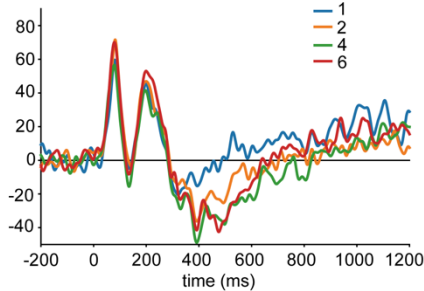


400-1200 ms

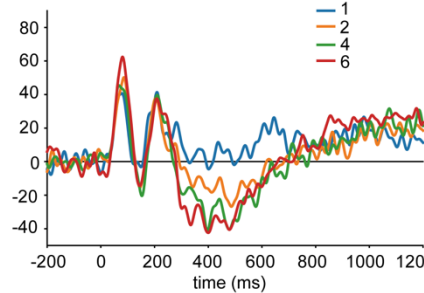
neural responses (fT)



left ROI: colored squares
neural responses (fT)



left ROI: complex shapes
neural responses (fT)



neural responses (fT)

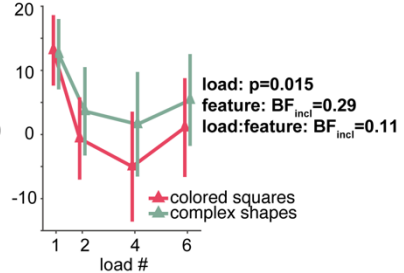


Figure 5.3 A summary of the MEG results in Experiment 5.2.

Error bars stand for standard error. ROI: region of interest. The ROI channels are marked with green triangles.

Pooling colored square data from all 40 participants across Experiments 1 and 2 together allowed us to conduct a high-powered analysis to determine whether the two ROIs exhibited the monotonic increase up to 4 followed by a plateau, as predicted if there is a limited pointer capacity of ~ 3 (Figure 5.4). For the right ROI, the Bayesian paired-sample t-test revealed moderate evidence for differences between loads 1 vs. 2 ($BF_{10} = 16.07$), 1 vs. 4 ($BF_{10} = 28.00$), 2 vs. 6 ($BF_{10} = 5.50$) and strong evidence for a difference between 1 vs. 6 ($BF_{10} = 134.59$). There was also moderate evidence against a difference between 4 vs. 6 ($BF_{10} = 0.19$). Evidence for the difference between 2 vs. 4 was weak ($BF_{10} = 2.18$) but was still in the direction of a difference. Overall, this suggests a plateau beyond ~ 4 items for the right ROI.

For the left ROI, the Bayesian paired-sample t-test revealed strong evidence for a difference between loads 1 vs. 2 ($BF_{10} = 12.76$) and moderate evidence for a difference between 1 vs. 6 ($BF_{10} = 7.94$), and moderate evidence against a difference between loads 2 vs. 4 ($BF_{10} = 0.17$), 2 vs. 6 ($BF_{10} = 0.17$), and 4 vs. 6 ($BF_{10} = 0.17$). Evidence for the difference between 1 vs. 4 was weak ($BF_{10} = 2.44$) but was still in the direction of a difference. Overall, this suggests a plateau beyond ~ 2 items for the left ROI.

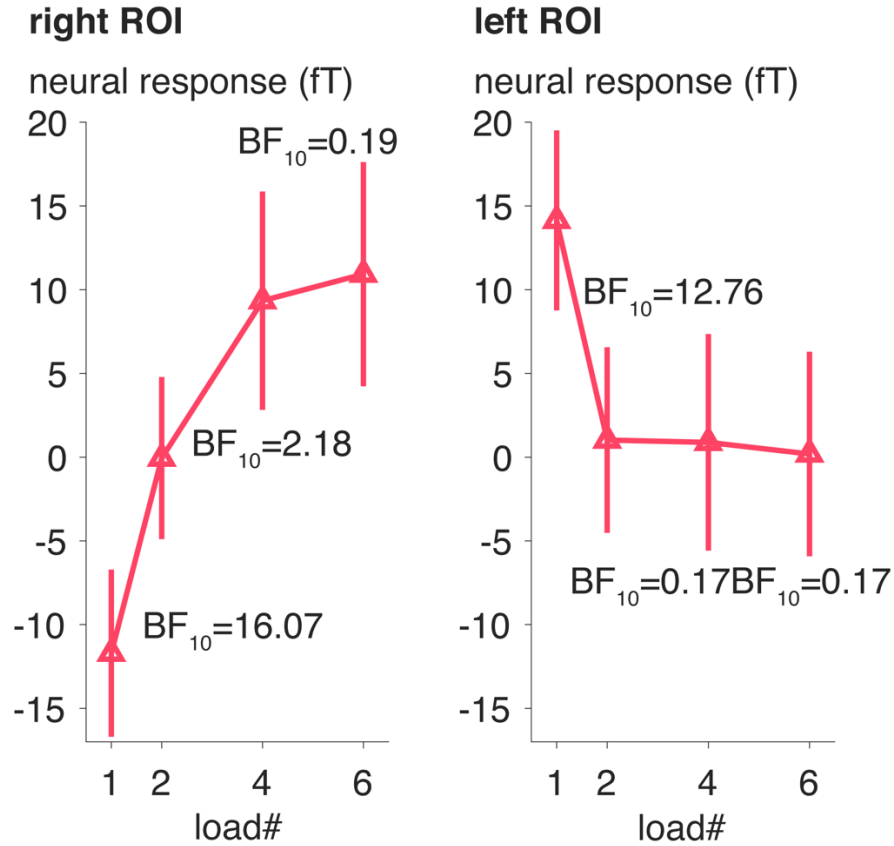


Figure 5.4 MEG responses during the VWM period (400:1200 ms) for the colored squares in the right and left ROIs, pooled across all 40 participants in Experiments 5.1 and 5.2. Error bars stand for standard error. ROI: region of interest.

5.4. Discussion

5.4.1. Conclusion

In this study, we observed univariate responses in MEG over right posterior sites (the right posterior activity, RPA) that met all three criteria for VWM pointers: load-dependent, capacity-limited, and content-free. This neural response changed monotonically as VWM load increased, reached a plateau at ~ 4 objects (in line with the classic VWM capacity limit), and was insensitive to the difference in type and number of features bound to each object in the colored

square vs. complex shape conditions even though these differences impacted behavioral performance. Given its maxima over right posterior channels, this response is likely to be of posterior parietal origin, in line with previous fMRI effects attributed to VWM pointers (Naughtin et al., 2016; Y. Xu & Chun, 2006), and consistent with the involvement of the posterior parietal cortex in VWM maintenance in general (Matsuyoshi et al., 2010; Todd & Marois, 2004; Y. Xu, 2023). This awaits to be confirmed by future studies with higher spatial-temporal resolution (e.g., combining MEG with structural MRI, intracranial recordings). Our current study serves as an important complement to prior fMRI studies, in that we confirmed VWM pointer-like responses during the VWM maintenance time window in particular. The right-lateralization of our current effects invites future neuroimaging research on the lateralization of VWM substrates. Intriguingly, this is in line with some neuropsychological reports suggesting that unilateral *right* posterior parietal lesions are sufficient for an impaired VWM (for a review and nuanced discussions see Olson & Berryhill, 2009).

We also found that left posterior channels exhibited a load-dependent, capacity-limited and content-free effect, but this response plateaued at only ~ 2 objects. While this different response profile awaits future investigation, it could reflect cognitive effort related to pattern-separating more than one feature of the same type (e.g., when one needs to hold multiple colors in VWM) or mechanisms related to attention allocation (Todd & Marois, 2005).

5.4.2. Limitations and future directions

As this is only the first MEG study exploring content-free neural signatures in VWM, future studies are certainly needed in order to address some limitations of the current work. Firstly, in

our current study the number of trials per condition was lower than previous MEG studies on VWM load (as noted in the “MEG results” section). It is yet unknown how many trials is sufficient to detect differences across VWM load conditions for the posterior activities in MEG (Ngiam et al., 2021). We only obtained weak evidence for the difference between 2 and 4 objects in the right posterior ROI ($BF_{10} = 2.18$), and more evidence is needed to confirm the existence of this difference. Secondly, in our current study, the number of pointers may covary with e.g., the number of locations in or the luminance of the memorandum display (Thyer et al., 2022), which could be controlled for in future studies. Furthermore, it would be interesting to examine whether RPA could predict individual differences in VWM capacity (Luria et al., 2016; Vogel & Machizawa, 2004).

6. Chapter 6

In this chapter, we report evidence demonstrating that inferior parietal regions host pointer-sensitive signals while being insensitive to feature-based load, which extends from VWM formation to maintenance, in contrast to more superior parietal regions that were only transiently sensitive to feature-based load during VWM formation.

6.1. Introduction

In the previous chapter, we reported a sensor-level MEG response, the right posterior activity (RPA), tracking the number of pointers while remaining insensitive to the number of features per object. This is in line with prior reports with fMRI, showing that the pointers and features are represented in dissociable brain regions (Naughtin et al., 2016; Y. Xu, 2009; Y. Xu & Chun, 2006): some regions are only sensitive to pointer-based load, while some other regions are sensitive to the number of features to be held. These studies have suggested that the inferior IPS (intraparietal sulcus) represents only pointers, while the superior IPS is sensitive to the number of features. In this Chapter, we ask a follow-up question: What are the spatiotemporal dynamics of this dissociation?

On the one hand, fMRI cannot clearly disentangle responses during perception or VWM formation vs. VWM maintenance due to its limited temporal resolution. This calls for the usage of methods with better spatiotemporal resolution such as MEG.

On the other hand, this research question is deeply relevant to the function of pointers. As pointers serve as neural indexes for the corresponding objects, they are hypothesized to be

implemented by sustained neural activities (Awh & Vogel, 2025). Since their function is to “point to” features, features do not have to be realized by sustained activities, as long as the pointers remain capable of fetching this information (via e.g., intracellular storage of feature address) when needed. As neural activations are energy-consuming, this appears to be an energy-efficient way to hold object information in moment-to-moment cognition. However, this difference across the representation of pointers and features has not been empirically tested. All the possibilities are outlined in Figure 6.1.

In this MEG study, with carefully digitized head shape data, we are able to offer analyses in the source space, in order to answer these questions.

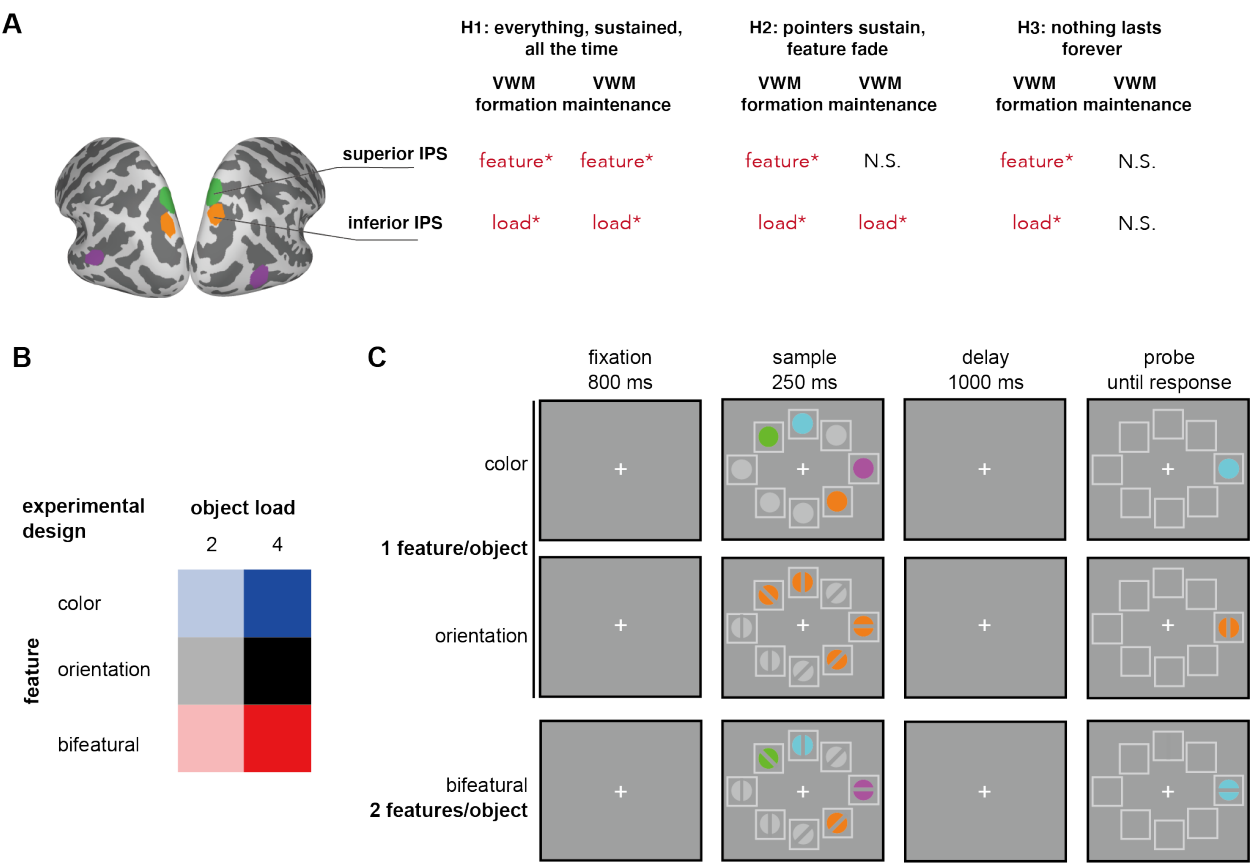


Figure 6.1 Experimental hypotheses (A), design (B) and procedure (C) of Experiment 6.1.

6.2. Experiment 6.1: Methods

6.2.1. Participants

27 participants (age 19-26, 14 female) entered our data analysis. One additional participant was excluded because of chance level (49%) performance in some blocks. The participants received monetary reimbursement or course credits for their participation. The procedures are approved by the IRB of University of Maryland College Park.

6.2.2. Procedure

Each participant was presented with 600 trials: 200 color trials, 200 orientation trials, and 200 bifeatural trials. Within each feature condition (color, orientation, bifeatural), half of the trials had a load of two objects and half had a load of four objects. These 600 trials were evenly divided into six blocks: two color blocks, two orientation blocks, and two bifeatural blocks. Each block contained 100 trials, and the order of the blocks was randomized for each participant, with the constraint that each type of blocks appeared once in the first three and last three blocks respectively. In other words, the trials in each block belonged to the same feature condition (color, orientation, bifeatural) but were intermixed in terms of memory load (two vs. four objects).

In each trial, the participants were first presented with a fixation cross with a duration of 800 ms against a grey background, followed by a sample memory array with a duration of 250 ms (see Figure 6.1). The sample consisted of two or four unique target objects. For the color trials, the colors were sampled without replacement from the following seven colors: red

[255,0,0], green [0,255,0], blue [0,0,255], yellow [255,255,0], purple [255,0,255], teal [0,255,255], and orange [255,128,0]. For the orientation trials, the bar orientation was sampled without replacement from the following four orientations: 0°, 45°, 90°, 135°. The color of the disk behind these target orientations was randomly sampled from the above colors for each participant and was kept consistent throughout that experimental session for that participant. For bifeatural trials, each target object was a combination of a color and an orientation sampled in the aforementioned way. The target objects ($2.1^\circ \times 2.1^\circ$) occupied two or four out of all eight boxes (light gray [191,191,191], $2.5^\circ \times 2.5^\circ$ each), located at eight fixed locations evenly distributed 4.5° from the center of the screen. The box contours were employed to discourage grouping (Xie & Zhang, 2017a). Filler gray disks ([166,166,166]) occupied the rest of the locations for all trials, additionally with an orientation bar at a random orientation for the orientation and bifeatural trials. The orientation bars had a width of 0.5° . A blank screen delay of 1000 ms followed the sample, after which the probe appeared on the screen. At the probe, only one object was presented at one of the locations previously occupied by a target object. Half of the trials had a match probe, and the other half had a change probe. On “match” trials, this object was the same one as the one in the sample for that location; on “change” trials, this probe object was randomly selected from other target features used in this block. For example, for a change trial in the color condition, participants were probed with a color that appeared in the sample but not at the probed location. For a change trial in the bifeatural condition, participants were probed with an object with the right color for that location but a wrong – yet previously presented – orientation (or the right orientation yet a wrong color). Participants pressed one button to indicate “match”, and pressed another button to indicate “change” (buttons randomized for each participant), with no time limit for response. After their response, they were shown a line of

asterisks on the screen, where they would press any button to proceed to the next trial in a self-paced manner. Prior to the experiment, participants were instructed to blink only during this period, and to avoid blinking during the trials. While the fixation cross was presented on the screen, the participants were instructed to fixate at it. They were also instructed against verbal rehearsal. At least 24 practice trials were run for each participant prior to the experiment. The setup was broadly consistent with the EEG study of Thyer et al. (2022), and the experiment was run in MATLAB R2009a with Psychtoolbox 3 (Brainard, 1997; Kleiner et al., 2007).

6.2.3. Behavioral analyses

Accuracy was calculated for each participant for each condition. The responses were subject to a 2×3 repeated-measures ANOVA (load: two, four $\times$ feature: color, orientation, bifeatural). Conditions were compared with a paired t test whenever necessary. The statistics reported in the current chapter are all based on two-tailed tests.

6.2.4. MEG recording

Prior to recording, five head position indicator coils were affixed to each participant's head, and the position of these coils relative to the nasion and tragus, as well as the participant's head-shape, was digitized using a Polhemus 3SPACE FASTRAK system in order to determine the participant's accurate placement in the MEG dewar. During the experimental sessions, participants laid supine in a dimly lit magnetically shielded room (Vacuumschmelze GmbH & Co. KG, Hanau, Germany). Continuous MEG recording was executed using a 160-channel axial

gradiometer whole-head system (Kanazawa Institute of Technology, Kanazawa, Japan), and data was sampled at 1000 Hz (60 Hz online notch filter, DC-200 Hz recording bandwidth).

6.2.5. MEG analyses

6.2.5.1. Sensor-level analysis

MEG sensor-level data was analyzed by customized code in MATLAB R2020b with MNE-Python 1.5.1 (Gramfort et al., 2013), with a similar procedure as (Yu & Lau, 2025). First, noisy and flat channels were identified for each participant with Maxwell filtering, and these channels were interpolated with the MNE algorithm. Then the environmental magnetic interferences were suppressed using temporal signal space separation (tSSS; Taulu & Simola, 2006). A low-pass infinite impulse response (IIR) filter with an upper cutoff frequency of 40 Hz was applied to the MEG data. Ocular and cardiac artifacts were removed using independent component analysis (ICA; fastica algorithm); 2-7 ICA components were removed from each participant. Epochs of -200:1200ms time-locked to the presentation of sample stimuli in each trial were extracted. The 200 ms pre-stimulus interval was used as the baseline interval. Trials with a maximum peak-to-peak signal amplitude higher than 3000 fT were excluded. For each participant, the number of epochs per condition was matched with the constraint of minimizing timing differences (the *mintime* algorithm). For each participant and each condition (i.e., a certain load condition in a certain feature block), we calculated the mean event-related field (ERF) response for each channel. Following Yu & Lau (2025), the RPA (right posterior activities) ROI is defined as the following channels: MEG054, MEG055, MEG071. Mean ERFs were calculated within this ROI.

Comparisons across conditions were statistically evaluated with repeated-measures ANOVA, and post hoc comparisons were Bonferroni-corrected.

6.2.5.2. Source space analysis

For the source analyses, the ‘fsaverage’ template brain from the FreeSurfer software package (<http://surfer.nmr.mgh.harvard.edu>) was used as an approximation for individual anatomy (four-fold icosahedral subdivision), and digitizer data was used to manually coregister individual head position to the template brain. A three-compartment boundary element model computed for this template brain with the linear collocation approach was used in the forward calculation (Hämäläinen & Sarvas, 1989; Mosher et al., 1999). Then we created an inverse operator, for each participant and for each condition. Noise covariance estimates were derived from data recorded in the 200 msec baseline period. The orientations of the dipoles subjected to a loose constraint with weighting factor 0.2 (F. H. Lin et al., 2006), and depth was weighted at 0.8. Then the inverse operator was applied to the corresponding averaged evoked responses for the corresponding participant and condition, with regularization $\lambda = 1/9$ and the dSPM algorithm (Dale et al., 2000). Individual estimates were then smoothed using 7 iterative steps to spread estimated activity to neighboring vertices.

The source space ROIs were a priori defined based on the fMRI study of Xu & Chun (2006), converting the original Talairach coordinates to MNI coordinates (Papademetris et al., 2006): inferior IPS, left [-25,-74,30], right [25,-69,36]; superior IPS, left [-14,-73,51], right [15,-72,55]; LOC, left [-46,-50,4], right [41,-68,2]. Each ROI grew from the nearest vertex to these MNI coordinates, within a maximum radius of 10 mm on the white matter surface based on geodesic distance, constrained to the corresponding hemisphere. The responses in each ROI were

averaged across two time windows: 200-400 ms (VWM formation) and 400-1200 ms (VWM maintenance). For each time window, the responses were subject to a 2×3 repeated-measures ANOVA (load: two, four $\times$ feature: color, orientation, bifeatural). For visual examination, a contrast map based on t-values was created by comparing across load conditions (paired t test) vertex by vertex with a threshold of $\alpha = 0.05$ across a series of 200 ms time windows from 200 ms to 1200 ms.

6.3. Results

6.3.1. Behavioral results

For each of the six conditions, participants had above-chance performance, with bootstrapped 95% CI: color, 2 (93.1%-96.0%), color, 4 (77.2%-85.1%), orientation, 2 (83.9%-89.7%), orientation, 4 (66.4%-71.7%), bifeatural, 2 (84.6%-89.1%), bifeatural, 4 (64.9%-69.6%). A 2 (load: two, four) $\times$ 3 (feature: color, orientation, bifeatural) repeated-measures ANOVA revealed a significant main effect of load, $F(1,26) = 403.28, p < 0.001, \eta_p^2 = 0.94$, a significant main effect of feature, $F(2,52) = 81.96, p < 0.001, \eta_p^2 = 0.76$, as well as a significant interaction effect, $F(2,52) = 5.37, p = 0.008, \eta_p^2 = 0.17$. The interaction effect appeared to be driven by a larger amount of impact of load on orientation and bifeatural trials compared to the color trials: comparing the drop in accuracy across load 2 and 4 conditions, we observed a significant difference between color and orientation trials (paired t test, $t(26) = -2.13, p = 0.04$, Cohen's $d = -0.41$), and a significant difference between color and bifeatural trials ($t(26) = -2.98, p = 0.006$, Cohen's $d = -0.57$), but not a significant difference between orientation and bifeatural trials

($t(26) = -1.07, p = 0.30$, Cohen's $d = -0.21$). Regardless, the drop in accuracy for color trials across loads 2 and 4 was statistically significant, $t(26) = 7.89, p < 0.001$, Cohen's $d = 1.52$.

6.3.2. Sensor-level results

Similar to our prior study (Yu & Lau, 2025), we analyzed the ERF responses in the 400-1200 ms time window in the RPA ROI. A 2 (load: two, four) $\times$ 3 (feature: color, orientation, bifeatural) repeated-measures ANOVA revealed a significant main effect of load, $F(1,26) = 21.88, p < 0.001, \eta_p^2 = 0.46$, as well as a significant interaction effect, $F(2,52) = 4.39, p = 0.02, \eta_p^2 = 0.14$. The main effect of feature was not statistically significant, $F(2,52) = 1.51, p = 0.23, \eta_p^2 = 0.06$. See Figure 6.2.

Follow-up paired t-tests across two load conditions revealed a significant difference for color trials, $t(26) = -4.23, p < 0.001$, Cohen's $d = -0.82$, as well as for the bifeatural trials, $t(26) = -3.19, p = 0.004$, Cohen's $d = -0.61$. The difference was only numerical for the orientation trials, $t(26) = -1.46, p = 0.16$, Cohen's $d = -0.28$.

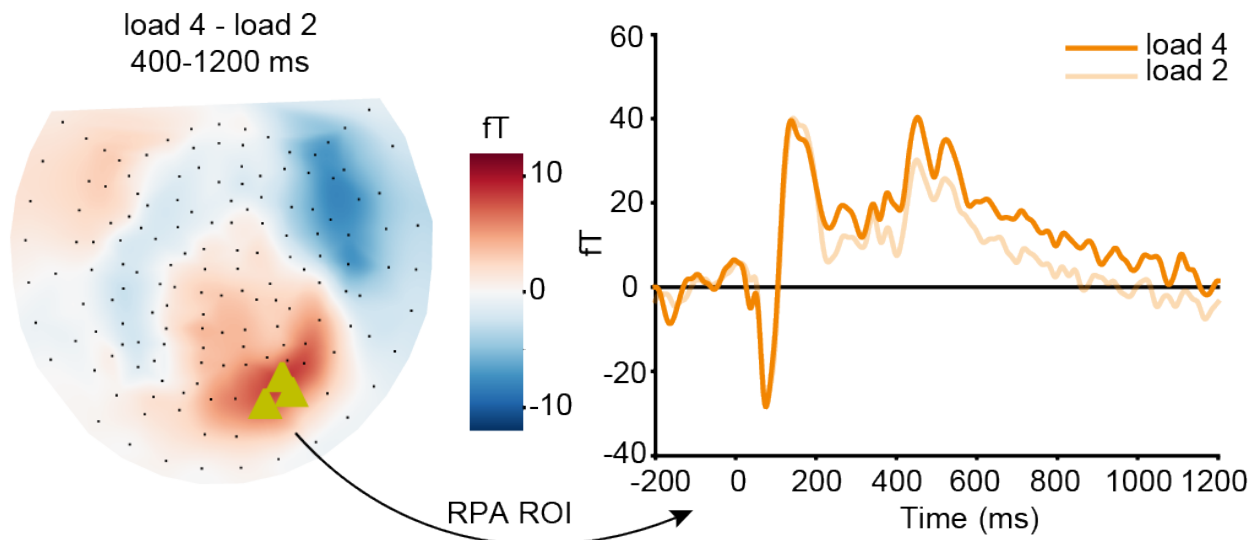


Figure 6.2 ERF responses at the RPA ROI for the two load conditions.

6.3.3. Source-level analysis

We ran 2 (load: two, four) $\times$ 3 (feature: color, orientation, bifeatural) repeated-measures ANOVA on three a priori ROIs (inferior IPS, superior IPS and LOC) across two time windows (200-400 ms: VWM formation; 400-1200 ms: VWM maintenance). See Figure 6.3.

For inferior IPS, a main effect of load emerged for VWM formation ($F(1,26) = 4.13, p = 0.05, \eta_p^2 = 0.14$), extending into VWM maintenance ($F(1,26) = 5.92, p = 0.02, \eta_p^2 = 0.19$). For VWM formation, neither the main effect of feature nor the interaction effect was statistically significant: feature, $F(2,52) = 0.60, p = 0.55, \eta_p^2 = 0.02$; interaction effect, $F(2,52) = 0.80, p = 0.46, \eta_p^2 = 0.03$. During VWM maintenance, neither the main effect of feature nor the interaction effect was statistically significant: feature, $F(2,52) = 0.66, p = 0.52, \eta_p^2 = 0.03$; interaction effect, $F(2,52) = 0.40, p = 0.67, \eta_p^2 = 0.02$.

For superior IPS, a main effect of feature manifested for VWM formation ($F(2,52) = 3.60, p = 0.03, \eta_p^2 = 0.12$). The main effect of load was only marginally significant, $F(1,26) = 3.61, p = 0.07, \eta_p^2 = 0.12$, and the interaction effect was non-significant, $F(2,52) = 0.09, p = 0.91, \eta_p^2 = 0.004$. Post hoc pairwise comparison (Bonferroni-corrected) across the three feature conditions revealed a significant difference between color and bifeatural conditions, $t(26) = -2.68, p = 0.03$, Cohen's $d = -0.16$. The differences between the other two pairs of feature conditions were not statistically significant. During VWM maintenance, none of the effects were statistical significant (F values < 2.48 , p values > 0.12).

For LOC, a main effect of feature manifested for VWM formation ($F(2,52) = 3.44, p = 0.04, \eta_p^2 = 0.12$). Neither the main effect of load nor the interaction effect was significant: load,

$F(1,26) = 0.48, p = 0.50, \eta_p^2 = 0.02$; the interaction effect, $F(2,52) = 0.11, p = 0.89, \eta_p^2 = 0.004$.

Post hoc pairwise comparison (Bonferroni-corrected) across the three feature conditions revealed a significant difference between color and bifeatural conditions, $t(26) = -2.62, p = 0.03$, Cohen's $d = -0.28$. The differences between the other two pairs of feature conditions were not statistically significant. During VWM maintenance, there was a significant interaction effect between feature and load, $F(2,52) = 3.18, p = 0.05, \eta_p^2 = 0.11$. The main effect of load was marginally significant, $F(1,26) = 3.48, p = 0.07, \eta_p^2 = 0.12$, and the main effect of feature was not significant, $F(2,52) = 0.37, p = 0.69, \eta_p^2 = 0.01$. The interaction effect was driven by differential impact of load on the responses: for color trials, paired t-test (Bonferroni-corrected) revealed a significant difference across the two load conditions, $t(26) = -2.63, p = 0.04$, Cohen's $d = -0.38$. The load-driven difference was not statistically significant for orientation trials and bifeatural trials ($|t \text{ values}| < 2$, Bonferroni-corrected $p \text{ values} > 0.18$).

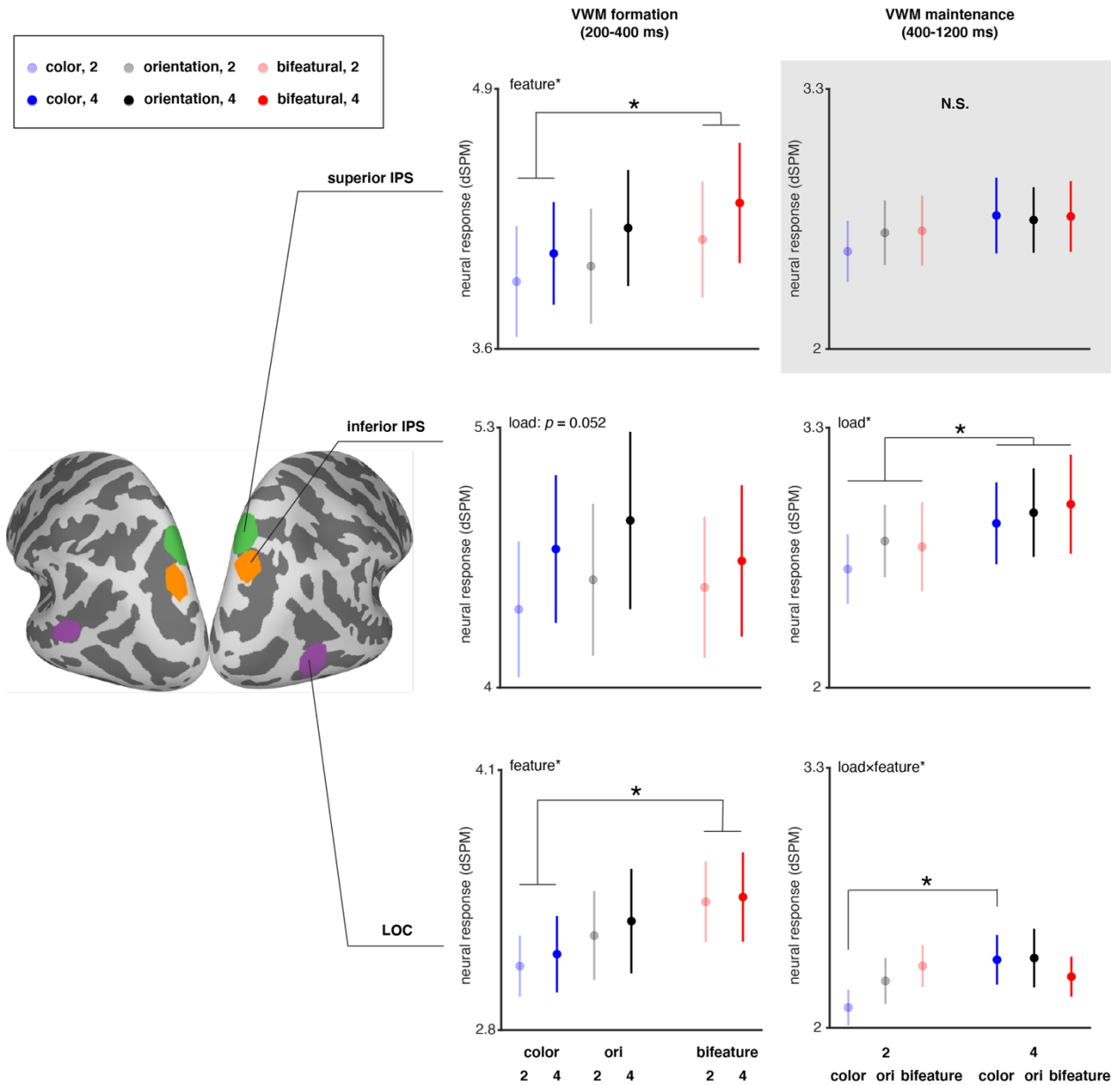


Figure 6.3 Responses during VWM formation (200-400 ms) and maintenance (400-1200 ms) in three a priori ROIs: inferior IPS, superior IPS and LOC.

*: $p \leq 0.05$.

6.4. Discussion and conclusion

The aim of this experiment was to study the spatiotemporal dynamics of pointer-based and feature-based neural representations across visual working memory formation and maintenance.

The results of our MEG source space analysis allow us to draw several conclusions. First, our MEG results replicated prior fMRI reports (Naughtin et al., 2016; Y. Xu, 2009; Y. Xu & Chun, 2006): while MEG responses in superior IPS and LOC were sensitive to the number of features, MEG responses in inferior IPS were primarily modulated by the number of objects – the number of pointers taxed. This is consistent with the proposal that superior IPS is responsible for object individuation and inferior IPS is responsible for object identification and thus is insensitive to the number of features per object (Y. Xu & Chun, 2009).

Second, the spatiotemporal resolution provided by MEG allowed us to deepen our understanding of this functional dissociation also in the temporal domain, by contrasting transient responses during initial VWM formation with longer-lasting responses during VWM maintenance. In contrast to a long-lasting effect of pointer-based load in the inferior IPS, superior IPS only transiently responded to feature-based load during VWM formation. This again suggests that in contrast with a transient object individuation signal for object individuation during VWM formation in superior IPS, inferior IPS represents a pointer-based signal that extends across VWM formation and maintenance. Taken together, our Hypothesis 2 was supported: that pointers persist, while features are less persistent across VWM maintenance.

Finally, the temporal dynamics of LOC were also intriguing: while it showed similar feature-sensitive responses as superior IPS during VWM formation, VWM maintenance no longer demonstrated a simple main effect between feature conditions but instead an interaction effect between feature and load. While this can be interpreted as a weakened feature-sensitive response, this is also in line with recent fMRI reports demonstrating a temporal progression of representational geometry in lateral occipital regions, due to modulations from the parietal regions (Y. Xu, 2023). The temporal dynamics of LOC representations may suggest its role in

initial visual analysis based on prior visual knowledge (J. G. Kim et al., 2009; Y. Xu, 2018), instead of binding via pointers during VWM maintenance.

In sum, these results support a critical role for inferior IPS in instantiating pointers during VWM and indicate that MEG, even with its lower spatial resolution than fMRI, is able to resolve activity in the parietal cortex from subtle neighboring regions with different functional roles. These results thus offer a new neural marker for investigating visual working memory with improved spatiotemporal precision.

7. Chapter 7²⁰

In Chapters 5 and 6, we replicated and extended prior findings in fMRI with MEG, where sensor-level and source-level univariate responses align well with the hypothesized properties of pointers in visual working memory, additionally offering much finer temporal resolution that is helpful to disentangle different processing stages. While prior relevant studies primarily utilized EEG and fMRI, our findings reveal the immense potential of MEG as an ideal neuroimaging technology to study pointers for its acceptable resolutions in both space and time.

Recent years have seen the development of OPM-MEG (e.g., Brickwedde et al., 2024), in contrast to the more widely-used SQUID-MEG. OPM-MEG, unlike SQUID-MEG, does not rely on constant helium supply, and can access signals closer to the scalp. This raises broad interest in OPM-MEG as potentially the ‘next generation’ MEG system, as these properties may enable wider applications (due to the high cost of helium supply), as well as better signal quality (due to spatial proximity to the scalp and thus to the neural sources). However, as a new technology, data pre-processing pipelines have been much less established compared to SQUID-MEG. Moreover, wearable OPM-MEG systems face unique challenges due to residual magnetic field in magnetically-shielded rooms despite offering unprecedented opportunities in its mobility (see Chapter 7.1 below), calling for innovative ways to denoise data.

In this chapter, we report methodological developments in pre-processing OPM-MEG signals, when limited number of sensors are available.

²⁰ This chapter reflects work in collaboration with Dongxue Zhang, Thomas Carey, Ernst Pöppel, Tilmann Sander and Ellen Lau. Based on Dongxue Zhang’s contribution to the experimental script, I and her will be co-first authors for relevant publications.

7.1. Introduction

Magnetoencephalography (MEG) is a non-invasive neuroimaging technology with both high spatial and temporal resolution. MEG studies therefore offer important complementary evidence in the field of cognitive neuroscience, compared to e.g., EEG (electroencephalography; with good temporal resolution but bad spatial resolution) and fMRI (functional magnetic resonance imaging; with good spatial resolution but bad temporal resolution). MEG also possesses largely underexplored clinical potential including e.g., the diagnosis of concussion (Davenport et al., 2022; Lee & Huang, 2012) and dyslexia (Dikker et al., 2020; Salmelin, 2007; Tarkiainen et al., 2003). MEG sites existing for more than 10 years typically make use of SQUID-MEG systems (SQUID: superconducting quantum interference device), but recent years have seen an increasing transition to OPM-MEG systems (OPM: optically pumped magnetometers), which possess several properties that are more desirable than SQUID-MEG systems (Borna et al., 2017; Boto et al., 2017; Brickwedde et al., 2024; Brookes et al., 2022; H. Cheng et al., 2024; Kominis et al., 2003; Qin & Gao, 2021; Sander et al., 2012; Seymour et al., 2022). First, OPM-MEG systems do not require the constant liquid helium supply necessary for SQUIDs, which is expensive and is easily affected by helium shortages. Second, OPM sensors can be much closer to the scalp compared to SQUIDs, allowing for higher signal amplitudes (which is critical for detecting the faint magnetic signals from the brain, generally at the scale of femtoteslas [fT], 10^{-15} T). For example, in the current study the OPM sensors were attached to an EEG cap, which is much closer to neural sources in the brain compared to SQUIDs, which are encapsulated in a dewar that cannot fit flexibly to the scalp (in this case the minimum distance to cortical sources would be ~ 2 -4 cm; Andersen et al., 2020; Iivanainen et al., 2020). Third, most current OPM-

MEG systems are much more mobile than SQUID-MEG systems, which may allow for mild movements during data collection. Mild movements are generally inevitable, especially for certain populations such as young children, posing challenges to the inclusion of more populations for SQUID-MEG systems. Moreover, a mobile setup allows for MEG recordings in more naturalistic settings.

However, what comes along with OPM-MEG's mobility is a low-frequency artifact driven by the interaction between the magnetometers and the remnant field gradients present in all magnetically shielded rooms (Jazbinšek et al., 2022). Although participants are instructed not to move their head during experiments, it is impossible for their heads to be completely motionless due to e.g., breathing. Even seemingly small movements would drive considerable artifacts at the magnetometers, resulting in low-frequency artifacts in the data that are believed to be up to ~ 3 Hz (Marhl et al., 2022; Seymour et al., 2022). This contrasts with canonical SQUID-MEG set-ups and some OPM-MEG designs (Alem et al., 2023), where the SQUID dewar or the OPM helmet are rigidly attached to the ground, and thus could only be affected when the ground moves relative to the Earth's magnetic field (e.g., when someone is stomping on the ground close to the MEG scanner). Several previous studies would therefore apply a high-pass filter with a rather high cutoff of e.g., 4 Hz in order to ensure removal of low-frequency artifacts (Gialopsou et al., 2021; Jodko-Władzińska & Sander, 2024; Marhl et al., 2022; Seymour et al., 2021, 2022). While such a high high-pass filter (HPF) may not cause a problem if the goal is to examine the power of oscillatory bands above this filter, this practice may pose a problem for studying event-related fields (ERFs), as high-pass filters with a high cutoff are known to attenuate or eliminate ERP (event-related potential) amplitudes and ERP effects across conditions in EEG studies (Tanner et al., 2015; van Driel et al., 2021; G. Zhang et al., 2024a, 2024b). Moreover, since

visual sentence processing studies usually present each segment (usually a word) for several hundreds of milliseconds, many ERF components may form some regularities in a lower frequency range slightly above 1 Hz (N. Ding et al., 2015; Y. Lu et al., 2022), therefore a high-pass filter of 4 Hz would especially risk removing them (Seymour et al., 2022). On the other hand, some previous studies employed a moderate high-pass filter of e.g., 1 Hz, along with canonical preprocessing methods for SQUID-MEG data including ICA (Iivanainen et al., 2023). However, ICA may not be able remove the movement artifacts if they are statistically non-stationary. The fundamental challenge is that, while ICA procedures typically reject undesirable components related to e.g., eye blinks and heart beats through inspection of waveform and topographic field map (Puce & Hämäläinen, 2017), it is unclear what would be the canonical topography for low-frequency movement artifacts in OPM-MEG. Moreover, with experimental OPM-MEG setups with a handful of (e.g., ~ 10) sensors, it is challenging to establish full topographic field maps in the first place (Iivanainen et al., 2023; Sander et al., 2020). Furthermore, the ability for ICA to single out “clean” artifact components appears to degrade as the number of sensors decrease (Klug & Gramann, 2021). However, the number of OPM sensors available is usually constrained by the nontrivial price of OPM sensors, so that many labs starting on OPM-MEG only possess a limited number of sensors. Therefore, a reasonable pipeline for OPM-MEG needs to be developed to analyze ERFs with the goal of (1) attenuating low-frequency movement artifacts, (2) retaining major ERF deflections and ERF effects, and (3) accommodating OPM-MEG setups with a smaller set of sensors (for which inspecting the topographic field map is challenging). Developing methods for setups with small numbers of OPMs will lead to more cost-effective research tools, as OPM sensors are still very expensive.

Such pipelines would greatly unleash the potential of OPM-MEG to study ERFs, including in sentence processing experiments.

One existing denoising method, evoked-biased denoising source separation (evoked-biased DSS (de Cheveigné & Simon, 2008)), may resolve this challenge, since this method does not rely on visual inspection of topographic field maps. Evoked-biased DSS decomposes data from sensors into PCA (principal component analysis) components, ranking from more time-locked to less time-locked. Assuming that movements are generally not time-locked to the stimulus, by keeping the first few components and excluding the rest of the components, we do not risk removing evoked responses much, while attenuating the effects from movement artifacts. However, evoked-biased DSS has not been tested on OPM-MEG data.

In the current OPM-MEG study, with a canonical visual sentence processing paradigm in Mandarin Chinese, we compared across preprocessing pipelines with various high-pass filtering thresholds and with/without evoked-biased DSS, leveraging the well-studied N400 response (and N400 effects across conditions) observed in SQUID-MEG and EEG (Lau et al., 2008, 2016). In SQUID-MEG, N400 manifests as a deflection around 350 ms which is particularly observable in the sensors above the left hemisphere (Pylkkänen et al., 2002, 2006; Stockall et al., 2004). Apart from the N400 deflection per se, we also capitalize on the difference in N400 amplitude between verbs and nouns, which has been observed in SQUID-MEG and EEG systems as well, across a variety of languages including Mandarin Chinese (our stimuli language; (Federmeier et al., 2000; Khader et al., 2003; Tsigka et al., 2014; Xia et al., 2016; Xia & Peng, 2022)). To date, few sentence processing experiments have been conducted with OPM-MEG. Although one recent study employed an auditory sentence processing paradigm (Wu et al., 2024), they primarily

employed multivariate analysis methods, and so it was unclear whether the canonical univariate N400 deflection and N400 effects were preserved.

To preview, we observed that while a 4 Hz high-pass filter completely destroyed the N400 deflection and any N400 differences between verbs and nouns, a low (0.1 Hz) or moderate (1 Hz) high-pass filter alone retained too much low-frequency artifacts in ERF data. Combining a moderate high-pass filter (1 Hz) with evoked-biased DSS appeared to offer the best performance in attenuating low-frequency movement artifacts while retaining the N400 deflection and the N400 difference between verbs and nouns.

7.2. Experiment 7.1: Method

7.2.1. Participants

11 native speakers of Mandarin Chinese participated in this study (5 females, age $M = 25$, $SD = 3$). Informed consent was obtained from all participants and they received monetary reimbursement for their participation. The procedures were approved by the IRB of University of Maryland.

7.2.2. Stimuli and procedure

The experiment was modeled after the EEG experiment of (Yu, Mancha, et al., 2024), with more sentences and therefore more trials.

Stimuli. We first came up with two lists of 80 verbs and 80 nouns in Mandarin Chinese, which was a superset of the list used in Yu, Mancha, et al. (2024). The verbs consisted of 40 separable verbs and 40 inseparable verbs; the nouns consisted of 40 compound nouns and 40 simplex nouns. For their detailed syntactic properties, see Yu, Mancha, et al. (2024). Comparing the responses between the two types of verbs or the two types of nouns is beyond the purpose of the current paper. These verbs and nouns were all disyllabic, each corresponding to two Chinese characters (average number of strokes for each verb or noun item: verb: 16.2; noun: 15.3). Mean log frequency in the BCC corpus (Xun et al., 2016): verb: 3.9, noun: 3.5. Mean log frequency in the CCL corpus (Zhan et al., 2019): verb: 4.1, noun: 3.7.

Then we created 160 sentences based on the 160 verb and noun items. The verbs and nouns were embedded in sentences in a certain sentential structure respectively, demonstrated in Table 7.1. Two counterbalanced lists of sentences for the verbs were created with two counterbalanced list of people’s names. These two lists were administered to roughly half of the participants respectively. Another 60 filler sentences were adapted from another study (Liao & Lau, 2020), rendering in total 220 sentences for each participant.

Table 7.1 Example sentential stimuli for Experiment 7.1.
The critical segment (i.e., the critical word) for our analysis is marked in bold.

Nouns
<i>Sentence structure:</i> [因为] [compound/simplex noun] [是] [segment 4] [segment 5,] [所以/因此] [continuation.] <i>English translation:</i> [Because] [compound/simplex noun] [is] [segment 4] [segment 5,] [therefore/so] [continuation.]
Example: 因为/ 热狗 /是/一种/高热量的/食物, /所以/不是很健康。 Because/ hotdog /is/a.kind.of/high-calorie/food,/therefore/not.very.healthy.

“Because hotdog is a kind of high-calorie food, therefore it is not very healthy.”

Verbs

Sentence structure:

[Name] [想要/不想/很想] [**separable/inseparable verb,**] [所以/因此] [continuation.]

English translation:

[Name] [wants to/doesn't want to/really wants to] [**separable/inseparable verb,**]
[so/therefore] [continuation.]

Example:

小范/很想/造反, /所以/他秘密组织了一支军队。

Fan[name]/really.wants.to/**rebel**,/therefore/he.gathered.an.army.secretly.

“Fan really wants to rebel, therefore he gathered an army secretly.”

Procedure. The procedure was also generally consistent with Yu, Mancha, et al. (2024).

Participants sat ~ 90 cm in front of the screen. Due to the setup of the projector in the lab, the screen had to be ~ 30° tilted from the ground; the participants' high performance suggested that the tilted screen did not affect sentence comprehension. 220 trials were run for each participant, with a different sentence presented in each trial. Each trial began with a fixation cross of 1000 ms with a 200 ms blank screen, followed by presentation of the segments of this sentence in a segment-by-segment (i.e., RSVP, rapid serial visual presentation) manner, with a presentation duration for 600 ms for each segment, and a 200 ms interval between segments, for a total of 800 ms stimulus-onset asynchrony (Figure 7.1). The segments were presented in white characters against a black background, in ~ 57 pt, Yahei UI font, at the center of the screen. In ~ 1/3 of the trials (76 out of 220 trials), the last segment of the sentence (i.e., the continuation) was colored in green and the subjects were instructed to respond to whether this continuation was congruent in this sentence, by pressing the left (congruent) or the right button (incongruent) on a response pad (the responses did not have a time-limit). In the other ~ 2/3 of the trials, the last segment was presented on the screen in white briefly for 1000 ms, and subjects did not need to respond. The

inter-trial intervals were self-paced; subjects pressed any button when they were ready to start the next trial. The participants completed a SQUID-MEG session with similar materials prior to the OPM-MEG visit. Since the two visits were interleaved by on average 10.5 (range 5-18) days, the SQUID-MEG session would unlikely affect the OPM-MEG session.

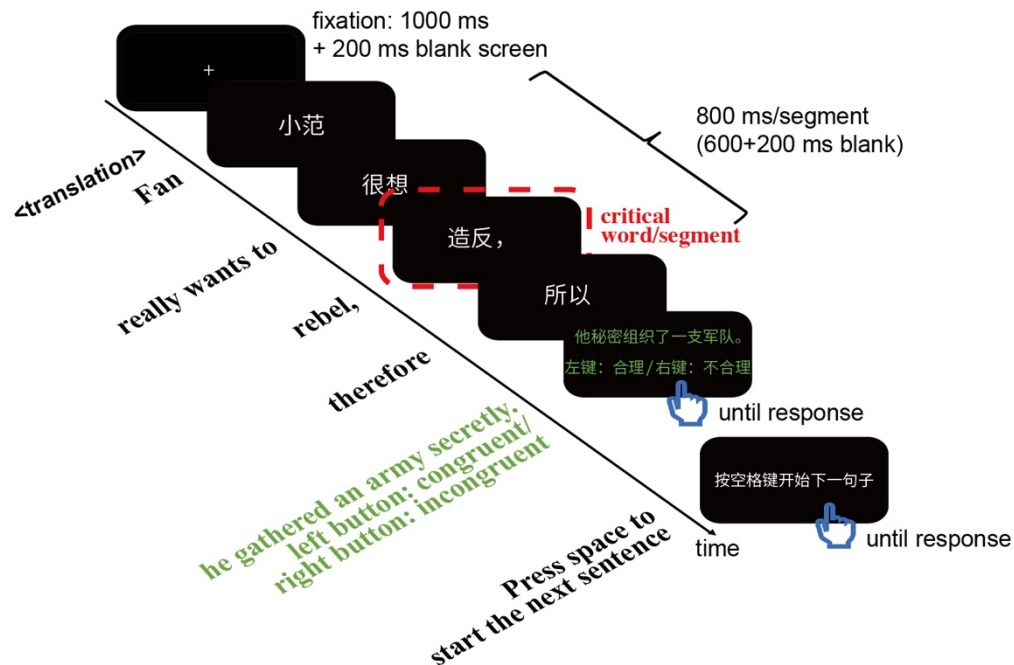


Figure 7.1 Illustration of a trial of our sentence processing study (Experiment 7.1). The English translation of the stimuli presented in Chinese is below the arrow.

7.2.3. Behavioral data analysis

Behavioral accuracy for each participant was calculated for the trials with a continuation judgement task, as the proportion of correct responses.

7.2.4. OPM Setup

The OPM measurements were conducted in a two-layered Ak3b magnetically shielded room (Vacuumschmelze GmbH & Co. KG, Hanau, Germany). Nine OPM magnetometer sensors (QZFM generation 2, QuSpin Inc., Louisville, Colorado, USA) were positioned across the scalp using a modified EEG cap (easycap 21-channel 10-20 cap). The cap had clip in OPM holders attached to the fabric (similar to the approach in <https://quspin.com/experimental-meg-cap/>). Such fabric caps modified for OPM usage have been used for clinical works as well (Feys et al., 2022). OPMs were placed especially on the left hemisphere, at T8, F4, Fpz, F3, F7, C3, T7, P7, O1 and O2 (Figure 7.2). Because the easycap 21-channel cap doesn't have an Fpz location, we created one at the appropriate location between Fp1 and Fp2. Each sensor recorded from two axes (z and y: radial head field and tangential field). The direction of the y axis was not consistent across all subjects due to rotational uncertainties, but was constant within each subject. In the current paper, we only analyze the z axis (radial direction relative to the head), which is in line with the direction recorded in most existing SQUID-MEG devices. Although SQUID devices mainly consist of gradiometers, these gradiometers are designed to be sensitive to radial magnetic field information.

7.2.5. OPM data analysis

OPM data analyses was conducted with customized code in Python and MATLAB, with MNE-Python (Gramfort et al., 2013). Statistical tests (paired t-test and repeated-measures ANOVA) were administered in JASP 0.18.1 (JASP Team, 2023).

Preprocessing. Upon inspecting the power spectrum (Figure 7.2) for individual participants, there were generally a 50 Hz component as well as a ~ 27 Hz component. The 50 Hz component was most likely line noise. A 20-30 Hz component is generally present in other OPM systems as well (Alem et al., 2023; Seymour et al., 2022; X. Zhang et al., 2020) with a yet-unclear source. Therefore, we applied notch filters at 50 and 27 Hz (transitional bandwidth = 4 Hz) to remove these components. Then we tested six different treatments: (1) Data was passed with a 0.1-40 Hz IIR (infinite impulse response) filter (i.e., high-pass filter, HPF = 0.1 Hz) only, (2) Data was passed with a 0.1-40 Hz IIR filter, then an evoked-biased denoising source separation (evoked-biased DSS), (3) a 1-40 Hz IIR filter only, (4) a 1-40 Hz IIR filter followed by evoked-biased DSS, (5) a 4-40 Hz IIR filter only, (6) a 4-40 Hz IIR filter followed by evoked-biased DSS. The other parameters for the IIR filters were determined by default algorithms in MNE-Python.

The procedure of evoked-biased DSS (de Cheveigné & Simon, 2008) is as follows. Evoked-biased DSS transforms normalized data from 9 sensors into 9 components, ranking from more stimulus-evoked-like to more noise-like. By keeping the first few components and excluding the rest of the components, we do not risk removing evoked responses much, while attenuating the effects from movement artifacts. The final epoch size that we wanted to have is -200:800 ms, time-locked to critical word onset. Each epoch was baselined based on the -200:0 ms baseline. In order to reduce the influence of the baseline period in DSS (note: in the current paper, “DSS” refers to evoked-biased DSS), we ran DSS on longer epochs of -200:1600 ms (S. G. Kim et al., 2020, 2022). Evoked-biased DSS was conducted with all epochs across the four conditions (epoch data were normalized for each sensor prior to DSS), resulting in 9 DSS components. We projected the first two DSS components back to the sensor space to create corrected epochs for analysis.

ERFs (event-related fields) were computed as the median of the epochs for each participant and for each condition respectively: separable verb, inseparable verb, compound noun, simplex noun, aggregated verb condition across separable and inseparable verbs, as well as aggregated noun condition across compound and simplex nouns. Because the mean is more subject to the influence of outlier data points, computing mean ERFs would require us to apply an additional outlier cutoff processing step which would likely result in different numbers of excluded trials across preprocessing pipelines. Median ERFs are less sensitive to outliers, therefore computing the ERFs with the median mitigates the need for applying a cutoff.

Calculation of cross-trial variance for ERFs. As a means of quantifying overall noise level in the results of each pipeline, we computed a cross-trial variance measure for the ERFs. For each pipeline, for each participant and at each time point, we first calculated *one* variance value across all verb epochs and all noun epochs respectively. This variance value was calculated in the following way. First, we calculated the standard deviation at each sensor. Then, we averaged the standard deviations across all 9 sensors. After that, for each time point, we averaged the standard deviations across all participants, for the noun epochs and verb epochs respectively. This way, for noun epochs and verb epochs respectively, we derived an 1D array across time, with one averaged variance value at each time point.

Calculation of RMS (root mean square). To determine whether the pipelines adequately captured the typical peaks in the ERF, we used an RMS measure. We calculated the RMS from ERFs that were first averaged across the verb and noun conditions, then averaged across participants. RMS was calculated as the root mean square across all 9 sensors.

Calculation of N400 amplitude. As described below, ERFs collapsed across conditions indicated a clear N400 peak between 300-400ms that was maximal at sensor F7. Therefore, we calculated N400 amplitude for each condition and for each participant at this sensor as the mean response amplitude across the 300-400 ms time window upon segment onset.

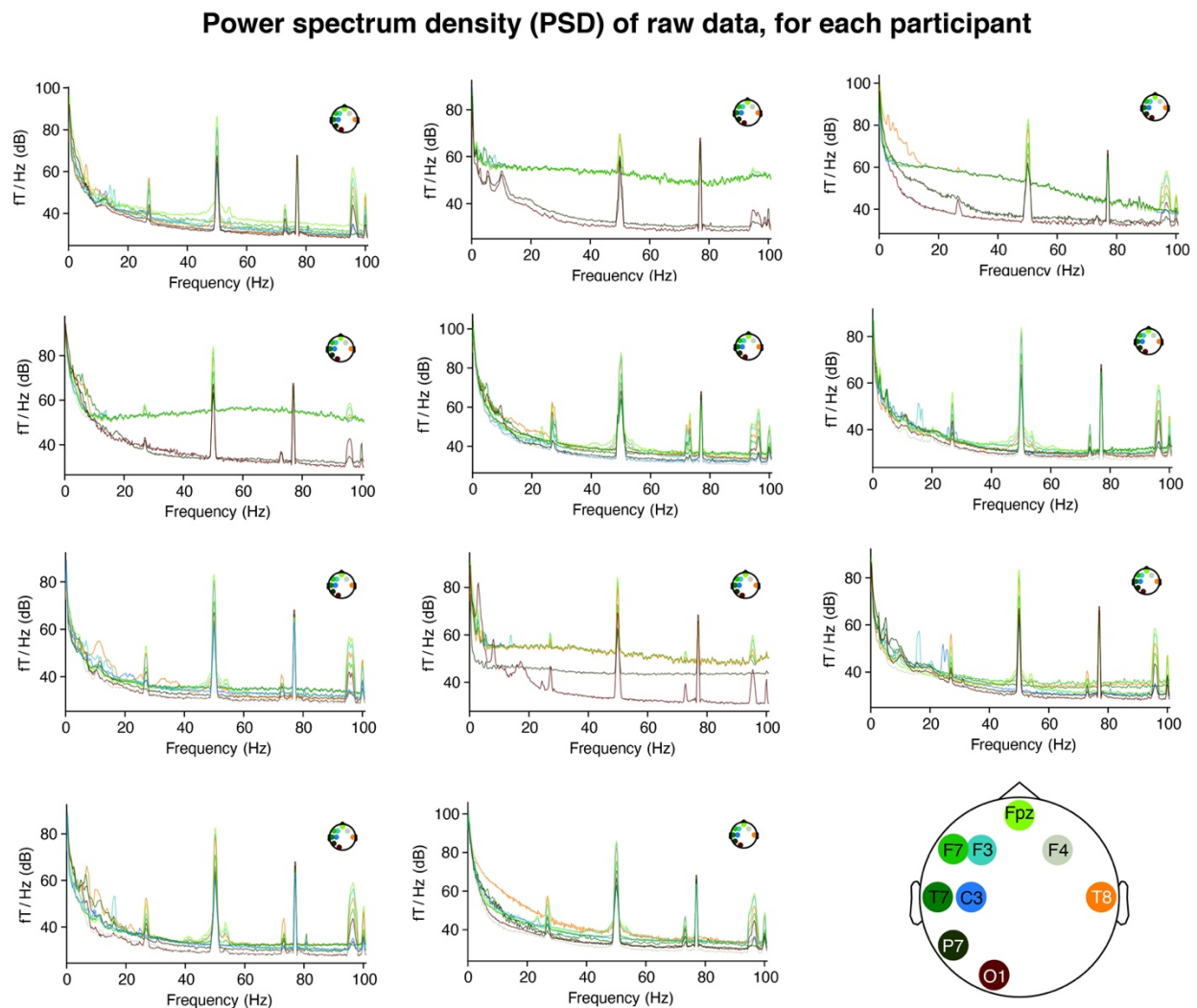


Figure 7.2 Power spectrum density (PSD) of raw OPM-MEG data for each of the 11 participants.

The OPM sensors were attached to an EEG cap, whose layout is illustrated in the bottom right corner. Consistent with prior OPM-MEG studies, we observed two spectral peaks around 27 Hz (cause yet unknown) and 50 Hz (line noise) across most participants.

7.3. Results

7.3.1. Behavioral results

Accuracy. Participants generally achieved ceiling accuracy (range 99%-100%) in judging the semantic congruity of sentence continuations. Therefore, all trials were included in the MEG analysis.

7.3.2. MEG results

In this section we evaluate the processing pipelines based on several qualitative and quantitative aspects.

Shape of ERF. In pipelines utilizing a 1 Hz high-pass filter (HPF), we were able to identify a clear deflection around 400 ms in left centro-frontal sensors, including F7, T7 and C3 (Figure 7.3). The deflection had the highest amplitude at F7, from which we calculate the amplitude of N400 responses for further analysis. This deflection was faintly identifiable when using a 0.1 Hz HPF, and was practically invisible when applying a 4 Hz HPF. The same observation generally holds when inspecting individual participants' data (Figure 7.4). Notably, the N400 deflection we observed was at the magnitude of hundreds of fT, which is higher compared to the typical magnitude observed in SQUID-MEG systems, at the magnitude of tens of fT (Lau et al., 2016; Lau & Nguyen, 2015), as expected given the much closer proximity of OPM-MEG sensors to the

underlying neural generators. Additionally, this difference in magnitude may also be relevant to the fact that, while OPM sensors are magnetometers, SQUIDs are often designed in a gradiometer configuration.

Cross-trial variance for ERFs. Across the entire epoch, the pipeline with HPF = 4 Hz, with DSS had the lowest cross-trial variance, which holds for both verb trials and noun trials (Figure 7.5). This was followed by the pipeline with HPF = 1 Hz, with DSS. The pipeline with HPF = 0.1 Hz without DSS had the highest variance among the pipelines.

Shape of RMS. Previous MEG sentence or single-word processing studies tend to observe a ~ 400 ms peak, along with earlier peak(s) around ~ 200 ms in RMS (Assadollahi & Pulvermüller, 2003; Maess et al., 2016; Tsigka et al., 2014). The ~ 200 ms peak(s) likely reflect visual-orthographic processing (Gwilliams et al., 2016; Salmelin, 2007). The ~ 400 ms peak or bump was clearly visible when we employed a 1 Hz HPF, while being faint or invisible for 0.1 or 4 Hz HPFs (Figure 7.6). A peak ~ 200 ms can only be clearly visually identified when a 1 Hz HPF is combined with DSS, but not when only applying a 1 Hz HPF.

N400 effect between verbs and nouns. Paired t-test (two-tailed) only revealed a significant difference between the N400 amplitudes of verbs and nouns for the pipeline with HPF = 1 Hz and DSS, $t(10) = -2.93, p = 0.015$, Cohen's $d = -0.89$ (Figure 7.7). This difference was not significantly different for the other pipelines, although the effects were generally in the same direction. HPF = 0.1 Hz without DSS: $t(10) = -0.36, p = 0.73$, Cohen's $d = -0.11$; HPF = 0.1 Hz with DSS: $t(10) = 1.48, p = 0.17$, Cohen's $d = 0.45$; HPF = 1 Hz without DSS: $t(10) = -1.67, p =$

0.13, Cohen's $d = -0.50$; HPF = 4 Hz without DSS: $t(10) = -1.27, p = 0.23$, Cohen's $d = -0.38$; HPF = 4 Hz with DSS: $t(10) = -0.83, p = 0.43$, Cohen's $d = -0.25$. Repeated-measures ANOVA across the four conditions (the two verb conditions and the two noun conditions) showed similar results.

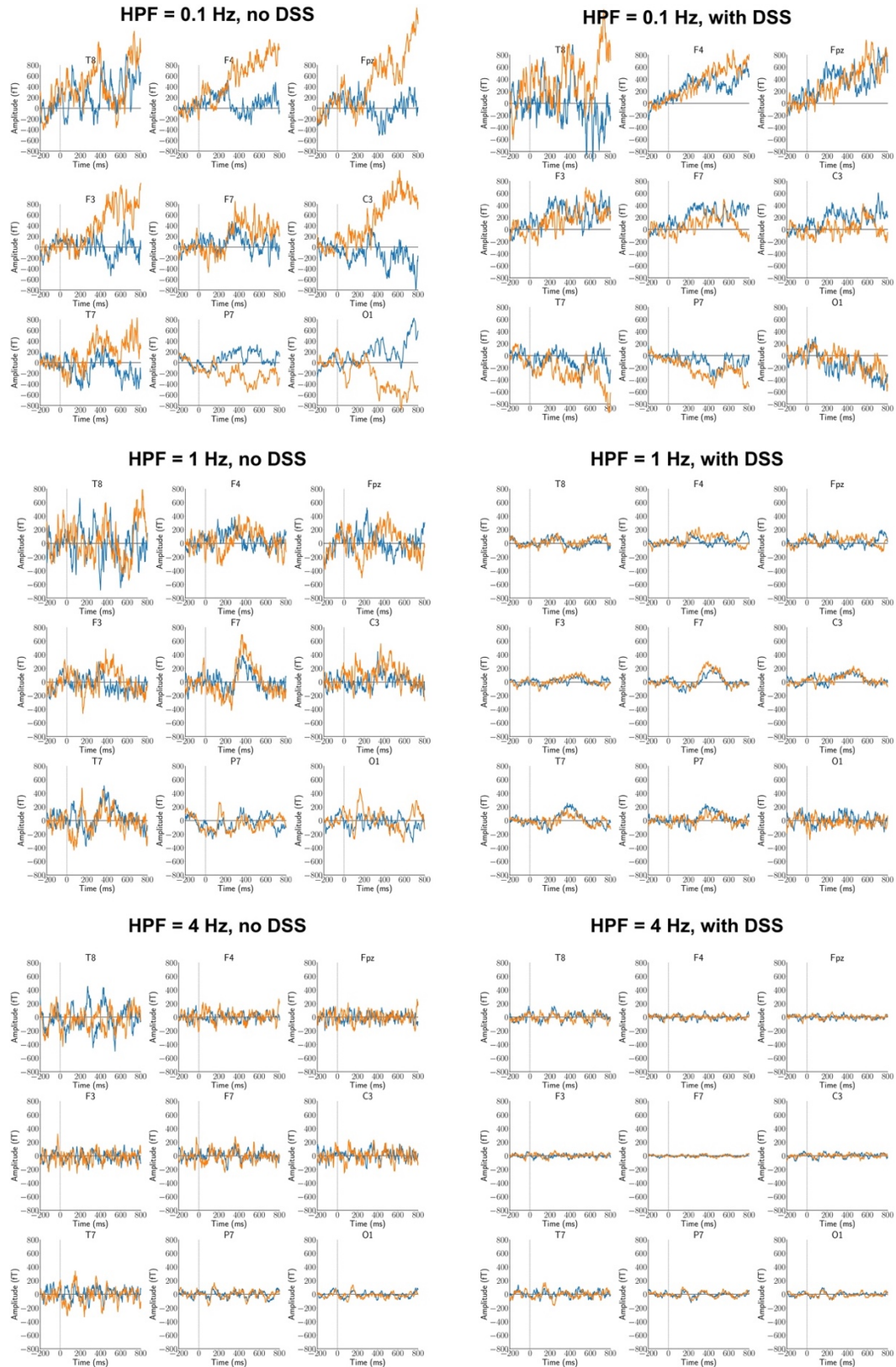


Figure 7.3 Grand average ERFs across all participants (blue line: verb; orange line: noun). ERF: event-related field.

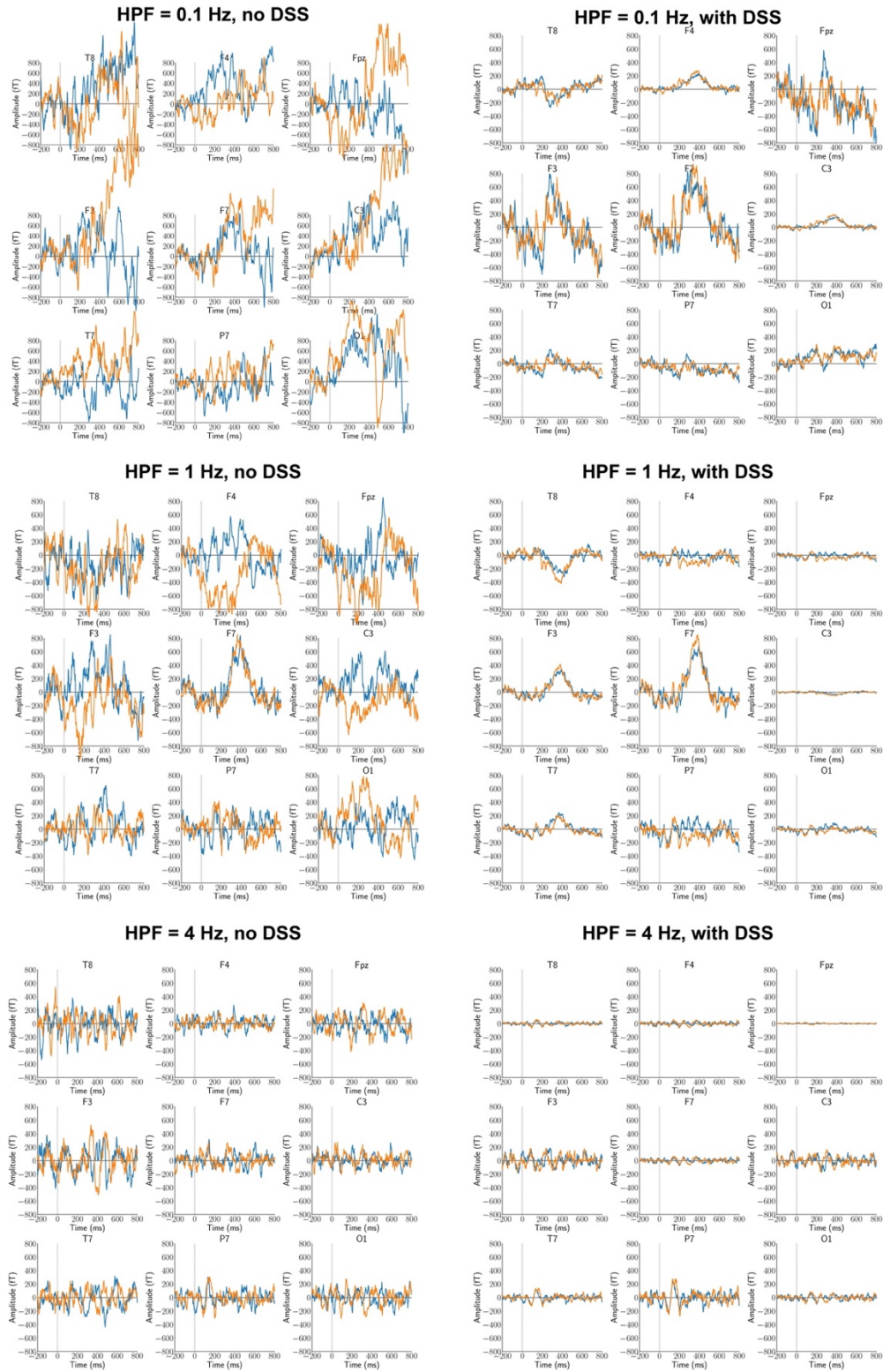


**Figure 7.4 Median ERFs for noun and verb epochs in one participant (blue line: verb; orange line: noun).
ERF: event-related field.**

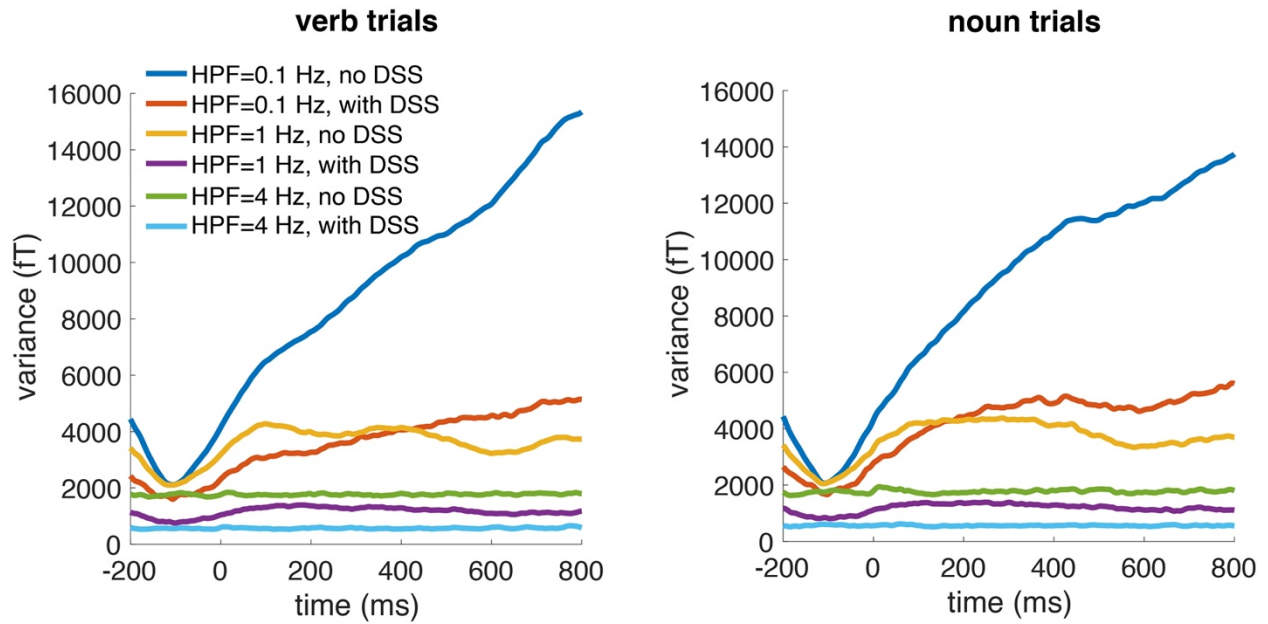


Figure 7.5 Cross-trial variance for each pipeline among verb trials (left) and noun trials (right), averaged across all participants.

HPF: high-pass filter; DSS: denoising source separation.

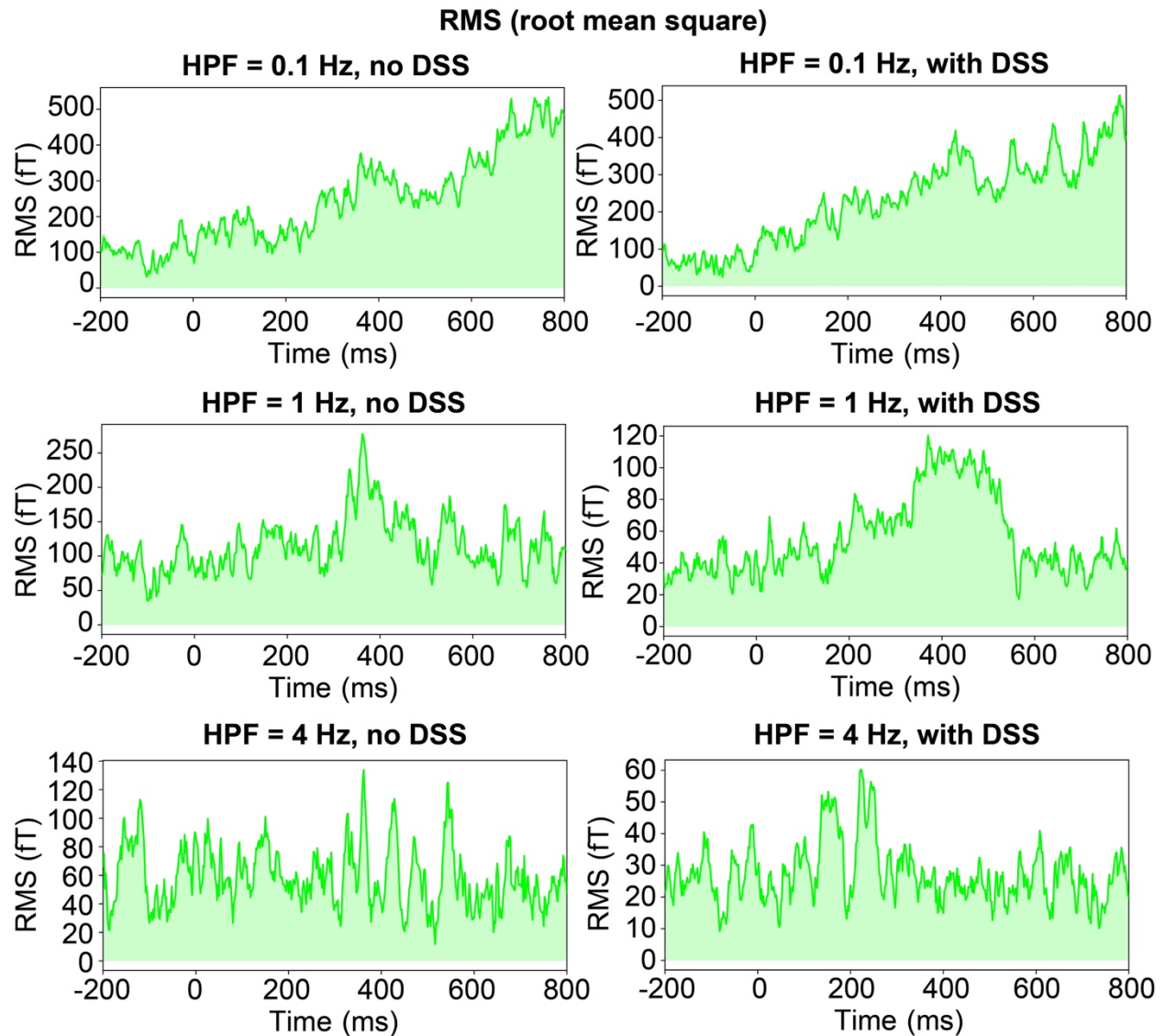


Figure 7.6 RMS (root mean square) across all electrodes, based on the ERFs averaged across verbs and nouns.

Only when combining a 1 Hz HPF and DSS are we able to clearly visually identify peaks around 200 ms and 400 ms. ERF: event-related field; HPF: high-pass filter; DSS: denoising source separation.

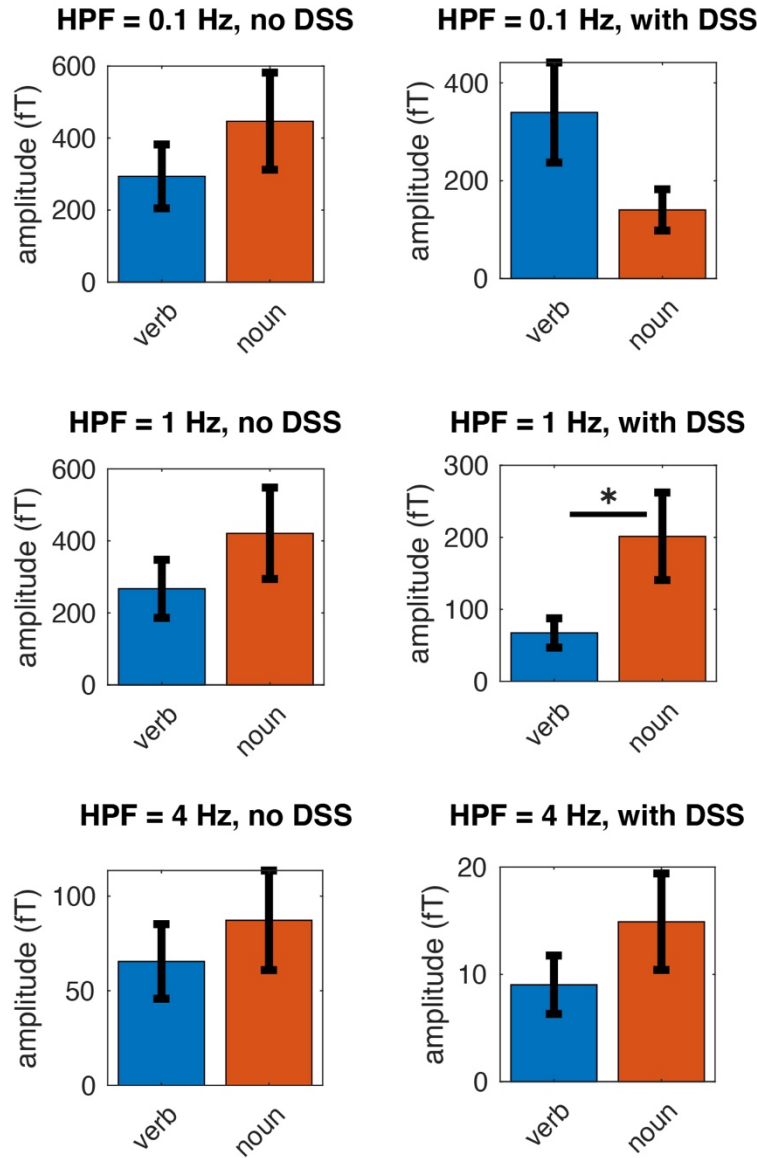


Figure 7.7 Comparison of N400 response (300-400 ms) amplitude at sensor F7 across verbs and nouns.

Only when an HPF (high-pass filter) of 1 Hz was combined with DSS, were we able to observe a significant difference across verbs and nouns. Error bars stand for standard error. *: $p < 0.05$.

HPF: high-pass filter; DSS: denoising source separation.

7.4. Conclusion

In the current paper, by comparing different preprocessing pipelines for OPM-MEG data, we demonstrated that combining a moderate high-pass filter (HPF) of 1 Hz with evoked-biased DSS

preserves the N400 deflection and allowed us to successfully replicate prior observations of an N400 difference between nouns and verbs. Moreover, this pipeline revealed canonical RMS peaks, and had a relatively low cross-trial variance among the pipelines. This pipeline thus unleashes the potential of OPM-MEG to study ERFs against low-frequency artifacts, especially for OPM-MEG setups with fewer sensors, where methods like ICA would be inapplicable. This can allow a beneficial ‘convergent methods’ approach for labs to obtain high-quality MEG data in closer proximity to the neural generators than traditional systems, even in the face of current constraints on the number of OPM sensors.

7.5. Future directions

7.5.1. On applying higher high-pass filters

Although a 4 Hz HPF largely destroyed the N400 deflection and N400 effects in our study, such higher HPFs may still be applicable in studying some other ERFs. For example, the auditory evoked responses of M100, which is a peak ~ 100 ms upon auditory stimuli onset, appeared to survive such higher HPFs in previous studies (Marhl et al., 2022). Notably, in (Marhl et al., 2022), auditory pure tones were presented at a duration of 500 ms, with intervals of 1200 ms between the tones. Thus, the stimuli presentation would form regularities below the 4 Hz HPF. Nevertheless, M100 was observed in that study. This is not surprising since M100 is a transient response that only lasted for ~ 50 ms, compared to our N400 deflections, which appeared to extend for ~ 300 ms. Exactly when higher HPFs can/cannot be used to study which ERF

component remains to be further elucidated in future studies, through empirical investigations and simulations.

7.5.2. Other pipelines

Beyond our current pipeline, some other pipelines may also achieve satisfactory performances, and future studies could work on comparing across these pipelines. Of note, HFC (homogeneous field correction) may be a promising method for this purpose (Tierney et al., 2021). We did not perform HFC for our current data because it requires the sensor orientation angles, which we did not record in this study. Future studies could compare across our current pipeline and HFC for the metrics that we used in the current work, or even combine them together.

7.5.3. Exploring OPM-MEG recording with a limited number of sensors

In the current study, we demonstrated a pipeline that accommodates OPM-MEG systems with a limited number of sensors (in our case, nine sensors), and obtained meaningful ERF responses in a sentence processing study. In practice, the number of sensors may not only be constrained by financial costs, but also by preparation time. For example, attaching more OPM sensors to the cap increases setup time, which may be a limiting factor in certain clinical contexts where shorter preparation times are preferred. Notably, previous research has shown that as few as four dual-axis OPM sensors can be sufficient for a brain-computer interface for motor imagery decoding (Tan et al., 2025). This, along with our findings, suggests that even with a reduced sensor count, OPM-MEG can still yield meaningful results. Taken together, our study marks one of the initial

steps in validating OPM-MEG systems with fewer sensors, paving the way for more cost-effective solutions which would enable wider clinical usage of OPM-MEG.

8. Chapter 8

8.1. Conclusion

In the current dissertation, we took the initial steps towards a rather new and exciting research program centering on “pointers” (or indexicals) in moment-to-moment cognition. The central goal for the first phase of this research program is to uncover the extent to which a same set of pointers is shared across different tasks and source modalities. Of note, by “the same set”, we not only mean the same numerical capacity limit, but shared neural representations.

Through a series of behavioral and EEG experiments in Chapters 2 to 4, we showed initial evidence for a shared set of pointers that doesn’t care about the category of visually-derived objects (across biological and non-biological visual objects), doesn’t care about whether the objects are derived from vision or language, and perhaps even doesn’t care about whether what is being represented is an object or a relation between objects. Although certainly more future studies are needed, I believe I have demonstrated evidence against a very fragmented object indexing system in the mind, which is largely dependent on source modality or the ontological nature of objects or relations. To this day, object representation has largely been studied with respect to specific tasks: for example, the classic delay-match-to-sample visual working memory task, or the MOT task. However, it is not guaranteed that the representations tested are necessarily encapsulated in the visual system, despite being a “visual” task in name – that is, the source modality is vision. Instead, some representations may be shared across a variety of tasks and source modalities. For example, although object files and discourse referents have been proposed in vision science and language science respectively, it is unclear whether

they recruit truly distinct neural representations. My studies provide one of the first attempts to address questions in this research direction.

In Chapters 5 to 6, I demonstrated evidence for an MEG marker for pointers in visual working memory tasks. This neural marker manifests as a right posterior activity (RPA) at the sensors, and reflects sustained activities in the inferior IPS bilaterally. These studies, on the one hand, suggest that inferior IPS is hosting pointers in visual working memory tasks, and on the other hand, serve as the basis of testing whether these neural signatures hold for other tasks and source modalities. Chapter 7 marks methodological development in OPM-MEG, a new generation of MEG systems. As OPM-MEG offers signals with higher amplitude compared to SQUID-MEG systems, this paves the way to better localization of pointer-related signals in the future.

In all, as language and vision are often studied in silos due to separate research communities, we still know little about the nature of shared representations or computations across visual and linguistic processing. In the current dissertation, we asked a concrete research question of this kind: to what extent do linguistic and visual input share the same set of pointers in moment-to-moment processing, and more generally, to what extent do these pointers generalize across different input, tasks, and source modalities? The works presented in this dissertation provide initial theorization and evidence towards this research program, and I hope these works would inspire further investigations into this topic, as well as asking new questions across vision and language. For example, does shared neural computational principles account for aesthetic appreciation (Yu, Raghavan, et al., 2025) or stimulus memorization across visual and linguistic media? We are in a great time of proposing and answering these largely unasked questions, with the development of both neuroimaging methods and computational neuroscience.

I hope the current dissertation will inspire more future works to bridge language and vision in novel ways, generating new insights about shared and unique computations in the mind and brain.

8.2. Limitations

The results reported in Chapters 2 to 4 have limitations and invite follow-up studies. As behavioral results can be affected by both the representations of pointers and features, and the neural markers used therein (NSW and posterior alpha power) may not exclusively respond to pointer-based load, and the results may still be confounded with feature-based competition (in the behavioral study) or feature-based effects (for NSW). The study with posterior alpha power in Chapter 4 was relatively free of this concern as the number of features was controlled for, yet the spatial extent of attention also differed across object load conditions. The number of pointers needs to be disentangled from the spatial extent of attention in future studies (Jones, Thyer, et al., 2024).

One solution for future studies is to test these scenarios with MEG, and examine the more targeted neural markers for pointers (RPA and inferior IPS responses) that we developed in Chapters 5 and 6. If these neural markers are also immune to the spatial extent of attention and could reflect reassignability, these neural markers would serve as ideal hallmarks for pointers in moment-to-moment cognition.

That said, our current analysis on MEG data primarily reflects the scale of cohort of neurons, instead of single-neuron responses. A higher amplitude captured from a cohort of neurons may translate transparently into higher firing rates for each neuron, or less transparently

into more sophisticated synchronization patterns (Yu & Lau, 2023). While future studies with single-neuron recordings may offer important and complementary insights, we consider these “systems-level” analyses as offering unique insights as well: while a neuron within a certain neural circuit may encode unique information, downstream circuits may read out from cohort-level properties instead of from one or a handful of neurons (X. X. Lin et al., 2023).

For the studies on object relations reported in Chapter 3, we should note that if we think of the pointers for objects as deeply intertwined with intuitive physics (Balaban et al., 2024), it is unlikely a priori that these pointers should stand in for relations between objects or events as a whole, as relations or events per se seldom move in space. Therefore, even if relations or events could be represented by pointers, it remains an open question whether these pointers are qualitatively the same as the object pointers.

Another limitation is about the imagery experiment reported in Chapter 4: the task is still about “visual” imagery. While this study was already less “visual” compared to visual working memory studies as the imagery was not derived from visual input in recent history, further studies could examine whether this pattern generalizes to even more non-visual tasks, where visual imagery is not encouraged (e.g., processing sentences whose content are harder to visualize) or even consciously unavailable (e.g., aphantasia).

8.3. Discussion

8.3.1. Why do we need pointers?

Up to now, we have talked about pointers as a part of mental representations for objects. What are the functions of pointers? What may the pointers do, and what may they not do? *Why* do we need pointers? Suppose some cognitive-neural systems may be well-off using only less content-free and less reassignable variables, why would some systems use pointers that are more radically content-free and reassignable?

One system that appears to be using pointers is the perception-action (i.e., visuo-motor) circuit: pointers appear to be useful when we interact with objects around us with our motor system (Ballard et al., 1997). First, posterior parietal cortex -- a critical region hosting pointers (see Chapters 5 and 6) -- lies in the dorsal stream of visual processing, projecting to the motor cortex and deeply relevant to executing motor actions (Genovesio et al., 2014; Goodale & Milner, 1992). This is in contrast to the ventral pathway that is primarily responsible for feature detection. Second, motor reaching trajectories for visually-available stimuli have been found to be modulated by a similar capacity limit of ~ 4 objects (Gallivan et al., 2011), although the link between this capacity limit and capacity limits for more canonical visual working memory tasks remains to be established. Ecologically, the objects we reach for are often no longer relevant after a short while: for example, we reach for an apple and eat it up, or we reach for a pen and put it in the drawer after work -- done. In case we care about the pen again the next day, we would be retrieving the location of this pen from long-term memory instead of using short-term memory representations, given its observed stark capacity limit. Similarly, when we (and likely our distant ancestors) forage for berries, we tend to either eat them right away, or put them in a bag so that we can “forget about it” in short-term memory (and any important information is shifted to long-term formats). A radical content-free and reassignable, “burn-after-reading” representational format appears to be sufficient in this context. This is also in line with the

intuition that solving the Hanoi tower (Ballard et al., 1997) or playing the board game of *Geistertreppe* (English: Spooky Stairs) is rather effortful: we need to keep track of multiple objects and reach for them from time to time -- likely while everything is sustained in short-term memory. This is rather non-ecological in the sense that this is not what our perception-action system is “built for” (i.e., shaped by evolution). That is not to say that a less content-free and reassignable variable system cannot do this task, but that the content-free, reassignable pointer system that did evolve is sufficient. Physical reaching appears to share functions and neural bases (e.g., parietal regions) with directing overt/covert attention (Ballard et al., 1997; Culham & Valyear, 2006). Such commonality may also extend to describing scenes via language (which also translates into motor output in the end), in particular with demonstratives (e.g., this, that). Overall, based on the abovementioned rationale, pointers likely underlie computations in the visuo-motor pathway, which would be an ideal starting point to study *how* pointers are maneuvered in specific mental programs, as well as when pointers are recruited and when not.

8.3.2. Future directions

With the neural signatures for pointers developed here, we could ask further research questions. For example, how many pointers are taxed when we group two or more objects together? When two objects are grouped together, are we recruiting only one pointer because they are in the same group, two pointers because grouping do not affect the number of pointers recruited, or even three pointers since we now also need to represent the whole group apart from the two individual objects? Although much research has been conducted on grouping, little has been shown about how grouping specifically affects the pointers. A pioneering fMRI study suggested that the posterior parietal response level (see Table 1.1) for two visually-grouped items (grouped by a rectangle in the background) was in between those for one item and two ungrouped items (Y. Xu & Chun, 2007). This may suggest that the participants sometimes represent two grouped items with one pointer, and sometimes with

two, so that the average level of response was in between. This is in line with recent observations from Thyer and colleagues (2022), that four items grouped in two groups (two notched discs facing each other) would tend to be classified as two separate items rather than four separate items by the mvLoad classifier, as well as earlier studies utilizing CDA (Luria & Vogel, 2014; Peterson et al., 2015). Items in these studies were grouped by relatively low-level visual cues; however, we also group objects together based on a variety of information apart from visual backgrounds and boundaries. For instance, we may group objects together based on (1) long-term semantic knowledge, for example, we may group a match and a candle together because typically, a match lights up a candle according to our long-term knowledge (O'Donnell et al., 2018; Yu, Thakurdesai, et al., 2025), or (2) whether they are involved in the same (social) interaction, for example, we may group two individuals talking with each other, compared to they are standing there in isolation (Papeo, 2020; Yu, Li, et al., 2024; C. Zhou et al., 2019). Grouping may also be trained through associative learning (Ngiam, Thyer, et al., 2023). Furthermore, grouping may be biased via task goals or instructions (Balaban & Luria, 2016; Knowlton et al., 2023; Rabbitt et al., 2017), the current cognitive effort available (Xie & Zhang, 2023), or the number of items one need to memorize in total (Feigenson & Carey, 2005; Kutter et al., 2023). Much less is known about how these different sources or driving forces of grouping affects the number of pointers taxed. Interestingly, mental objects stored in long-term memory have also been proposed to be instantiated by pointers, leaving open also the intriguing question of how pointers in moment-to-moment processing interacts with long-term memory pointers (Cowan, 2019). Moreover, the neural signatures for pointers also allow us to characterize the development of pointer representations across the lifespan, and examine the relationship between these markers and individual traits or atypicalities.

How do we *compute with* these pointers within and across different cognitive-neural systems? In other words, what do the “computer programs” in different cognitive-neural systems look like? For example, swapping two mental objects (registered by variables/pointers x_1 and x_2) can be realized with only one additional variable x_3 :

Program 1

```
x3 = x1  
x1 = x2  
x2 = x3
```

Or by using two additional variables x_3 and x_4 :

Program 2

```
x3 = x1  
x4 = x2  
x1 = x4  
x2 = x3
```

In a recent paper, Tian and colleagues (2024) discovered that in macaque frontal cortex, two neural subspaces additional to the two subspaces for the two original objects are recruited when swapping the rank order for visual locations in a sequence. If a neural subspace meets both our criteria for neural signatures for pointers, one would be able to infer that the mental program (i.e., LoT procedures/routines) in macaque frontal cortex for swapping is Program 2. Thus, neural signatures for pointers in moment-to-moment cognition could serve as a window into even the actual LoT *procedures* that the brain is implementing, or at least narrow down the space for possible LoT procedures.

In all, we outlined a novel research program spanning the territories of vision, auditory, and language (neuro)science and beyond, calling for a quest for neural signatures for pointers in

moment-to-moment cognition, as well as for deciphering how these neural signatures generalize (or not) across paradigms, task goals and source modalities. In this way, we will be able to delineate the neural taxonomy of pointers in moment-to-moment cognition.

We are in an exciting time for this research question given unprecedented advances in human intracranial recording that could achieve unit-level resolution (Chung et al., 2022; Eisenkolb et al., 2023), as well as accompanying and increasing novel evidence demonstrating *digital* neural signatures in human cognition (e.g., coding by the rank order of neural firing sequences across neuronal populations; Xie et al., 2024). This research program could not only incubate novel evidence and perspectives towards the long-standing debate between symbolic and connectionist cognition in cognitive psychology/neuroscience (Fodor & Pylyshyn, 1988; Smolensky, 1988), but also towards the debate over many vs. one LoT(s) in the philosophy of mind (Fodor, 1975; Mandelbaum et al., 2022). These neural signatures may finally crack open the stiff window towards answering how our brain and cognition is programmed neurally, how the choice of mental programs may be determined by external and internal factors, and even towards actually writing and fixing these programs.

Bibliography

- 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
- Adam, K. C. S., Vogel, E. K., & Awh, E. (2020). Multivariate analysis reveals a generalizable human electrophysiological signature of working memory load. *Psychophysiology*, 57(12), 1–17. <https://doi.org/10.1111/psyp.13691>
- Akhlaghpour, H. (2022). An RNA-based theory of natural universal computation. *Journal of Theoretical Biology*, 537, 110984. <https://doi.org/10.1016/j.jtbi.2021.110984>
- Alem, O., Hughes, K. J., Buard, I., Cheung, T. P., Maydew, T., Griesshammer, A., Holloway, K., Park, A., Lechuga, V., Coolidge, C., Gerginov, M., Quigg, E., Seames, A., Kronberg, E., Teale, P., & Knappe, S. (2023). An integrated full-head OPM-MEG system based on 128 zero-field sensors. *Frontiers in Neuroscience*, 17(June), 1–8. <https://doi.org/10.3389/fnins.2023.1190310>
- Altmann, G. T. M., & Ekves, Z. (2019). Events as intersecting object histories: A new theory of event representation. *Psychological Review*, 126(6), 817–840. <https://doi.org/10.1037/rev0000154>
- Altmann, L. J. P., Saleem, A., Kendall, D., Heilman, K. M., & Rothi, L. J. G. (2006). Orthographic directionality and thematic role illustration in English and Arabic. *Brain and Language*, 97(3), 306–316. <https://doi.org/10.1016/j.bandl.2005.12.003>
- Alunni-Menichini, K., Guimond, S., Bermudez, P., Nolden, S., Lefebvre, C., & Jolicoeur, P. (2014). Saturation of auditory short-term memory causes a plateau in the sustained anterior negativity event-related potential. *Brain Research*, 1592, 55–64. <https://doi.org/10.1016/j.brainres.2014.09.047>
- Andersen, L. M., Pfeiffer, C., Ruffieux, S., Riaz, B., Winkler, D., Schneiderman, J. F., & Lundqvist, D. (2020). On-scalp MEG SQUIDS are sensitive to early somatosensory activity unseen by conventional MEG. *NeuroImage*, 221(July). <https://doi.org/10.1016/j.neuroimage.2020.117157>
- Arend, A. M., & Zimmer, H. D. (2011). What does ipsilateral delay activity reflect? Inferences from slow potentials in a lateralized visual working memory task. *Journal of Cognitive Neuroscience*, 23(12), 4048–4056. https://doi.org/10.1162/jocn_a_00068
- Assadollahi, R., & Pulvermüller, F. (2003). Early influences of word length and frequency: A group study using MEG. *NeuroReport*, 14(8), 1183–1187. <https://doi.org/10.1097/00001756-200306110-00016>

- 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>
- Baillargeon, R., & Carey, S. (2012). Core cognition and beyond: The acquisition of physical and numerical knowledge. In S. M. Pauen (Ed.), *Early childhood development and later outcome* (pp. 33–65). Cambridge University Press.
- Balaban, H., Drew, T., & Luria, R. (2018). Delineating resetting and updating in visual working memory based on the object-to-representation correspondence. *Neuropsychologia*, 113(March), 85–94. <https://doi.org/10.1016/j.neuropsychologia.2018.03.038>
- 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., Drew, T., & Luria, R. (2023). Dissociable online integration processes in visual working memory. *Cerebral Cortex*, 33(23), 11420–11430. <https://doi.org/10.1093/cercor/bhad378>
- Balaban, H., & Luria, R. (2016). Object representations in visual working memory change according to the task context. *Cortex*, 81, 1–13. <https://doi.org/10.1016/j.cortex.2016.04.004>
- 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., 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. (2024). The capacity limits of mental simulation. *PsyArXiv*. <https://doi.org/https://doi.org/10.31234/osf.io/xzcmb>
- Baldassano, C., Chen, J., Zadbood, A., Pillow, J. W., Hasson, U., & Norman, K. A. (2017). Discovering Event Structure in Continuous Narrative Perception and Memory. *Neuron*, 95(3), 709–721.e5. <https://doi.org/10.1016/j.neuron.2017.06.041>
- Balestrieri, E., Ronconi, L., & Melcher, D. (2021). Shared resources between visual attention and visual working memory are allocated through rhythmic sampling. *European Journal of Neuroscience*, April, 1–14. <https://doi.org/10.1111/ejn.15264>
- Ballard, D. H., Hayhoe, M. M., Pook, P. K., & Rao, R. P. N. (1997). Deictic codes for the

- embodiment of cognition. *Behavioral and Brain Sciences*, 20(4), 723–767.
<https://doi.org/10.1017/S0140525X97001611>
- Bapat, A. N., Shafer-Skelton, A., Kupitz, C. N., & Golomb, J. D. (2017). Binding object features to locations: Does the “spatial congruency bias” update with object movement? *Attention, Perception, & Psychophysics*, 79(6), 1682–1694. <https://doi.org/10.3758/s13414-017-1350-5>
- Barrett, A. M., Kim, M., Crucian, G. P., & Heilman, K. M. (2002). Spatial bias: Effects of early reading direction on Korean subjects. *Neuropsychologia*, 40(7), 1003–1012.
[https://doi.org/10.1016/S0028-3932\(01\)00147-6](https://doi.org/10.1016/S0028-3932(01)00147-6)
- Bays, P. M., Schneegans, S., Ma, W. J., & Brady, T. F. (2024). Representation and computation in visual working memory. *Nature Human Behaviour*, 8(6), 1016–1034.
<https://doi.org/10.1038/s41562-024-01871-2>
- Becke, A., Müller, N., Vellage, A., Schoenfeld, M. A., & Hopf, J. M. (2015). Neural sources of visual working memory maintenance in human parietal and ventral extrastriate visual cortex. *NeuroImage*, 110, 78–86. <https://doi.org/10.1016/j.neuroimage.2015.01.059>
- 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 and Review*, 23(1), 265–270. <https://doi.org/10.3758/s13423-015-0891-7>
- Berggren, N., & Eimer, M. (2016). Does Contralateral Delay Activity Reflect Working Memory Storage or the Current Focus of Spatial Attention within Visual Working Memory? *Journal of Cognitive Neuroscience*, 28(12), 2003–2020. https://doi.org/10.1162/jocn_a_01019
- Bettencourt, K. C., & Xu, Y. (2016). Understanding location- and feature-based processing along the human intraparietal sulcus. *Journal of Neurophysiology*, 116(3), 1488–1497.
<https://doi.org/10.1152/jn.00404.2016>
- Beukers, A., Buschman, T., Cohen, J. D., & Norman, K. A. (2020). Is activity silent working memory just episodic memory? *Trends in Cognitive Sciences*, 1–10.
<https://doi.org/10.1016/j.tics.2021.01.003>
- Bidelman, G. M., Brown, J. A., & Bashivan, P. (2021). Auditory cortex supports verbal working memory capacity. *NeuroReport*, 163–168.
<https://doi.org/10.1097/WNR.0000000000001570>
- Block, N. (2023). Perception is iconic, perceptual working memory is discursive. *Behavioral and Brain Sciences*, 46, e265. <https://doi.org/10.1017/S0140525X23001899>

- Boersma, P. (2001). Praat, a system for doing phonetics by computer. *Glott International*, 5(9/10), 341–345.
- Bonatti, L., Frot, E., Zangl, R., & Mehler, J. (2002). The human first hypothesis: Identification of conspecifics and individuation of objects in the young infant. *Cognitive Psychology*, 44(4), 388–426. <https://doi.org/10.1006/cogp.2002.0779>
- Borna, A., Carter, T. R., Goldberg, J. D., Colombo, A. P., Jau, Y. Y., Berry, C., McKay, J., Stephen, J., Weisend, M., & Schwindt, P. D. D. (2017). A 20-channel magnetoencephalography system based on optically pumped magnetometers. *Physics in Medicine and Biology*, 62(23), 8909–8923. <https://doi.org/10.1088/1361-6560/aa93d1>
- Boto, E., Meyer, S. S., Shah, V., Alem, O., Knappe, S., Kruger, P., Fromhold, T. M., Lim, M., Glover, P. M., Morris, P. G., Bowtell, R., Barnes, G. R., & Brookes, M. J. (2017). A new generation of magnetoencephalography: Room temperature measurements using optically-pumped magnetometers. *NeuroImage*, 149(January), 404–414. <https://doi.org/10.1016/j.neuroimage.2017.01.034>
- 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>
- Bowman, H., & Wyble, B. (2007). The simultaneous type, serial token model of temporal attention and working memory. *Psychological Review*, 114(1), 38–70. <https://doi.org/10.1037/0033-295X.114.1.38>
- Brady, T. F., & 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. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 41(3), 921–929. <https://doi.org/10.1037/xlm0000075>
- Brady, T. F., Konkle, T., Alvarez, G. A., & Oliva, A. (2008). Visual long-term memory has a massive storage capacity for object details. *Proceedings of the National Academy of Sciences of the United States of America*, 105(38), 14325–14329. <https://doi.org/10.1073/pnas.0803390105>
- Brady, T. F., Störmer, V. S., & Alvarez, G. A. (2016). Working memory is not fixed-capacity: More active storage capacity for real-world objects than for simple stimuli. *Proceedings of the National Academy of Sciences of the United States of America*, 113(27), 7459–7464. <https://doi.org/10.1073/pnas.1520027113>
- Brainard, D. H. (1997). The Psychophysics Toolbox. *Spatial Vision*, 10(4), 433–436. <https://doi.org/10.1163/156856897X00357>
- Brickwedde, M., Anders, P., Kühn, A. A., Lofredi, R., Holtkamp, M., Kaindl, A. M., Grent-‘t-Jong, T., Krüger, P., Sander, T., & Uhlhaas, P. J. (2024). Applications of OPM-MEG for

- translational neuroscience: a perspective. *Translational Psychiatry*, 14(1), 1–12.
<https://doi.org/10.1038/s41398-024-03047-y>
- Brodbeck, C., Gwilliams, L., & Pylkkänen, 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. (2025). Discourse referents in infancy. *Psychological Review*.
<https://doi.org/10.1037/rev0000545>
- Broers, N., & Busch, N. A. (2021). Correction to: The effect of intrinsic image memorability on recollection and familiarity. *Memory & Cognition*, 49(5), 1019–1019.
<https://doi.org/10.3758/s13421-021-01158-1>
- Brookes, M. J., Leggett, J., Rea, M., Hill, R. M., Holmes, N., Boto, E., & Bowtell, R. (2022). Magnetoencephalography with optically pumped magnetometers (OPM-MEG): the next generation of functional neuroimaging. *Trends in Neurosciences*, 45(8), 621–634.
<https://doi.org/10.1016/j.tins.2022.05.008>
- 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)
- Bueti, D., & Walsh, V. (2009). The parietal cortex and the representation of time, space, number and other magnitudes. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 364(1525), 1831–1840. <https://doi.org/10.1098/rstb.2009.0028>
- Cai, Y., Fulvio, J. M., Samaha, J., & 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(6), 1–15. <https://doi.org/10.1523/ENEURO.0207-22.2022>
- Cao, R., & Deouell, L. Y. (2023). Task-Dependent Representation of Feature Binding within the Same Dimension in Visual Working Memory. *BioRxiv*.
<https://doi.org/10.1101/2023.12.19.572444>
- Capitani, E., Laiacona, M., Mahon, B., & Caramazza, A. (2003). WHAT ARE THE FACTS OF SEMANTIC CATEGORY-SPECIFIC DEFICITS? A CRITICAL REVIEW OF THE CLINICAL EVIDENCE. *Cognitive Neuropsychology*, 20(3–6), 213–261.
<https://doi.org/10.1080/02643290244000266>
- Carey, S. (2009). *The origin of concepts*. Oxford University Press.
- Carey, S., & Xu, F. (2001). Infants’ knowledge of objects: beyond object files and object tracking. *Cognition*, 80(1–2), 179–213. [https://doi.org/10.1016/S0010-0277\(00\)00154-2](https://doi.org/10.1016/S0010-0277(00)00154-2)

- Carlson, T. A., Hogendoorn, H., Kanai, R., Mesik, J., & Turret, J. (2011). High temporal resolution decoding of object position and category. *Journal of Vision*, *11*(10), 1–17. <https://doi.org/10.1167/11.10.1>
- Carlson, T. A., Tovar, D. A., Alink, A., & Kriegeskorte, N. (2013). Representational dynamics of object vision: The first 1000 ms. *Journal of Vision*, *13*(10), 1–1. <https://doi.org/10.1167/13.10.1>
- Cavanagh, P., Hunt, A. R., Afraz, A., & Rolfs, M. (2010). Visual stability based on remapping of attention pointers. *Trends in Cognitive Sciences*, *14*(4), 147–153. <https://doi.org/10.1016/j.tics.2010.01.007>
- Centelles, L., Assaiante, C., Nazarian, B., Anton, J. L., & 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*(1). <https://doi.org/10.1371/journal.pone.0015749>
- Chatterjee, A., Maher, L. M., & Heilman, K. M. (1995). Spatial characteristics of thematic role representation. *Neuropsychologia*, *33*(5), 643–648. [https://doi.org/10.1016/0028-3932\(94\)00134-B](https://doi.org/10.1016/0028-3932(94)00134-B)
- 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>
- Cheng, C., & Kibbe, M. M. (2022). Development of updating in working memory in 4–7-year-old children. *Developmental Psychology*, *58*(5), 902–912. <https://doi.org/10.1037/dev0001337>
- Cheng, H., He, K., Li, C., Ma, X., Zheng, F., Xu, W., Liao, P., Yang, R., Li, D., Qin, L., Na, S., Lyu, B., & Gao, J. H. (2024). Wireless optically pumped magnetometer MEG. *NeuroImage*, *300*(April), 120864. <https://doi.org/10.1016/j.neuroimage.2024.120864>
- Chota, S., Leto, C., van Zantwijk, L., & Van der Stigchel, S. (2022). Attention rhythmically samples multi-feature objects in working memory. *Scientific Reports*, *12*(1), 14703. <https://doi.org/10.1038/s41598-022-18819-z>
- Chung, J. E., Sellers, K. K., Leonard, M. K., Gwilliams, L., Xu, D., Dougherty, M. E., Kharazia, V., Metzger, S. L., Welkenhuysen, M., Dutta, B., & Chang, E. F. (2022). High-density single-unit human cortical recordings using the Neuropixels probe. *Neuron*, *110*(15), 2409–2421.e3. <https://doi.org/10.1016/j.neuron.2022.05.007>
- Clevenger, P. E., & Hummel, J. E. (2014). Working memory for relations among objects. *Attention, Perception, and Psychophysics*, *76*(7), 1933–1953. <https://doi.org/10.3758/s13414-013-0601-3>

- Cohn, N., Paczynski, M., & Kutas, M. (2017). Not so secret agents: Event-related potentials to semantic roles in visual event comprehension. *Brain and Cognition*, 119(April), 1–9. <https://doi.org/10.1016/j.bandc.2017.09.001>
- Conway, A. R. A., Cowan, N., Bunting, M. F., Theriault, D. J., & Minkoff, S. R. B. (2002). A latent variable analysis of working memory capacity, short-term memory capacity, processing speed, and general fluid intelligence. *Intelligence*, 30(2), 163–183. [https://doi.org/10.1016/S0160-2896\(01\)00096-4](https://doi.org/10.1016/S0160-2896(01)00096-4)
- Coslett, H. B., & Saffran, E. (1991). Simultanagnosia. To see but not two see. *Brain*, 114(4), 1523–1545. <https://doi.org/10.1093/brain/114.4.1523>
- 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>
- Cowan, N. (2019). Short-term memory based on activated long-term memory: A review in response to Norris (2017). *Psychological Bulletin*, 145(8), 822–847. <https://doi.org/10.1037/bul0000199>
- Cruz Heredia, A. A. L., Dickerson, B., & Lau, E. (2021). Towards Understanding Sustained Neural Activity Across Syntactic Dependencies. *Neurobiology of Language*, 1–22. https://doi.org/10.1162/nol_a_00050
- Culham, J. C., & Valyear, K. F. (2006). Human parietal cortex in action. *Current Opinion in Neurobiology*, 16(2), 205–212. <https://doi.org/10.1016/j.conb.2006.03.005>
- Dale, A. M., Liu, A. K., Fischl, B. R., Buckner, R. L., Belliveau, J. W., Lewine, J. D., & Halgren, E. (2000). Dynamic statistical parametric mapping: Combining fMRI and MEG for high-resolution imaging of cortical activity. *Neuron*, 26(1), 55–67. [https://doi.org/10.1016/S0896-6273\(00\)81138-1](https://doi.org/10.1016/S0896-6273(00)81138-1)
- Davenport, E. M., Urban, J. E., Vaughan, C., DeSimone, J. C., Wagner, B., Espeland, M. A., Powers, A. K., Whitlow, C. T., Stitzel, J. D., & Maldjian, J. A. (2022). MEG measured delta waves increase in adolescents after concussion. *Brain and Behavior*, 12(9), 1–9. <https://doi.org/10.1002/brb3.2720>
- de Cheveigné, A., & Simon, J. Z. (2008). Denoising based on spatial filtering. *Journal of Neuroscience Methods*, 171(2), 331–339. <https://doi.org/10.1016/j.jneumeth.2008.03.015>
- de la Rosa, S., Choudhery, R. N., Curio, C., Ullman, S., Assif, L., & Bühlhoff, H. H. (2015). Visual categorization of social interactions. *Visual Cognition*, 22(January), 1233–1271. <https://doi.org/10.1080/13506285.2014.991368>
- 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>
- Decarli, G., Franchin, L., Piazza, M., & Surian, L. (2020). Infants' use of motion cues in object individuation processes. *Journal of Experimental Child Psychology*, 197, 104868. <https://doi.org/10.1016/j.jecp.2020.104868>
- Delorme, A., & Makeig, S. (2004). EEGLAB: an open source toolbox for analysis of single-trial EEG dynamics including independent component analysis. *Journal of Neuroscience Methods*, 134(1), 9–21. <https://doi.org/10.1016/j.jneumeth.2003.10.009>
- 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
- Dikker, S., Assaneo, M. F., Gwilliams, L., Wang, L., & Kösem, A. (2020). Magnetoencephalography and Language. *Neuroimaging Clinics of North America*, 30(2), 229–238. <https://doi.org/10.1016/j.nic.2020.01.004>
- Ding, N., Melloni, L., Zhang, H., Tian, X., & Poeppel, D. (2015). Cortical tracking of hierarchical linguistic structures in connected speech. *Nature Neuroscience*, 19(1), 158–164. <https://doi.org/10.1038/nn.4186>
- Ding, X., Gao, Z., & Shen, M. (2017). Two Equals One: Two Human Actions During Social Interaction Are Grouped as One Unit in Working Memory. *Psychological Science*, 28(9), 1311–1320. <https://doi.org/10.1177/0956797617707318>
- Dobel, C., Diesendruck, G., & Bölte, J. (2007). How writing system and age influence spatial representations of actions: A developmental, cross-linguistic study: Research report. *Psychological Science*, 18(6), 487–491. <https://doi.org/10.1111/j.1467-9280.2007.01926.x>
- Dobel, C., Enriquez-Geppert, S., Zwitserlood, P., & Bölte, J. (2014). Literacy shapes thought: The case of event representation in different cultures. *Frontiers in Psychology*, 5(APR), 1–5. <https://doi.org/10.3389/fpsyg.2014.00290>
- Dobel, C., Gumnior, H., Bölte, J., & Zwitserlood, P. (2007). Describing scenes hardly seen. *Acta Psychologica*, 125(2), 129–143. <https://doi.org/10.1016/j.actpsy.2006.07.004>
- Dowty, D. (1991). Thematic proto-roles and argument selection. *Language*, 67(3), 547–619. <https://doi.org/10.1353/lan.1991.0021>
- Drew, T., Horowitz, T. S., & Vogel, E. K. (2013). Swapping or dropping? Electrophysiological measures of difficulty during multiple object tracking. *Cognition*, 126(2), 213–223. <https://doi.org/10.1016/j.cognition.2012.10.003>
- Drew, T., Horowitz, T. S., Wolfe, J. M., & Vogel, E. K. (2012). Neural measures of dynamic changes in attentive tracking load. *Journal of Cognitive Neuroscience*, 24(2), 440–450.

https://doi.org/10.1162/jocn_a_00107

- Duncan, D. H., van Moorselaar, D., & Theeuwes, J. (2023). Pinging the brain to reveal the hidden attentional priority map using encephalography. *Nature Communications*, 14(1), 4749. <https://doi.org/10.1038/s41467-023-40405-8>
- 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)
- Eichenbaum, H., & Cohen, N. J. (2014). Can We Reconcile the Declarative Memory and Spatial Navigation Views on Hippocampal Function? *Neuron*, 83(4), 764–770. <https://doi.org/10.1016/j.neuron.2014.07.032>
- Eisenkolb, V. M., Held, L. M., Utzschmid, A., Lin, X. X., Krieg, S. M., Meyer, B., Gempt, J., & Jacob, S. N. (2023). Human acute microelectrode array recordings with broad cortical access, single-unit resolution, and parallel behavioral monitoring. *Cell Reports*, 42(5), 112467. <https://doi.org/10.1016/j.celrep.2023.112467>
- Fan, Y., Han, Q., Guo, S., & Luo, H. (2021). Distinct neural representations of content and ordinal structure in auditory sequence memory. *Journal of Neuroscience*, 41(29), 6290–6303. <https://doi.org/10.1523/JNEUROSCI.0320-21.2021>
- Fan, Y., & Luo, H. (2023). Reactivating ordinal position information from auditory sequence memory in human brains. *Cerebral Cortex*, 33(10), 5924–5936. <https://doi.org/10.1093/cercor/bhac471>
- Fan, Z., Muthukumaraswamy, S. D., Singh, K. D., & Shapiro, K. (2012). The role of sustained posterior brain activity in the serial chaining of two cognitive operations: A MEG study. *Psychophysiology*, 49(8), 1133–1144. <https://doi.org/10.1111/j.1469-8986.2012.01391.x>
- 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)
- Faul, F., Erdfelder, E., Lang, A.-G., & Buchner, A. (2007). G*Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior Research Methods*, 39(2), 175–191. <https://doi.org/10.3758/BF03193146>
- Federmeier, K. D., Segal, J. B., Lombrozo, T., & Kutas, M. (2000). Brain responses to nouns, verbs and class-ambiguous words in context. *Brain*, 123(12), 2552–2566. <https://doi.org/10.1093/brain/123.12.2552>

- Feigenson, L. (2011). Objects, Sets, and Ensembles. In *Space, Time and Number in the Brain* (pp. 13–22). Elsevier. <https://doi.org/10.1016/B978-0-12-385948-8.00002-5>
- Feigenson, L., & Carey, S. (2005). On the limits of infants' quantification of small object arrays. *Cognition*, 97(3), 295–313. <https://doi.org/10.1016/j.cognition.2004.09.010>
- Feigenson, L., Carey, S., & Hauser, M. (2002). The Representations Underlying Infants' Choice of More: Object Files Versus Analog Magnitudes. *Psychological Science*, 13(2), 150–156. <https://doi.org/10.1111/1467-9280.00427>
- 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>
- Feldmann-Wüstefeld, T., Vogel, E. K., & Awh, E. (2018). Contralateral Delay Activity Indexes Working Memory Storage, Not the Current Focus of Spatial Attention. *Journal of Cognitive Neuroscience*, 30(8), 1185–1196. https://doi.org/10.1162/jocn_a_01271
- Feys, O., Corvilain, P., Aeby, A., Sculier, C., Christiaens, F., Holmes, N., Brookes, M., Goldman, S., Wens, V., & De Tiège, X. (2022). On-Scalp Optically Pumped Magnetometers versus Cryogenic Magnetoencephalography for Diagnostic Evaluation of Epilepsy in School-aged Children. *Radiology*, 304(2), 429–434. <https://doi.org/10.1148/radiol.212453>
- Fitch, W. T. (2020). Animal cognition and the evolution of human language: why we cannot focus solely on communication. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 375(1789), 20190046. <https://doi.org/10.1098/rstb.2019.0046>
- Fodor, J. A. (1975). *The language of thought (Vol. 5)*. Harvard University Press.
- Fodor, J. A., & Pylyshyn, Z. W. (1988). Connectionism and cognitive architecture: A critical analysis. *Cognition*, 28(1–2), 3–71. [https://doi.org/10.1016/0010-0277\(88\)90031-5](https://doi.org/10.1016/0010-0277(88)90031-5)
- Fornaciai, M., & Park, J. (2020). Disentangling feedforward versus feedback processing in numerosity representation. *Cortex*, 104743. <https://doi.org/10.1016/j.cortex.2020.11.013>
- 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>
- Fougnie, D., & Marois, R. (2011). What Limits Working Memory Capacity? Evidence for Modality-Specific Sources to the Simultaneous Storage of Visual and Auditory Arrays. *Journal of Experimental Psychology: Learning Memory and Cognition*, 37(6), 1329–1341. <https://doi.org/10.1037/a0024834>
- Fougnie, D., Zughni, S., Godwin, D., & Marois, R. (2015). Working memory storage is intrinsically domain specific. *Journal of Experimental Psychology: General*, 144(1), 30–47.

<https://doi.org/10.1037/a0038211>

- Frankland, S. M., & Greene, J. D. (2015). An architecture for encoding sentence meaning in left mid-superior temporal cortex. *Proceedings of the National Academy of Sciences of the United States of America*, 112(37), 11732–11737. <https://doi.org/10.1073/pnas.1421236112>
- Frankland, S. M., & Greene, J. D. (2020). Two Ways to Build a Thought: Distinct Forms of Compositional Semantic Representation across Brain Regions. *Cerebral Cortex*, 30(6), 3838–3855. <https://doi.org/10.1093/cercor/bhaa001>
- Friedman-Hill, S. R., Robertson, L. C., & Treisman, A. (1995). Parietal Contributions to Visual Feature Binding: Evidence from a Patient with Bilateral Lesions. *Science*, 269(5225), 853–855. <https://doi.org/10.1126/science.7638604>
- Fukuda, K., Awh, E., & Vogel, E. K. (2010). Discrete capacity limits in visual working memory. *Current Opinion in Neurobiology*, 20(2), 177–182. <https://doi.org/10.1016/j.conb.2010.03.005>
- Fukuda, K., Mance, I., & Vogel, E. K. (2015). A 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., Pereira, A. E., Saito, J. M., Tang, T. Y., Tsubomi, H., & Bae, G. Y. (2022). Working Memory Content Is Distorted by Its Use in Perceptual Comparisons. *Psychological Science*, 33(5), 816–829. <https://doi.org/10.1177/09567976211055375>
- 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>
- Gallistel, C. R., & Cramer, A. E. (1996). Computations on metric maps in mammals: Getting oriented and choosing a multi-destination route. *Journal of Experimental Biology*, 199(1), 211–217. <https://doi.org/10.1242/jeb.199.1.211>
- Gallistel, C. R., & King, A. P. (2009). *Memory and the Computational Brain: Why Cognitive Science Will Transform Neuroscience*. Wiley-Blackwell.
- Gallivan, J. P., Chapman, C. S., Wood, D. K., Milne, J. L., Ansari, D., Culham, J. C., & Goodale, M. A. (2011). One to four, and nothing more: Nonconscious parallel individuation of objects during action planning. *Psychological Science*, 22(6), 803–811. <https://doi.org/10.1177/0956797611408733>
- Gao, Z., Ding, X., Yang, T., Liang, J., & Shui, R. (2013). Coarse-to-Fine Construction for High-Resolution Representation in Visual Working Memory. *PLoS ONE*, 8(2). <https://doi.org/10.1371/journal.pone.0057913>

- 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>
- Gaston, P. (2020). *The role of syntactic prediction in auditory word recognition*. University of Maryland.
- 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>
- Geminiani, G., Bisiach, E., Berti, A., & Rusconi, M. L. (1995). Analogical representation and language structure. *Neuropsychologia*, 33(11), 1565–1574. [https://doi.org/10.1016/0028-3932\(95\)00081-D](https://doi.org/10.1016/0028-3932(95)00081-D)
- Genovesio, A., Wise, S. P., & Passingham, R. E. (2014). Prefrontal-parietal function: From foraging to foresight. *Trends in Cognitive Sciences*, 18(2), 72–81. <https://doi.org/10.1016/j.tics.2013.11.007>
- Gerber, E. M. (2021). *Permutest*. <https://www.mathworks.com/matlabcentral/fileexchange/71737-permutest>
- Gershman, S. J. (2023). The molecular memory code and synaptic plasticity: A synthesis. *BioSystems*, 224(October 2022), 104825. <https://doi.org/10.1016/j.biosystems.2022.104825>
- Gialopsou, A., Abel, C., James, T. M., Coussens, T., Bason, M. G., Puddy, R., Lorenzo, F. Di, Rolfs, K., Voigt, J., Sander, T., Cercignani, M., & Krüger, P. (2021). Improved spatio - temporal measurements of visually evoked fields using optically - pumped magnetometers. *Scientific Reports*, 1–11. <https://doi.org/10.1038/s41598-021-01854-7>
- Golomb, J. D., Kupitz, C. N., & Thiemann, C. T. (2014). The influence of object location on identity: A “spatial congruency bias”. *Journal of Experimental Psychology: General*, 143(6), 2262–2278. <https://doi.org/10.1037/xge0000017>
- Goodale, M. A., & Milner, A. D. (1992). Separate visual pathways for perception and action. *Trends in Neurosciences*, 15(1), 20–25. [https://doi.org/10.1016/0166-2236\(92\)90344-8](https://doi.org/10.1016/0166-2236(92)90344-8)
- Gramfort, A., Luessi, M., Larson, E., Engemann, D. A., Strohmeier, D., Brodbeck, C., Goj, R., Jas, M., Brooks, T., Parkkonen, L., & Hämäläinen, M. (2013). MEG and EEG data analysis with MNE-Python. *Frontiers in Neuroscience*, 7(7 DEC), 1–13. <https://doi.org/10.3389/fnins.2013.00267>
- Green, E. J., & Quilty-Dunn, J. (2021). What Is an Object File? *The British Journal for the Philosophy of Science*, 72(3), 665–699. <https://doi.org/10.1093/bjps/axx055>

- Griffiths, T. D., & Warren, J. D. (2004). What is an auditory object? *Nature Reviews Neuroscience*, 5(11), 887–892. <https://doi.org/10.1038/nrn1538>
- Grootswagers, T., Robinson, A. K., Shatek, S. M., & Carlson, T. A. (2019). Untangling featural and conceptual object representations. *NeuroImage*, 202, 116083. <https://doi.org/10.1016/j.neuroimage.2019.116083>
- Gwilliams, L., Lewis, G. A., & Marantz, A. (2016). Functional characterisation of letter-specific responses in time, space and current polarity using magnetoencephalography. *NeuroImage*, 132, 320–333. <https://doi.org/10.1016/j.neuroimage.2016.02.057>
- Hafri, A., Bonner, M. F., Landau, B., & Firestone, C. (2024). A phone in a basket looks like a knife in a cup: Role-filler independence in visual processing. *Open Mind*, 8, 766–794. https://doi.org/10.1162/opmi_a_00146
- Hafri, A., Papafragou, A., & Trueswell, J. C. (2013). Getting the gist of events: Recognition of two-participant actions from brief displays. *Journal of Experimental Psychology: General*, 142(3), 880–905. <https://doi.org/10.1037/a0030045>
- Hafri, A., Trueswell, J. C., & Epstein, R. A. (2017). Neural representations of observed actions generalize across static and dynamic visual input. *Journal of Neuroscience*, 37(11), 3056–3071. <https://doi.org/10.1523/JNEUROSCI.2496-16.2017>
- Hakim, N., Adam, K. C. S., Gunseli, E., Awh, E., & Vogel, E. K. (2019). Dissecting the Neural Focus of Attention Reveals Distinct Processes for Spatial Attention and Object-Based Storage in Visual Working Memory. *Psychological Science*, 30(4), 526–540. <https://doi.org/10.1177/0956797619830384>
- Halford, G. S., Phillips, S., Wilson, W. H., McCredden, J., Glenda, A., Birney, D., Baker, R., & Bain, J. D. (2007). Relational processing is fundamental to the central executive and is limited to four variables. In *The Cognitive Neuroscience of Working Memory* (pp. 261–280). Oxford University Press. <https://doi.org/10.1093/acprof:oso/9780198570394.003.0016>
- Halford, G. S., Wilson, W. H., & Phillips, S. (1998). Processing capacity defined by relational complexity: Implications for comparative, developmental, and cognitive psychology. *Behavioral and Brain Sciences*, 21(6), 803–864. <https://doi.org/10.1017/S0140525X98001769>
- Hämäläinen, M. S., & Sarvas, J. (1989). Realistic Conductivity Geometry Model of the Human Head for Interpretation of Neuromagnetic Data. *IEEE Transactions on Biomedical Engineering*, 36(2), 165–171. <https://doi.org/10.1109/10.16463>
- Harrison, A., Jolicoeur, P., & Marois, R. (2010). “What” and “Where” in the Intraparietal Sulcus: An fMRI Study of Object Identity and Location in Visual Short-Term Memory. *Cerebral Cortex*, 20(10), 2478–2485. <https://doi.org/10.1093/cercor/bhp314>

- Hebart, M. N., Zheng, C. Y., Pereira, F., & Baker, C. I. (2020). Revealing the multidimensional mental representations of natural objects underlying human similarity judgements. *Nature Human Behaviour*, 4(11), 1173–1185. <https://doi.org/10.1038/s41562-020-00951-3>
- Hedayati, S., O'Donnell, R. E., & Wyble, B. (2022). A model of working memory for latent representations. *Nature Human Behaviour*, 6(5), 709–719. <https://doi.org/10.1038/s41562-021-01264-9>
- Heim, I. (1982). *The Semantics of Definite and Indefinite Noun Phrases*. Umass Amherst.
- Howard, W. A. (1980). The formulae-as-types notion of construction. In *To H. B. Curry: Essays on Combinatory Logic, Lambda Calculus, and Formalism*. (pp. 479–490).
- Hu, Y., & Yu, Q. (2023). Spatiotemporal dynamics of self-generated imagery reveal a reverse cortical hierarchy from cue-induced imagery. *Cell Reports*, 42(10), 113242. <https://doi.org/10.1016/j.celrep.2023.113242>
- Huang, Q., Zhang, H., & Luo, H. (2021). Sequence structure organizes items in varied latent states of working memory neural network. *ELife*, 10. <https://doi.org/10.7554/eLife.67589>
- Hultén, A., Karvonen, L., Laine, M., & Salmelin, R. (2014). Producing Speech with a Newly Learned Morphosyntax and Vocabulary: An Magnetoencephalography Study. *Journal of Cognitive Neuroscience*, 26(8), 1721–1735. https://doi.org/10.1162/jocn_a_00558
- Hummel, J. E., & Holyoak, K. J. (2003). A Symbolic-Connectionist Theory of Relational Inference and Generalization. *Psychological Review*, 110(2), 220–264. <https://doi.org/10.1037/0033-295X.110.2.220>
- Hyun, J., Woodman, G. F., Vogel, E. K., Hollingworth, A., & Luck, S. J. (2009). The comparison of visual working memory representations with perceptual inputs. *Journal of Experimental Psychology: Human Perception and Performance*, 35(4), 1140–1160. <https://doi.org/10.1037/a0015019>
- Iivanainen, J., Carter, T. R., Trumbo, M. C. S., McKay, J., Taulu, S., Wang, J., Stephen, J. M., Schwindt, P. D. D., & Borna, A. (2023). Single-trial classification of evoked responses to auditory tones using OPM- and SQUID-MEG. *Journal of Neural Engineering*, 20(5), 056032. <https://doi.org/10.1088/1741-2552/acfd9>
- Iivanainen, J., Zetter, R., & Parkkonen, L. (2020). Potential of on-scalp MEG: Robust detection of human visual gamma-band responses. *Human Brain Mapping*, 41(1), 150–161. <https://doi.org/10.1002/hbm.24795>
- Isik, L., Koldewyn, K., Beeler, D., & Kanwisher, N. (2017). Perceiving social interactions in the posterior superior temporal sulcus. *Proceedings of the National Academy of Sciences*, 114(43), E113–E114. <https://doi.org/10.1073/pnas.1714471114>

- 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
- Jackendoff, R. (1987). On beyond Zebra: The relation of linguistic and visual information. *Cognition*, 26(2), 89–114. [https://doi.org/10.1016/0010-0277\(87\)90026-6](https://doi.org/10.1016/0010-0277(87)90026-6)
- JASP Team. (2022). *JASP 0.16.1*.
- JASP Team. (2023). *JASP 0.18.1*.
- Jazbinšek, V., Marhl, U., & Sander, T. (2022). SERF-OPM Usability for MEG in Two-Layer-Shielded Rooms. In *Flexible High Performance Magnetic Field Sensors* (pp. 179–193). Springer International Publishing. https://doi.org/10.1007/978-3-031-05363-4_10
- Jessop, A., & Chang, F. (2020). Thematic role information is maintained in the visual object-tracking system. *Quarterly Journal of Experimental Psychology*, 73(1), 146–163. <https://doi.org/10.1177/1747021819882842>
- Jessop, A., & Chang, F. (2022). Thematic role tracking difficulties across multiple visual events influences role use in language production. *Visual Cognition*, 30(3), 151–173. <https://doi.org/10.1080/13506285.2021.2013374>
- Ji, H., & Scholl, B. J. (2024). “Visual verbs”: Dynamic event types are extracted spontaneously during visual perception. *Journal of Experimental Psychology. General*, 153(10), 2441–2453. <https://doi.org/10.1037/xge0001636>
- Jia, J., Liu, L., Fang, F., & Luo, H. (2017). Sequential sampling of visual objects during sustained attention. *PLoS Biology*, 15(6), 1–19. <https://doi.org/10.1371/journal.pbio.2001903>
- Jia, K., Li, Y., Gong, M., Huang, H., Wang, Y., & Li, S. (2021). Perceptual Learning beyond Perception: Mnemonic Representation in Early Visual Cortex and Intraparietal Sulcus. *The Journal of Neuroscience*, 41(20), 4476–4486. <https://doi.org/10.1523/jneurosci.2780-20.2021>
- Jodko-Włodzińska, A., & Sander, T. (2024). Emotionally Charged Visually Evoked Magnetic Fields. *Acta Physica Polonica A*, 146(4), 521–525. <https://doi.org/10.12693/APhysPolA.146.521>
- Johnson, M. G. (1967). Syntactic position and rated meaning. *Journal of Verbal Learning and Verbal Behavior*, 6(2), 240–246. [https://doi.org/10.1016/S0022-5371\(67\)80103-8](https://doi.org/10.1016/S0022-5371(67)80103-8)
- 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*, 23(9), 5617. <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. *The Journal of Neuroscience : The Official Journal of the Society for Neuroscience*, 44(44), 1–18. <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). 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. <https://doi.org/10.1016/j.neuroimage.2014.11.024>
- Kaan, E., Harris, A., Gibson, E., & Holcomb, P. (2000). The P600 as an index of syntactic integration difficulty. *Language and Cognitive Processes*, 15(2), 159–201. <https://doi.org/10.1080/016909600386084>
- Kahneman, D., Treisman, A., & Gibbs, B. J. (1992). The reviewing of object files: Object-specific integration of information. *Cognitive Psychology*, 24(2), 175–219. [https://doi.org/10.1016/0010-0285\(92\)90007-O](https://doi.org/10.1016/0010-0285(92)90007-O)
- Kako, E. (2006a). The semantics of syntactic frames. *Language and Cognitive Processes*, 21(5), 562–575. <https://doi.org/10.1080/01690960500101967>
- Kako, E. (2006b). Thematic role properties of subjects and objects. *Cognition*, 101(1), 1–42. <https://doi.org/10.1016/j.cognition.2005.08.002>
- Kamiński, J., & Rutishauser, U. (2020). Between persistently active and activity-silent frameworks: novel vistas on the cellular basis of working memory. *Annals of the New York Academy of Sciences*, 1464(1), 64–75. <https://doi.org/10.1111/nyas.14213>
- Kamp, H. (1981). A theory of truth and semantic representation. In J. A. G. Groenendijk, T. M. V. Janssen, & M. B. J. Stokhof (Eds.), *Formal methods in the study of language* (pp. 277–322). Mathematisch Centrum.
- Kanwisher, N. (1991). Repetition Blindness and Illusory Conjunctions: Errors in Binding Visual Types With Visual Tokens. *Journal of Experimental Psychology: Human Perception and Performance*, 17(2), 404–421. <https://doi.org/10.1037/0096-1523.17.2.404>
- 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>
- Kazandjian, S., Gaash, E., Love, I. Y., Zivotofsky, A. Z., & Chokron, S. (2011). Spatial representation of action phrases among bidirectional readers: The effect of language environment and sentence complexity. *Social Psychology*, 42(3), 249–258. <https://doi.org/10.1027/1864-9335/a000069>
- Kazanina, N., & Poeppel, D. (2023). The neural ingredients for a language of thought are

available. *Trends in Cognitive Sciences*, 27(11), 996–1007.
<https://doi.org/10.1016/j.tics.2023.07.012>

Khader, P., Scherag, A., Streb, J., & Rösler, F. (2003). Differences between noun and verb processing in a minimal phrase context: A semantic priming study using event-related brain potentials. *Cognitive Brain Research*, 17(2), 293–313. [https://doi.org/10.1016/S0926-6410\(03\)00130-7](https://doi.org/10.1016/S0926-6410(03)00130-7)

Kibbe, M. M., & Leslie, A. M. (2011). What do infants remember when they forget? location and identity in 6-month-olds' memory for objects. *Psychological Science*, 22(12), 1500–1505. <https://doi.org/10.1177/0956797611420165>

Kibbe, M. M., & Leslie, A. M. (2013). What's the object of object working memory in infancy? Unraveling 'what' and 'how many.' *Cognitive Psychology*, 66(4), 380–404. <https://doi.org/10.1016/j.cogpsych.2013.05.001>

Kibbe, M. M., & Leslie, A. M. (2019). Conceptually Rich, Perceptually Sparse: Object Representations in 6-Month-Old Infants' Working Memory. *Psychological Science*, 30(3), 362–375. <https://doi.org/10.1177/0956797618817754>

Kim, J. G., Biederman, I., Lescroart, M. D., & Hayworth, K. J. (2009). Adaptation to objects in the lateral occipital complex (LOC): Shape or semantics? *Vision Research*, 49(18), 2297–2305. <https://doi.org/10.1016/j.visres.2009.06.020>

Kim, S. G., Overath, T., Sedley, W., Kumar, S., Teki, S., Kikuchi, Y., Patterson, R., & Griffiths, T. D. (2022). MEG correlates of temporal regularity relevant to pitch perception in human auditory cortex. *NeuroImage*, 249(January). <https://doi.org/10.1016/j.neuroimage.2022.118879>

Kim, S. G., Poeppel, D., & Overath, T. (2020). Modulation change detection in human auditory cortex: evidence for asymmetric, nonlinear edge detection. *The European Journal of Neuroscience*, January, 1–16. <https://doi.org/10.1111/ejn.14707>

Klaver, P., Smid, H. G. O. ., & Heinze, H.-J. (1999). Representations in human visual short-term memory: an event-related brain potential study. *Neuroscience Letters*, 268(2), 65–68. [https://doi.org/10.1016/S0304-3940\(99\)00380-8](https://doi.org/10.1016/S0304-3940(99)00380-8)

Klaver, P., Talsma, D., Wijers, A. A., Heinze, H. J., & Mulder, G. (1999). An event-related brain potential correlate of visual short-term memory. *NeuroReport*, 10(10), 2001–2005. <https://doi.org/10.1097/00001756-199907130-00002>

Kleiner, M., Brainard, D., Pelli, D., Ingling, A., Murray, R., & Broussard, C. (2007). What's new in psychtoolbox-3. *Perception*, 36(1_suppl), 1–235. <https://doi.org/10.1177/03010066070360S101>

Klug, M., & Gramann, K. (2021). Identifying key factors for improving ICA-based

- decomposition of EEG data in mobile and stationary experiments. *European Journal of Neuroscience*, 54(12), 8406–8420. <https://doi.org/10.1111/ejn.14992>
- Knops, A., Piazza, M., Sengupta, R., Eger, E., & Melcher, D. (2014). A shared, flexible neural map architecture reflects capacity limits in both visual short-term memory and enumeration. *Journal of Neuroscience*, 34(30), 9857–9866. <https://doi.org/10.1523/JNEUROSCI.2758-13.2014>
- Knowlton, T. Z., & Gomes, V. (2022). Linguistic and non-linguistic cues to acquiring the strong distributivity of each. *Proceedings of the Linguistic Society of America*, 7(1), 5236. <https://doi.org/10.3765/plsa.v7i1.5236>
- Knowlton, T. Z., Halberda, J., Pietroski, P., & Lidz, J. (2023). Individuals versus ensembles and “each” versus “every”: linguistic framing affects performance in a change detection task. *Glossa Psycholinguistics*, 2(1), 1–30. <https://doi.org/10.5070/g6011181>
- Kominis, I. K., Kornack, T. W., Allred, J. C., & Romalis, M. V. (2003). A subfemtotesla multichannel atomic magnetometer. *Nature*, 422(6932), 596–599. <https://doi.org/10.1038/nature01484>
- Konkel, A., & Cohen, N. J. (2009). Relational memory and the hippocampus: Representations and methods. *Frontiers in Neuroscience*, 3(SEP), 166–174. <https://doi.org/10.3389/neuro.01.023.2009>
- 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., & Shwartz, 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)
- Kreither, J., Papaioannou, O., & Luck, S. J. (2022). Active Working Memory and Simple Cognitive Operations. *Journal of Cognitive Neuroscience*, 34(2), 313–331. https://doi.org/10.1162/jocn_a_01791
- Küçük, E., Foxwell, M., Kaiser, D., & Pitcher, D. (2024). Moving and Static Faces, Bodies, Objects, and Scenes Are Differentially Represented across the Three Visual Pathways. *Journal of Cognitive Neuroscience*, 36(12), 2639–2651. https://doi.org/10.1162/jocn_a_02139
- Kutter, E. F., Dehnen, G., Borger, V., Surges, R., Mormann, F., & Nieder, A. (2023). Distinct neuronal representation of small and large numbers in the human medial temporal lobe. *Nature Human Behaviour*, 7(11), 1998–2007. <https://doi.org/10.1038/s41562-023-01709-3>
- Lalisse, M., & Smolensky, P. (2021). Distributed neural encoding of binding to thematic roles. *ArXiv*, 8. <http://arxiv.org/abs/2110.12342>

- Lapinskaya, N., Uzomah, U., Bedny, M., & Lau, E. (2016). Electrophysiological signatures of event words: Dissociating syntactic and semantic category effects in lexical processing. *Neuropsychologia*, 93(June), 151–157. <https://doi.org/10.1016/j.neuropsychologia.2016.10.014>
- Lau, E. (2018). Neural Indices of Structured Sentence Representation. In *Psychology of Learning and Motivation* (Vol. 68, pp. 117–142). Elsevier Ltd. <https://doi.org/10.1016/bs.plm.2018.08.004>
- Lau, E., & Liao, C. H. (2018). Linguistic structure across time: ERP responses to coordinated and uncoordinated noun phrases. *Language, Cognition and Neuroscience*, 33(5), 633–647. <https://doi.org/10.1080/23273798.2017.1400081>
- Lau, E., & Nguyen, E. (2015). The role of temporal predictability in semantic expectation: An MEG investigation. *Cortex*, 68, 8–19. <https://doi.org/10.1016/j.cortex.2015.02.022>
- Lau, E., Phillips, C., & Poeppel, D. (2008). A cortical network for semantics: (De)constructing the N400. *Nature Reviews Neuroscience*, 9(12), 920–933. <https://doi.org/10.1038/nrn2532>
- Lau, E., Weber, K., Gramfort, A., Hämäläinen, M. S., & Kuperberg, G. R. (2016). Spatiotemporal Signatures of Lexical-Semantic Prediction. *Cerebral Cortex*, 26(4), 1377–1387. <https://doi.org/10.1093/cercor/bhu219>
- Lau, E., & Yang, X. (2023). Neural response to coordination reflects interpretive processes, not syntactic connectedness [conference presentation]. *Proceedings of the Human Sentence Processing 2023*.
- Leavitt, M. L., Mendoza-Halliday, D., & Martinez-Trujillo, J. C. (2017). Sustained Activity Encoding Working Memories: Not Fully Distributed. *Trends in Neurosciences*, 40(6), 328–346. <https://doi.org/10.1016/j.tins.2017.04.004>
- Lee, R. R., & Huang, M. (2012). Magnetoencephalography in the diagnosis of concussion. *Concussion*, November, 94–111. <https://doi.org/10.1159/000358768>
- Lefebvre, C., & Jolicœur, P. (2016). Memory for pure tone sequences without contour. *Brain Research*, 1640, 222–231. <https://doi.org/10.1016/j.brainres.2016.02.025>
- Leslie, A. M., Xu, F., Tremoulet, P. D., & Scholl, B. J. (1998). Indexing and the object concept: Developing “what” and “where” systems. *Trends in Cognitive Sciences*, 2(1), 10–18. [https://doi.org/10.1016/S1364-6613\(97\)01113-3](https://doi.org/10.1016/S1364-6613(97)01113-3)
- Lewis-Peacock, J. A., Drysdale, A. T., Oberauer, K., & Postle, B. R. (2012). Neural evidence for a distinction between short-term memory and the focus of attention. *Journal of Cognitive Neuroscience*, 24(1), 61–79. https://doi.org/10.1162/jocn_a_00140
- Li, C., Tian, W., He, Y., Wang, C., Wang, X., Xu, X., Bai, L., Xue, T., Liao, Y., Xu, T., Liu, X.,

- & Wu, S. (2023). How are patterned movements stored in working memory? *Frontiers in Psychology*, 14. <https://doi.org/10.3389/fpsyg.2023.1074520>
- Li, Y., & Noguchi, Y. (2022). Neural correlates of a load-dependent decline in visual working memory. *Cerebral Cortex Communications*, 3(2). <https://doi.org/10.1093/texcom/tgac015>
- Liao, C. H., & Lau, E. (2020). Enough time to get results? An ERP investigation of prediction with complex events. *Language, Cognition and Neuroscience*, 35(9), 1162–1182. <https://doi.org/10.1080/23273798.2020.1733626>
- Lin, F. H., Belliveau, J. W., Dale, A. M., & Hämäläinen, M. S. (2006). Distributed current estimates using cortical orientation constraints. *Human Brain Mapping*, 27(1), 1–13. <https://doi.org/10.1002/hbm.20155>
- Lin, X. X., Nieder, A., & Jacob, S. N. (2023). The neuronal implementation of representational geometry in primate prefrontal cortex. *Science Advances*, 9(50). <https://doi.org/10.1126/sciadv.adh8685>
- Liu, C., Wang, R., Li, L., Ding, G., Yang, J., & Li, P. (2020). Effects of encoding modes on memory of naturalistic events. *Journal of Neurolinguistics*, 53(July 2019). <https://doi.org/10.1016/j.jneuroling.2019.100863>
- 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>
- Liu, Y., Lu, X., Wu, F., Shen, M., & Gao, Z. (2019). Biological motion is stored independently from bound representation in working memory. *Visual Cognition*, 27(9–10), 701–713. <https://doi.org/10.1080/13506285.2019.1638479>
- Lo, C. W., & Brennan, J. R. (2021). EEG Correlates of Long-Distance Dependency Formation in Mandarin Wh-Questions. *Frontiers in Human Neuroscience*, 15(February), 1–18. <https://doi.org/10.3389/fnhum.2021.591613>
- Lobina, D. J., & García-Albea, J. E. (2017). On Language and Thought: A Question of Formats. In *On Concepts, Modules, and Language: Cognitive Science at Its Core* (pp. 249–268). Oxford University Press. <https://doi.org/10.1093/oso/9780190464783.003.0012>
- Long, B., Störmer, V. S., & Alvarez, G. A. (2017). Mid-level perceptual features contain early cues to animacy. *Journal of Vision*, 17(6), 20. <https://doi.org/10.1167/17.6.20>
- Lopez-Calderon, J., & Luck, S. J. (2014). ERPLAB: an open-source toolbox for the analysis of event-related potentials. *Frontiers in Human Neuroscience*, 8. <https://doi.org/10.3389/fnhum.2014.00213>
- Lu, X., Dai, A., Guo, Y., Shen, M., & Gao, Z. (2022). Is the social chunking of agent actions in

- working memory resource-demanding? *Cognition*, 229, 105249.
<https://doi.org/10.1016/j.cognition.2022.105249>
- Lu, X., Huang, J., Yi, Y., Shen, M., Weng, X., & Gao, Z. (2016). Holding biological motion in working memory: An fMRI study. *Frontiers in Human Neuroscience*, 10(June).
<https://doi.org/10.3389/fnhum.2016.00251>
- Lu, X., Ma, X., Zhao, Y., Gao, Z., & Shen, M. (2019). Retaining event files in working memory requires extra object-based attention than the constituent elements. *Quarterly Journal of Experimental Psychology*, 72(9), 2225–2239. <https://doi.org/10.1177/1747021819839993>
- Lu, Y., Jin, P., Pan, X., & Ding, N. (2022). Delta-band neural activity primarily tracks sentences instead of semantic properties of words. *NeuroImage*, 118979.
<https://doi.org/10.1016/j.neuroimage.2022.118979>
- Luck, S. J. (2014). *An Introduction to the Event-related Potential Technique* (Second Edi). The MIT Press.
- Luck, S. J., & Vogel, E. K. (1997). The capacity of visual working memory for features and conjunctions. *Nature*, 390(6657), 279–281. <https://doi.org/10.1038/36846>
- Luck, S. J., & Vogel, E. K. (2013). Visual working memory capacity: From psychophysics and neurobiology to individual differences. *Trends in Cognitive Sciences*, 17(8), 391–400.
<https://doi.org/10.1016/j.tics.2013.06.006>
- Lundqvist, M., Brincat, S. L., Rose, J., Warden, M. R., Buschman, T. J., Miller, E. K., & Herman, P. (2023). Working memory control dynamics follow principles of spatial computing. *Nature Communications*, 14(1), 1–12. <https://doi.org/10.1038/s41467-023-36555-4>
- Lundqvist, M., Rose, J., Warden, M. R., Buschman, T., Miller, E. K., & Herman, P. (2021). A Hot-Coal theory of Working Memory. *BioRxiv*, 2020.12.30.424833.
<https://doi.org/10.1101/2020.12.30.424833>
- Luria, R., Balaban, H., Awh, E., & Vogel, E. K. (2016). The contralateral delay activity as a neural measure of visual working memory. *Neuroscience and Biobehavioral Reviews*, 62, 100–108. <https://doi.org/10.1016/j.neubiorev.2016.01.003>
- Luria, R., & Vogel, E. K. (2011). Shape and color conjunction stimuli are represented as bound objects in visual working memory. *Neuropsychologia*, 49(6), 1632–1639.
<https://doi.org/10.1016/j.neuropsychologia.2010.11.031>
- Luria, R., & Vogel, E. K. (2014). Come Together, Right Now: Dynamic Overwriting of an Object's History through Common Fate. *Journal of Cognitive Neuroscience*, 26(8), 1819–1828. https://doi.org/10.1162/jocn_a_00584

- Luyckx, F., Nili, H., Spitzer, B., & Summerfield, C. (2019). Neural structure mapping in human probabilistic reward learning. *ELife*, 8, 1–19. <https://doi.org/10.7554/eLife.42816>
- Lyu, C., Hu, S., Wei, L., Zhang, X., & Talhelm, T. (2015). Brain Activation of Identity Switching in Multiple Identity Tracking Task. *PLOS ONE*, 10(12), e0145489. <https://doi.org/10.1371/journal.pone.0145489>
- Ma, W. J., Husain, M., & Bays, P. M. (2014). Changing concepts of working memory. *Nature Neuroscience*, 17(3), 347–356. <https://doi.org/10.1038/nn.3655>
- Maass, A., & Russo, A. (2003). Directional Bias in the Mental Representation of Spatial Events: Nature or Culture? *Psychological Science*, 14(4), 296–301. <https://doi.org/10.1111/1467-9280.14421>
- Macnamara, J. (1999). *Through the rearview mirror: Historical reflections on psychology*. The MIT Press.
- Madnani, N., Boyd-Graber, J., & Resnik, P. (2010). Measuring transitivity using untrained annotators. *Proceedings of the NAACL HLT 2010 Workshop on Creating Speech and Language Data with Amazon's Mechanical Turk, June*, 188–194.
- Maess, B., Mamashli, F., Obleser, J., Helle, L., & Friederici, A. D. (2016). Prediction signatures in the brain: Semantic pre-activation during language comprehension. *Frontiers in Human Neuroscience*, 10(NOV2016), 1–11. <https://doi.org/10.3389/fnhum.2016.00591>
- Mahon, B. Z., & Kemmerer, D. (2020). Interactions between language, thought, and perception: Cognitive and neural perspectives. *Cognitive Neuropsychology*, 37(5–6), 235–240. <https://doi.org/10.1080/02643294.2020.1829578>
- 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>
- Marhl, U., Jodko-Władzińska, A., Brühl, R., Sander, T., & Jazbinšek, V. (2022). Transforming and comparing data between standard SQUID and OPM-MEG systems. *PLoS ONE*, 17(1 January), 1–22. <https://doi.org/10.1371/journal.pone.0262669>
- Maris, E., & Oostenveld, R. (2007). Nonparametric statistical testing of EEG- and MEG-data. *Journal of Neuroscience Methods*, 164(1), 177–190. <https://doi.org/10.1016/j.jneumeth.2007.03.024>
- Martin, A. E., & Doumas, L. A. A. (2017). A mechanism for the cortical computation of hierarchical linguistic structure. *PLoS Biology*, 15(3), 1–23. <https://doi.org/10.1371/journal.pbio.2000663>

- Martin, B., Wiener, M., & Van Wassenhove, V. (2017). A Bayesian Perspective on Accumulation in the Magnitude System. *Scientific Reports*, 7(1), 1–14. <https://doi.org/10.1038/s41598-017-00680-0>
- Matchin, W., Brodbeck, C., Hammerly, C., & Lau, E. (2019). The temporal dynamics of structure and content in sentence comprehension: Evidence from fMRI-constrained MEG. *Human Brain Mapping*, 40(2), 663–678. <https://doi.org/10.1002/hbm.24403>
- Matchin, W., Liao, C. H., Gaston, P., & Lau, E. (2019). Same words, different structures: An fMRI investigation of argument relations and the angular gyrus. *Neuropsychologia*, 125(January), 116–128. <https://doi.org/10.1016/j.neuropsychologia.2019.01.019>
- Matsuyoshi, D., Ikeda, T., Sawamoto, N., Kakigi, R., Fukuyama, H., & Osaka, N. (2010). Task-irrelevant memory load induces inattentional blindness without temporo-parietal suppression. *Neuropsychologia*, 48(10), 3094–3101. <https://doi.org/10.1016/j.neuropsychologia.2010.06.021>
- Matthey, L., Bays, P. M., & Dayan, P. (2015). A Probabilistic Palimpsest Model of Visual Short-term Memory. *PLOS Computational Biology*, 11(1), e1004003. <https://doi.org/10.1371/journal.pcbi.1004003>
- Mazza, V., & Caramazza, A. (2015). Multiple object individuation and subitizing in enumeration: A view from electrophysiology. *Frontiers in Human Neuroscience*, 9(APRIL), 1–7. <https://doi.org/10.3389/fnhum.2015.00162>
- McCants, C. W., Katus, T., & Eimer, M. (2020). Task goals modulate the activation of part-based versus object-based representations in visual working memory. *Cognitive Neuroscience*, 11(1–2), 92–100. <https://doi.org/10.1080/17588928.2019.1642864>
- Mccollough, A. W., Machizawa, M. G., & Vogel, E. K. (2007). Electrophysiological Measures of Maintaining Representations in Visual Working Memory. *Cortex*, 43, 77–94.
- McKinnon, R., & Osterhout, L. (1996). Constraints on Movement Phenomena in Sentence Processing: Evidence from Event-related Brain Potentials. *Language and Cognitive Processes*, 11(5), 495–524. <https://doi.org/10.1080/016909696387132>
- Mecklinger, A., & Pfeifer, E. (1996). Event-related potentials reveal topographical and temporal distinct neuronal activation patterns for spatial and object working memory. *Cognitive Brain Research*, 4(3), 211–224. [https://doi.org/10.1016/S0926-6410\(96\)00034-1](https://doi.org/10.1016/S0926-6410(96)00034-1)
- Menenti, L., Segaert, K., & Hagoort, P. (2012). The neuronal infrastructure of speaking. *Brain and Language*, 122(2), 71–80. <https://doi.org/10.1016/j.bandl.2012.04.012>
- Mitchell, D. J., & Cusack, R. (2008). Flexible, capacity-limited activity of posterior parietal cortex in perceptual as well as visual short-term memory tasks. *Cerebral Cortex*, 18(8), 1788–1798. <https://doi.org/10.1093/cercor/bhm205>

- Mitchell, D. J., & Cusack, R. (2011). The temporal evolution of electromagnetic markers sensitive to the capacity limits of visual short-term memory. *Frontiers in Human Neuroscience*, 5(FEBRUARY), 1–20. <https://doi.org/10.3389/fnhum.2011.00018>
- 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>
- Morey, C. C., & Cowan, N. (2004). When visual and verbal memories compete: Evidence of cross-domain limits in working memory. *Psychonomic Bulletin and Review*, 11(2), 296–301. <https://doi.org/10.3758/BF03196573>
- Morey, C. C., & Cowan, N. (2005). When do visual and verbal memories conflict? The importance of working-memory load and retrieval. *Journal of Experimental Psychology: Learning Memory and Cognition*, 31(4), 703–713. <https://doi.org/10.1037/0278-7393.31.4.703>
- Morrison, C., Kamal, F., Le, K., & Taler, V. (2020). Monolinguals and bilinguals respond differently to a delayed matching-to-sample task: An ERP study. *Bilingualism*, 23(4), 858–868. <https://doi.org/10.1017/S1366728919000816>
- Mosher, J. C., Leahy, R. M., & Lewis, P. S. (1999). EEG and MEG: Forward solutions for inverse methods. *IEEE Transactions on Biomedical Engineering*, 46(3), 245–259. <https://doi.org/10.1109/10.748978>
- Murez, M., & Smortchkova, J. (2014). Singular thought: Object-files, person-files, and the sortal PERSON. *Topics in Cognitive Science*, 6(4), 632–646. <https://doi.org/10.1111/tops.12110>
- 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).
- Nau, M., Schmid, A. C., Kaplan, S. M., Baker, C. I., & Kravitz, D. J. (2024). Centering cognitive neuroscience on task demands and generalization. *Nature Neuroscience*, 27(September), 1656–1667. <https://doi.org/10.1038/s41593-024-01711-6>
- 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>

- Ngiam, W. X. Q., Foster, J. J., Adam, K. C. S., & Awh, E. (2023). Distinguishing guesses from fuzzy memories: Further evidence for item limits in visual working memory. *Attention, Perception, & Psychophysics*, 85(5), 1695–1709. <https://doi.org/10.3758/s13414-022-02631-y>
- Ngiam, W. X. Q., Khaw, K. L. C., Holcombe, A. O., & Goodbourn, P. T. (2019). Visual Working Memory for Letters Varies With Familiarity but Not Complexity. *Journal of Experimental Psychology: Learning Memory and Cognition*, 45(10), 1761–1775. <https://doi.org/10.1037/xlm0000682>
- Ngiam, W. X. Q., Thyer, W., Epstein, W. S., Jones, H. M., Suplica, D., & Awh, E. (2023). Associative learning changes multivariate neural signatures of working memory load. *OPAM 2023*.
- O’Connell, R. G., Dockree, P. M., & Kelly, S. P. (2012). A supramodal accumulation-to-bound signal that determines perceptual decisions in humans. *Nature Neuroscience*, 15(12), 1729–1735. <https://doi.org/10.1038/nn.3248>
- O’Connell, R. G., & Kelly, S. P. (2021). Neurophysiology of Human Perceptual Decision-Making. *Annual Review of Neuroscience*, 44(1). <https://doi.org/10.1146/annurev-neuro-092019-100200>
- O’Donnell, R. E., Clement, A., & Brockmole, J. R. (2018). Semantic and functional relationships among objects increase the capacity of visual working memory. *Journal of Experimental Psychology: Learning Memory and Cognition*, 44(7), 1151–1158. <https://doi.org/10.1037/xlm0000508>
- O’Reilly, R. C., Ranganath, C., & Russin, J. L. (2022). The Structure of Systematicity in the Brain. *Current Directions in Psychological Science*, 31(2), 124–130. <https://doi.org/10.1177/09637214211049233>
- Oberauer, K. (2017). What is working memory capacity? / ¿Qué es la capacidad de la memoria de trabajo? *Estudios de Psicología*, 38(2), 338–384. <https://doi.org/10.1080/02109395.2017.1295579>
- Odic, D., & Starr, A. (2018). An Introduction to the Approximate Number System. *Child Development Perspectives*, 12(4), 223–229. <https://doi.org/10.1111/cdep.12288>
- Oldfield, R. C. (1971). The assessment and analysis of handedness: The Edinburgh inventory. *Neuropsychologia*, 9(1), 97–113. [https://doi.org/10.1016/0028-3932\(71\)90067-4](https://doi.org/10.1016/0028-3932(71)90067-4)
- Olson, I. R., & Berryhill, M. (2009). Some surprising findings on the involvement of the parietal lobe in human memory. *Neurobiology of Learning and Memory*, 91(2), 155–165. <https://doi.org/10.1016/j.nlm.2008.09.006>

- Papademetris, X., Jackowski, M., Rajeevan, N., DiStasio, M., Okuda, H., Constable, R. T., & Staib, L. (2006). BioImage Suite: An integrated medical image analysis suite: An update. *The Insight Journal*, 1–8. <https://doi.org/10.54294/2g80r4>
- Paparella, I., & Papeo, L. (2022). Chunking by social relationship in working memory. *Visual Cognition*, 30(5), 354–370. <https://doi.org/10.1080/13506285.2022.2064950>
- Papeo, L. (2020). Twos in human visual perception. *Cortex*, 132, 473–478. <https://doi.org/10.1016/j.cortex.2020.06.005>
- Papeo, L., Vettori, S., Serraille, E., Odin, C., Rostami, F., & Hochmann, J. R. (2024). Abstract thematic roles in infants' representation of social events. *Current Biology*, 4294–4300. <https://doi.org/10.1016/j.cub.2024.07.081>
- Park, B., B. Walther, D., & Fukuda, K. (2020). Dynamic Representations in Visual Working Memory. *Journal of Vision*, 20(11), 900. <https://doi.org/10.1167/jov.20.11.900>
- Park, H.-B., Zhang, W., & Hyun, J.-S. (2017). Dissociating models of visual working memory by reaction-time distribution analysis. *Acta Psychologica*, 173, 21–31. <https://doi.org/10.1016/j.actpsy.2016.12.003>
- Pätzold, W., & Liszkowski, U. (2019). Pupillometry reveals communication-induced object expectations in 12- but not 8-month-old infants. *Developmental Science*, 22(6), 1–15. <https://doi.org/10.1111/desc.12832>
- Peng, Y., Javangula, P. R., Lu, H., & Holyoak, K. J. (2017). Behavioral Oscillations in Verification of Relational Role Bindings. *Proceedings of the 40th Annual Conference of the Cognitive Science Society*, 2223–2228.
- Perez, V. B., & Vogel, E. K. (2012). What ERPs Can Tell Us about Working Memory. In *The Oxford Handbook of Event-Related Potential Components* (pp. 361–372). <https://doi.org/10.1093/oxfordhb/9780195374148.013.0180>
- Peters, B., & Kriegeskorte, N. (2021). Capturing the objects of vision with neural networks. *Nature Human Behaviour*, 5(9), 1127–1144. <https://doi.org/10.1038/s41562-021-01194-6>
- Peterson, D. J., Gözenman, F., Arciniega, H., & Berryhill, M. E. (2015). Contralateral delay activity tracks the influence of Gestalt grouping principles on active visual working memory representations. *Attention, Perception, and Psychophysics*, 77(7), 2270–2283. <https://doi.org/10.3758/s13414-015-0929-y>
- Piantadosi, S. T. (2021). The Computational Origin of Representation. *Minds and Machines*, 31(1), 1–58. <https://doi.org/10.1007/s11023-020-09540-9>
- Pion-Tonachini, L., Kreutz-Delgado, K., & Makeig, S. (2019). ICLabel: An automated electroencephalographic independent component classifier, dataset, and website.

- NeuroImage*, 198(May), 181–197. <https://doi.org/10.1016/j.neuroimage.2019.05.026>
- Pitcher, D., & Ungerleider, L. G. (2021). Evidence for a Third Visual Pathway Specialized for Social Perception. *Trends in Cognitive Sciences*, 25(2), 100–110. <https://doi.org/10.1016/j.tics.2020.11.006>
- Pomper, U., & Ansorge, U. (2021). Theta-Rhythmic Oscillation of Working Memory Performance. *Psychological Science*, 32(11), 1801–1810. <https://doi.org/10.1177/09567976211013045>
- Poncet, M., Caramazza, A., & Mazza, V. (2016). Individuation of objects and object parts rely on the same neuronal mechanism. *Scientific Reports*, 6(December), 1–10. <https://doi.org/10.1038/srep38434>
- Prasada, S. (2016). Mechanisms for Thinking about Kinds, Instances of Kinds, and Kinds of Kinds. In *Core Knowledge and Conceptual Change* (pp. 209–224). Oxford University Press. <https://doi.org/10.1093/acprof:oso/9780190467630.003.0012>
- Prasada, S., & Hall, D. G. (2019). Instance-of-object-kind representations. *Cognition*, 189(August 2014), 209–220. <https://doi.org/10.1016/j.cognition.2019.04.002>
- Proskovec, A. L., Heinrichs-Graham, E., & Wilson, T. W. (2019). Load modulates the alpha and beta oscillatory dynamics serving verbal working memory. *NeuroImage*, 184(July 2018), 256–265. <https://doi.org/10.1016/j.neuroimage.2018.09.022>
- Puce, A., & Hämäläinen, M. S. (2017). A review of issues related to data acquisition and analysis in EEG/MEG studies. *Brain Sciences*, 7(6). <https://doi.org/10.3390/brainsci7060058>
- Pykkänen, L., Llinás, R., & Murphy, G. L. (2006). The representation of polysemy: MEG evidence. *Journal of Cognitive Neuroscience*, 18(1), 97–109. <https://doi.org/10.1162/089892906775250003>
- Pykkänen, L., Stringfellow, A., & Marantz, A. (2002). Neuromagnetic evidence for the timing of lexical activation: An MEG component sensitive to phonotactic probability but not to neighborhood density. *Brain and Language*, 81(1–3), 666–678. <https://doi.org/10.1006/brln.2001.2555>
- Pylyshyn, Z. W. (1989). The role of location indexes in spatial perception. *Cognition*, 32, 65–97.
- Pylyshyn, Z. W. (2001). Visual indexes, preconceptual objects, and situated vision. *Cognition*, 80(1–2), 127–158. [https://doi.org/10.1016/S0010-0277\(00\)00156-6](https://doi.org/10.1016/S0010-0277(00)00156-6)
- Qian, J., Zhang, K., Liu, S., & Lei, Q. (2019). The transition from feature to object: Storage unit in visual working memory depends on task difficulty. *Memory & Cognition*, 47(8), 1498–1514. <https://doi.org/10.3758/s13421-019-00956-y>

- Qin, L., & Gao, J. H. (2021). New avenues for functional neuroimaging: ultra-high field MRI and OPM-MEG. *Psychoradiology*, 1(4), 165–171. <https://doi.org/10.1093/psyrad/kkab014>
- Quadflieg, S., Gentile, F., & Rossion, B. (2015). The neural basis of perceiving person interactions. *Cortex*, 70, 5–20. <https://doi.org/10.1016/j.cortex.2014.12.020>
- Quilty-Dunn, J. (2020). Is Iconic Memory Iconic? *Philosophy and Phenomenological Research*, 101(3), 660–682. <https://doi.org/10.1111/phpr.12625>
- Quilty-Dunn, J., Porot, N., & Mandelbaum, E. (2023). The best game in town: The reemergence of the language-of-thought hypothesis across the cognitive sciences. *Behavioral and Brain Sciences*, 46. <https://doi.org/10.1017/S0140525X22002849>
- Rabbitt, L. R., Roberts, D. M., McDonald, C. G., & Peterson, M. S. (2017). Neural activity reveals perceptual grouping in working memory. *International Journal of Psychophysiology*, 113, 40–45. <https://doi.org/10.1016/j.ijpsycho.2017.01.005>
- Rafal, R. (2001). Balint’s syndrome. In M. Behrmann (Ed.), *Handbook of neuropsychology*, 2nd Edn, Vol. 4 (pp. 121–141). Elsevier Science.
- Reisinger, D., Rudinger, R., Ferraro, F., Harman, C., Rawlins, K., & Van Durme, B. (2015). Semantic Proto-Roles. *Transactions of the Association for Computational Linguistics*, 3, 475–488. https://doi.org/10.1162/tacl_a_00152
- Rissman, L., & Lupyan, G. (2021). A Dissociation Between Conceptual Prominence And Explicit Category Learning: Evidence From Agent And Patient Event Roles. *Journal of Experimental Psychology: General*. <https://doi.org/10.1037/xge0001146>
- Rissman, L., & Majid, A. (2019). Thematic roles: Core knowledge or linguistic construct? *Psychonomic Bulletin and Review*, 26(6), 1850–1869. <https://doi.org/10.3758/s13423-019-01634-5>
- Robitaille, N., Grimault, S., & Jolicoeur, P. (2009). Bilateral parietal and contralateral responses during maintenance of unilaterally encoded objects in visual short-term memory: Evidence from magnetoencephalography. *Psychophysiology*, 46(5), 1090–1099. <https://doi.org/10.1111/j.1469-8986.2009.00837.x>
- Robitaille, N., Marois, R., Todd, J., Grimault, S., Cheyne, D., & 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(4), 1334–1345. <https://doi.org/10.1016/j.neuroimage.2010.07.027>
- Rose, N. S., LaRocque, J. J., Riggall, A. C., Gosseries, O., Starrett, M. J., Meyering, E. E., & Postle, B. R. (2016). Reactivation of latent working memories with transcranial magnetic stimulation. *Science*, 354(6316), 1136–1139. <https://doi.org/10.1126/science.aah7011>

- Rouder, J. N., Morey, R. D., Morey, C. C., & Cowan, N. (2011). How to measure working memory capacity in the change detection paradigm. *Psychonomic Bulletin and Review*, 18(2), 324–330. <https://doi.org/10.3758/s13423-011-0055-3>
- Ruchkin, D. S., Johnson, R., Grafman, J., Canoune, H., & Ritter, W. (1992). Distinctions and similarities among working memory processes: an event-related potential study. *Cognitive Brain Research*, 1(1), 53–66. [https://doi.org/10.1016/0926-6410\(92\)90005-C](https://doi.org/10.1016/0926-6410(92)90005-C)
- Salmelin, R. (2007). Clinical neurophysiology of language: The MEG approach. *Clinical Neurophysiology*, 118(2), 237–254. <https://doi.org/10.1016/j.clinph.2006.07.316>
- Sander, T., Jodko-Włodzińska, A., Hartwig, S., Brühl, R., & Middelmann, T. (2020). Optically pumped magnetometers enable a new level of biomagnetic measurements. *Advanced Optical Technologies*, 9(5), 247–251. <https://doi.org/10.1515/aot-2020-0027>
- Sander, T., Preusser, J., Mhaskar, R., Kitching, J., Trahms, L., & Knappe, S. (2012). *Magnetoencephalography with a chip-scale atomic magnetometer*. 3(5), 27167–27172.
- Schneegans, S., & Bays, P. M. (2017). Neural Architecture for Feature Binding in Visual Working Memory. *The Journal of Neuroscience*, 37(14), 3913–3925. <https://doi.org/10.1523/JNEUROSCI.3493-16.2017>
- Schneegans, S., & Bays, P. M. (2019). New perspectives on binding in visual working memory. *British Journal of Psychology*, 110(2), 207–244. <https://doi.org/10.1111/bjop.12345>
- Scholl, B. J. (2009). What Have We Learned about Attention from Multiple-Object Tracking (and Vice Versa)? In *Computation, Cognition, and Pylyshyn* (Issue 2, pp. 49–78). The MIT Press. <https://doi.org/10.7551/mitpress/8135.003.0005>
- Scholl, B. J., & Pylyshyn, Z. W. (1999). Tracking Multiple Items Through Occlusion: Clues to Visual Objecthood. *Cognitive Psychology*, 38(2), 259–290. <https://doi.org/10.1006/cogp.1998.0698>
- Seymour, R. A., Alexander, N., Mellor, S., Neill, G. C. O., Tierney, T. M., Barnes, G. R., & Maguire, E. A. (2021). NeuroImage Using OPMs to measure neural activity in standing , mobile participants. *NeuroImage*, 244(August), 118604. <https://doi.org/10.1016/j.neuroimage.2021.118604>
- Seymour, R. A., Alexander, N., Mellor, S., O'Neill, G. C., Tierney, T. M., Barnes, G. R., & Maguire, E. A. (2022). Interference suppression techniques for OPM-based MEG: Opportunities and challenges. *NeuroImage*, 247(December 2021), 118834. <https://doi.org/10.1016/j.neuroimage.2021.118834>
- Shastri, L. (1999). Advances in SHRUTI—A Neurally Motivated Model of Relational Knowledge Representation and Rapid Inference Using Temporal Synchrony. *Applied Intelligence*, 11(1), 79–108. <https://doi.org/10.1023/A:1008380614985>

- Shastri, L. (2002). Episodic memory and cortico–hippocampal interactions. *Trends in Cognitive Sciences*, 6(4), 162–168. [https://doi.org/10.1016/S1364-6613\(02\)01868-5](https://doi.org/10.1016/S1364-6613(02)01868-5)
- Shastri, L. (2007). *SHRUTI: A Neurally Motivated Architecture for Rapid, Scalable Inference* (pp. 183–203). https://doi.org/10.1007/978-3-540-73954-8_8
- Shastri, L., & Ajanagadde, V. (1993). From simple associations to systematic reasoning: A connectionist representation of rules, variables and dynamic bindings using temporal synchrony. *Behavioral and Brain Sciences*, 16(3), 417–494. <https://doi.org/10.1017/s0140525x00030910>
- Sheahan, H., Luyckx, F., Nelli, S., Teupe, C., & Summerfield, C. (2020). Neural state space alignment for magnitude generalisation in humans and recurrent networks. *BioRxiv*, 1–13. <https://doi.org/10.1101/2020.07.22.215541>
- Shen, M., Chen, J., Yang, X., Dong, H., Chen, H., & Zhou, J. (2021). The storage mechanism of dynamic relations in visual working memory. *Cognition*, 209(January 2020), 104571. <https://doi.org/10.1016/j.cognition.2020.104571>
- Shen, M., Gao, Z., Ding, X., Zhou, B., & Huang, X. (2014). Holding biological motion information in working memory. *Journal of Experimental Psychology: Human Perception and Performance*, 40(4), 1332–1345. <https://doi.org/10.1037/a0036839>
- Sigurdardottir, H. M., Ólafsdóttir, I. M., & Devillez, H. (2021). Words as visual objects: Neural and behavioral evidence for high-level visual impairments in dyslexia. *Brain Sciences*, 11(11). <https://doi.org/10.3390/brainsci11111427>
- Simal, A., & Jolicoeur, P. (2020). Scanning acoustic short-term memory: Evidence for two subsystems with different time-course and memory strength. *International Journal of Psychophysiology*, 155(August 2019), 105–117. <https://doi.org/10.1016/j.ijpsycho.2020.06.004>
- Singer, W. (1993). Synchronization of Cortical Activity and its Putative Role in Information Processing and Learning. *Annual Review of Physiology*, 55(1), 349–374. <https://doi.org/10.1146/annurev.ph.55.030193.002025>
- Smolensky, P. (1988). On the proper treatment of connectionism. *Behavioral and Brain Sciences*, 11(1), 1–23. <https://doi.org/10.1017/S0140525X00052432>
- Song, J. H., & Jiang, Y. (2006). Visual working memory for simple and complex features: An fMRI study. *NeuroImage*, 30(3), 963–972. <https://doi.org/10.1016/j.neuroimage.2005.10.006>
- Spelke, E. S. (1990). Principles of object perception. *Cognitive Science*, 14(1), 29–56. [https://doi.org/10.1016/0364-0213\(90\)90025-R](https://doi.org/10.1016/0364-0213(90)90025-R)

- Spitzer, B., Waschke, L., & Summerfield, C. (2017). Selective overweighting of larger magnitudes during noisy numerical comparison. *Nature Human Behaviour*, 1(8), 1–8. <https://doi.org/10.1038/s41562-017-0145>
- Spriet, C., Abassi, E., Hochmann, J.-R., & Papeo, L. (2022). Visual object categorization in infancy. *Proceedings of the National Academy of Sciences*, 119(8). <https://doi.org/10.1073/pnas.2105866119>
- Sprouse, J., Kucerova, I., Park, J., Cerrone, P., Schrum, N., & Lau, E. (2021). *Sustained anterior negativities and movement dependencies in syntax*. Department of Linguistics, University of Connecticut.
- Stahl, A. E., & Feigenson, L. (2014). Social Knowledge Facilitates Chunking in Infancy. *Child Development*, 85(4), 1477–1490. <https://doi.org/10.1111/cdev.12217>
- Stockall, L., Stringfellow, A., & Marantz, A. (2004). The precise time course of lexical activation: MEG measurements of the effects of frequency, probability, and density in lexical decision. *Brain and Language*, 90(1–3), 88–94. [https://doi.org/10.1016/S0093-934X\(03\)00422-X](https://doi.org/10.1016/S0093-934X(03)00422-X)
- Stokes, M. G. (2015). “Activity-silent” working memory in prefrontal cortex: A dynamic coding framework. *Trends in Cognitive Sciences*, 19(7), 394–405. <https://doi.org/10.1016/j.tics.2015.05.004>
- Storm, R. W., & Pylyshyn, Z. W. (1988). Tracking multiple independent targets: Evidence for a parallel tracking mechanism. *Spatial Vision*, 3(3), 179–197.
- Strickland, B. (2017). Language Reflects “Core” Cognition: A New Theory About the Origin of Cross-Linguistic Regularities. *Cognitive Science*, 41(1), 70–101. <https://doi.org/10.1111/cogs.12332>
- Summerfield, C., Luyckx, F., & Sheahan, H. (2020). Structure learning and the posterior parietal cortex. *Progress in Neurobiology*, 184(July). <https://doi.org/10.1016/j.pneurobio.2019.101717>
- Surian, L., & Caldi, S. (2010). Infants’ individuation of agents and inert objects. *Developmental Science*, 13(1), 143–150. <https://doi.org/10.1111/j.1467-7687.2009.00873.x>
- Swan, G., & Wyble, B. (2014). The binding pool: A model of shared neural resources for distinct items in visual working memory. *Attention, Perception, and Psychophysics*, 76(7), 2136–2157. <https://doi.org/10.3758/s13414-014-0633-3>
- Tan, G., Gai, J., Li, F., Huang, Q., Yu, W., Huang, Y., Zhang, G. Y., Zhao, X., Lin, Q., & Hu, Z. (2025). A Brain-Computer Interface System based on Magnetoencephalography with Optically Pumped Magnetometers. *IEEE Transactions on Instrumentation and Measurement*, PP, 1. <https://doi.org/10.1109/TIM.2024.3522701>

- Tanner, D., Morgan-short, K., & Luck, S. J. (2015). *How inappropriate high-pass filters can produce artifactual effects and incorrect conclusions in ERP studies of language and cognition*. 52, 997–1009. <https://doi.org/10.1111/psyp.12437>
- Tarkiainen, A., Helenius, P., & Salmelin, R. (2003). Category-specific occipitotemporal activation during face perception in dyslexic individuals: An MEG study. *NeuroImage*, 19(3), 1194–1204. [https://doi.org/10.1016/S1053-8119\(03\)00161-7](https://doi.org/10.1016/S1053-8119(03)00161-7)
- Taulu, S., & Simola, J. (2006). Spatiotemporal signal space separation method for rejecting nearby interference in MEG measurements. *Physics in Medicine and Biology*, 51(7), 1759–1768. <https://doi.org/10.1088/0031-9155/51/7/008>
- Teichmann, L., Moerel, D., Rich, A. N., & Baker, C. I. (2022). The nature of neural object representations during dynamic occlusion. *Cortex*, 153, 66–86. <https://doi.org/10.1016/j.cortex.2022.04.009>
- ten Oever, S., De Weerd, P., & Sack, A. T. (2020). Phase-dependent amplification of working memory content and performance. *Nature Communications*, 11(1), 1832. <https://doi.org/10.1038/s41467-020-15629-7>
- Teng, C., & Kravitz, D. J. (2019). Visual working memory directly alters perception. *Nature Human Behaviour*, 3(8), 827–836. <https://doi.org/10.1038/s41562-019-0640-4>
- Thompson, C. K., Bonakdarpour, B., Fix, S. C., Blumenfeld, H. K., Parrish, T. B., Gitelman, D. R., & Mesulam, M. M. (2007). Neural correlates of verb argument structure processing. *Journal of Cognitive Neuroscience*, 19(11), 1753–1767. <https://doi.org/10.1162/jocn.2007.19.11.1753>
- 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>
- Tian, Z., Chen, J., Zhang, C., Min, B., Xu, B., & Wang, L. (2024). Mental programming of spatial sequences in working memory in the macaque frontal cortex. *Science*, 385(6716). <https://doi.org/10.1126/science.adp6091>
- Tierney, T. M., Alexander, N., Mellor, S., Holmes, N., Seymour, R., O'Neill, G. C., Maguire, E. A., & Barnes, G. R. (2021). Modelling optically pumped magnetometer interference in MEG as a spatially homogeneous magnetic field. *NeuroImage*, 244(July), 118484. <https://doi.org/10.1016/j.neuroimage.2021.118484>
- Todd, J. J., & Marois, R. (2004). Capacity limit of visual short-term memory in human posterior parietal cortex. *Nature*, 428, 751–754.
- Todd, J. J., & Marois, R. (2005). Posterior parietal cortex activity predicts individual differences

- in visual short-term memory capacity. *Cognitive, Affective, & Behavioral Neuroscience*, 5(2), 144–155. <https://doi.org/10.3758/CABN.5.2.144>
- Treisman, A. (1998). Feature binding, attention and object perception. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 353(1373), 1295–1306. <https://doi.org/10.1098/rstb.1998.0284>
- Tsigka, S., Papadelis, C., Braun, C., & Miceli, G. (2014). Distinguishable neural correlates of verbs and nouns: A MEG study on homonyms. *Neuropsychologia*, 54(1), 87–97. <https://doi.org/10.1016/j.neuropsychologia.2013.12.018>
- Ünal, E., Ji, Y., & Papafragou, A. (2019). From Event Representation to Linguistic Meaning. *Topics in Cognitive Science*, tops.12475. <https://doi.org/10.1111/tops.12475>
- van Bree, S. (2024). Why the neural ingredients for a language of thought are not like spatial cells (commentary on Kazanina & Poeppel, 2023). *European Journal of Neuroscience*, 59(10), 2552–2555. <https://doi.org/10.1111/ejn.16329>
- van den Berg, R., Awh, E., & Ma, W. J. (2014). Factorial comparison of working memory models. *Psychological Review*, 121(1), 124–149. <https://doi.org/10.1037/a0035234>
- van der Werf, O. J., Ten Oever, S., Schuhmann, T., & Sack, A. T. (2021). No evidence of rhythmic visuospatial attention at cued locations in a spatial cuing paradigm, regardless of their behavioural relevance. *European Journal of Neuroscience*, September 2020, 1–17. <https://doi.org/10.1111/ejn.15353>
- van Doorn, J., van den Bergh, D., Böhm, U., Dablander, F., Derks, K., Draws, T., Etz, A., Evans, N. J., Gronau, Q. F., Haaf, J. M., Hinne, M., Kucharský, Š., Ly, A., Marsman, M., Matzke, D., Gupta, A. R. K. N., Sarafoglou, A., Stefan, A., Voelkel, J. G., & Wagenmakers, E. J. (2021). The JASP guidelines for conducting and reporting a Bayesian analysis. *Psychonomic Bulletin and Review*, 28(3), 813–826. <https://doi.org/10.3758/s13423-020-01798-5>
- van Driel, J., Olivers, C. N. L., & Fahrenfort, J. J. (2021). High-pass filtering artifacts in multivariate classification of neural time series data. *Journal of Neuroscience Methods*, 352(November 2019), 109080. <https://doi.org/10.1016/j.jneumeth.2021.109080>
- Vanrie, J., & Verfaillie, K. (2004). Perception of biological motion: A stimulus set of human point-light actions. *Behavior Research Methods, Instruments, and Computers*, 36(4), 625–629. <https://doi.org/10.3758/BF03206542>
- Vergauwe, E., & Cowan, N. (2015). Working memory units are all in your head: Factors that influence whether features or objects are the favored units. *Journal of Experimental Psychology: Learning Memory and Cognition*, 41(5), 1404–1416. <https://doi.org/10.1037/xlm0000108>

- Vogel, E. K., & Machizawa, M. G. (2004). Neural activity predicts individual differences in visual working memory capacity. *Nature*, 428(6984), 748–751. <https://doi.org/10.1038/nature02447>
- Walbrin, J., & Koldewyn, K. (2019). Dyadic interaction processing in the posterior temporal cortex. *NeuroImage*, 198(January), 296–302. <https://doi.org/10.1016/j.neuroimage.2019.05.027>
- Walsh, V. (2003). A theory of magnitude: Common cortical metrics of time, space and quantity. *Trends in Cognitive Sciences*, 7(11), 483–488. <https://doi.org/10.1016/j.tics.2003.09.002>
- Wang, B., Cao, X., Theeuwes, J., Olivers, C. N. L., & Wang, Z. (2017). Separate capacities for storing different features in visual working memory. *Journal of Experimental Psychology: Learning Memory and Cognition*, 43(2), 226–236. <https://doi.org/10.1037/xlm0000295>
- Wang, J., Cherkassky, V. L., Yang, Y., Chang, K. min K., Vargas, R., Diana, N., & Just, M. A. (2016). Identifying thematic roles from neural representations measured by functional magnetic resonance imaging. *Cognitive Neuropsychology*, 33(3–4), 257–264. <https://doi.org/10.1080/02643294.2016.1182480>
- Wang, R., Janini, D., & Konkle, T. (2022). Mid-level Feature Differences Support Early Animacy and Object Size Distinctions: Evidence from Electroencephalography Decoding. *Journal of Cognitive Neuroscience*, 34(9), 1670–1680. https://doi.org/10.1162/jocn_a_01883
- Wang, X. J. (2021). 50 Years of Mnemonic Persistent Activity: Quo Vadis? *Trends in Neurosciences*, 44(11), 888–902. <https://doi.org/10.1016/j.tins.2021.09.001>
- 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>
- Wellwood, A., & Hunter, T. (2023). Linguistic meanings in mind. *Behavioral and Brain Sciences*, 46, e289. <https://doi.org/10.1017/S0140525X23001887>
- Wheeler, M. E., & Treisman, A. M. (2002). Binding in short-term visual memory. *Journal of Experimental Psychology: General*, 131(1), 48–64. <https://doi.org/10.1037/0096-3445.131.1.48>
- Whitney, D., & Yamanashi Leib, A. (2018). Ensemble Perception. *Annual Review of Psychology*, 69, 105–129. <https://doi.org/10.1146/annurev-psych-010416-044232>
- Williams, Adina, Reddigari, S., & Pykkänen, L. (2017). Early sensitivity of left perisylvian cortex to relationality in nouns and verbs. *Neuropsychologia*, 100(November 2016), 131–143. <https://doi.org/10.1016/j.neuropsychologia.2017.04.029>

- Williams, Alexander. (2015). *Arguments in Syntax and Semantics*. Cambridge University Press.
- Williams, Alexander. (2021). Events in Semantics. *The Cambridge Handbook of the Philosophy of Language*, 366–386. <https://doi.org/10.1017/9781108698283.021>
- Wilson, K. E., Adamo, M., Barense, M. D., & Ferber, S. (2012). To bind or not to bind: Addressing the question of object representation in visual short-term memory. *Journal of Vision*, 12(8), 1–16. <https://doi.org/10.1167/12.8.14>
- Wilson, V. A. D., Zuberbühler, K., & Bickel, B. (2022). The evolutionary origins of syntax: Event cognition in nonhuman primates. *Science Advances*, 1–13.
- Wolff, M. J., Ding, J., Myers, N. E., & Stokes, M. G. (2015). Revealing hidden states in visual working memory using electroencephalography. *Frontiers in Systems Neuroscience*, 9(september), 1–12. <https://doi.org/10.3389/fnsys.2015.00123>
- Wolff, M. J., Jochim, J., Akyürek, E. G., Buschman, T. J., & Stokes, M. G. (2020). Drifting codes within a stable coding scheme for working memory. *PLOS Biology*, 18(3), e3000625. <https://doi.org/10.1371/journal.pbio.3000625>
- Wolff, M. J., Jochim, J., Akyürek, E. G., & Stokes, M. G. (2017). Dynamic hidden states underlying working-memory-guided behavior. *Nature Neuroscience*, 20(6), 864–871. <https://doi.org/10.1038/nn.4546>
- Wolff, M. J., Kandemir, G., Stokes, M. G., & Akyürek, E. G. (2019). Impulse responses reveal unimodal and bimodal access to visual and auditory working memory. *BioRxiv*. <https://doi.org/10.1101/623835>
- Wood, J. N. (2007). Visual Working Memory for Observed Actions. *Journal of Experimental Psychology: General*, 136(4), 639–652. <https://doi.org/10.1037/0096-3445.136.4.639>
- Wood, J. N. (2011). “A core knowledge architecture of visual working memory”: Correction to Wood (2011). *Journal of Experimental Psychology: Human Perception and Performance*, 37(3), 633–633. <https://doi.org/10.1037/a0023751>
- Woodman, G. F., & Vogel, E. K. (2008). Selective storage and maintenance of an object’s features in visual working memory. *Psychonomic Bulletin and Review*, 15(1), 223–229. <https://doi.org/10.3758/PBR.15.1.223>
- 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
- Wu, H., Liang, X., Wang, R., Ma, Y., Gao, Y., & Ning, X. (2024). A Multivariate analysis on evoked components of Chinese semantic congruity: an OP-MEG study with EEG. *Cerebral*

Cortex, 34(4). <https://doi.org/10.1093/cercor/bhae108>

- Wurm, M. F., Ariani, G., Greenlee, M. W., & Lingnau, A. (2016). Decoding Concrete and Abstract Action Representations During Explicit and Implicit Conceptual Processing. *Cerebral Cortex*, 26(8), 3390–3401. <https://doi.org/10.1093/cercor/bhv169>
- Wurm, M. F., & Caramazza, A. (2019). Distinct roles of temporal and frontoparietal cortex in representing actions across vision and language. *Nature Communications*, 10(1), 1–10. <https://doi.org/10.1038/s41467-018-08084-y>
- Wurm, M. F., Caramazza, A., & Lingnau, A. (2017). Action categories in lateral occipitotemporal cortex are organized along sociality and transitivity. *Journal of Neuroscience*, 37(3), 562–575. <https://doi.org/10.1523/JNEUROSCI.1717-16.2016>
- Wurm, M. F., & Lingnau, A. (2015). Decoding actions at different levels of abstraction. *Journal of Neuroscience*, 35(20), 7727–7735. <https://doi.org/10.1523/JNEUROSCI.0188-15.2015>
- Wurm, M. F., Tagliabue, C. F., & Mazza, V. (2021). Decoding location-specific and location-invariant stages of numerosity processing in subitizing. *European Journal of Neuroscience*, 54(3), 4971–4984. <https://doi.org/10.1111/ejn.15352>
- Wutz, A., Zazio, A., & Weisz, N. (2020). Oscillatory bursts in parietal cortex reflect dynamic attention between multiple objects and ensembles. *Journal of Neuroscience*, 40(36), 6927–6937. <https://doi.org/10.1523/JNEUROSCI.0231-20.2020>
- Xia, Q., & Peng, G. (2022). The roles of object and action, and concreteness and imageability, in the distinction between nouns and verbs: An ERP study on monosyllabic words in Chinese. *Journal of Neurolinguistics*, 61(August 2021), 101026. <https://doi.org/10.1016/j.jneuroling.2021.101026>
- Xia, Q., Wang, L., & Peng, G. (2016). Nouns and verbs in Chinese are processed differently: Evidence from an ERP study on monosyllabic and disyllabic word processing. *Journal of Neurolinguistics*, 40, 66–78. <https://doi.org/10.1016/j.jneuroling.2016.06.002>
- Xie, W., Campbell, S., & Zhang, W. (2020). Working memory capacity predicts individual differences in social-distancing compliance during the COVID-19 pandemic in the United States. *Proceedings of the National Academy of Sciences of the United States of America*, 117(30), 17667–17674. <https://doi.org/10.1073/pnas.2008868117>
- Xie, W., Wittig, J. H., Chapeton, J. I., El-Kalliny, M., Jackson, S. N., Inati, S. K., & Zaghloul, K. A. (2024). Neuronal sequences in population bursts encode information in human cortex. *Nature*, 635(8040), 935–942. <https://doi.org/10.1038/s41586-024-08075-8>
- Xie, W., & Zhang, W. (2017a). Discrete item-based and continuous configural representations in visual short-term memory. *Visual Cognition*, 25(1–3), 21–33. <https://doi.org/10.1080/13506285.2017.1339157>

- Xie, W., & Zhang, W. (2017b). Dissociations of the number and precision of visual short-term memory representations in change detection. *Memory and Cognition*, 45(8), 1423–1437. <https://doi.org/10.3758/s13421-017-0739-7>
- Xie, W., & Zhang, W. (2023). Effortfulness of Visual Working Memory: Gauged by Physical Exertion. *Journal of Experimental Psychology: General*, 152(7), 2074–2093. <https://doi.org/10.1037/xge0001391>
- Xu, F. (1999). Object individuation and object identity in infancy: The role of spatiotemporal information, object property information, and language. *Acta Psychologica*, 102(2–3), 113–136. [https://doi.org/10.1016/S0001-6918\(99\)00029-3](https://doi.org/10.1016/S0001-6918(99)00029-3)
- Xu, F., & Carey, S. (1996). Infants' Metaphysics: The Case of Numerical Identity. *Cognitive Psychology*, 30(2), 111–153. <https://doi.org/10.1006/cogp.1996.0005>
- Xu, Y. (2007). The Role of the Superior Intraparietal Sulcus in Supporting Visual Short-Term Memory for Multifeature Objects. *Journal of Neuroscience*, 27(43), 11676–11686. <https://doi.org/10.1523/JNEUROSCI.3545-07.2007>
- 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. (2017). Reevaluating the Sensory Account of Visual Working Memory Storage. *Trends in Cognitive Sciences*, 21(10), 794–815. <https://doi.org/10.1016/j.tics.2017.06.013>
- Xu, Y. (2018). A tale of two visual systems: Invariant and adaptive visual information representations in the primate brain. *Annual Review of Vision Science*, 4, 311–336. <https://doi.org/10.1146/annurev-vision-091517-033954>
- Xu, Y. (2021). Towards a better understanding of information storage in visual working memory. *Visual Cognition*, 29(7), 437–445. <https://doi.org/10.1080/13506285.2021.1946230>
- Xu, Y. (2023). Parietal-driven visual working memory representation in occipito-temporal cortex. *Current Biology*, 33(20), 4516–4523.e5. <https://doi.org/10.1016/j.cub.2023.08.080>
- 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>
- Xu, Y., & Chun, M. M. (2007). Visual grouping in human parietal cortex. *Proceedings of the National Academy of Sciences of the United States of America*, 104(47), 18766–18771. <https://doi.org/10.1073/pnas.0705618104>
- Xu, Y., & Chun, M. M. (2009). Selecting and perceiving multiple visual objects. *Trends in Cognitive Sciences*, 13(4), 167–174. <https://doi.org/10.1016/j.tics.2009.01.008>

- Xun, E., Rao, G., Xiao, X., & Zang, J. (2016). Dashuju beijing xia BCC yuliaoku de yanzhi [The construction of the BCC corpus in the age of Big Data]. *Yuliaoku Yuyanxue [Corpora Linguistics]*, 3(1), 93–109.
- Yang, H., He, C., Han, Z., & Bi, Y. (2020). Domain-specific functional coupling between dorsal and ventral systems during action perception. *Scientific Reports*, 10(1), 1–14. <https://doi.org/10.1038/s41598-020-78276-4>
- Yano, M., & Koizumi, M. (2018). Processing of non-canonical word orders in (in)felicitous contexts: evidence from event-related brain potentials. *Language, Cognition and Neuroscience*, 33(10), 1340–1354. <https://doi.org/10.1080/23273798.2018.1489066>
- Yano, M., & Koizumi, M. (2021). The role of discourse in long-distance dependency formation. *Language, Cognition and Neuroscience*, 36(6), 711–729. <https://doi.org/10.1080/23273798.2021.1883694>
- Yeung, N., & Sanfey, A. G. (2004). Independent coding of reward magnitude and valence in the human brain. *Journal of Neuroscience*, 24(28), 6258–6264. <https://doi.org/10.1523/JNEUROSCI.4537-03.2004>
- Yu, X. (2025). Mapping the Neural Taxonomy of Mental Objects in Moment-to-Moment Cognition. *Journal of Cognitive Neuroscience*, 1–15. 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. (2024). Same set of visual pointers/indexicals for biological and non-biological objects in working memory. *PsyArXiv*. <https://doi.org/10.31234/osf.io/pd3n7>
- 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., Li, J., Zhu, H., Tian, X., & Lau, E. (2024). Electrophysiological hallmarks for event relations and event roles in working memory. *Frontiers in Neuroscience*, 17(January), 1–22. <https://doi.org/10.3389/fnins.2023.1282869>
- Yu, X., Mancha, S., Tian, X., & Lau, E. (2024). Shared neural computations for syntactic and morphological structures: evidence from Mandarin Chinese. *BioRxiv*. <https://doi.org/10.1101/2024.01.31.578104>
- Yu, X., Raghavan, P., Niu, X., Bao, Y., Pöppel, E., & Xie, W. (2025). Universal Beauty of Abstract Art Predicted by Deep Neural Networks. *PsyArXiv*. https://doi.org/10.31234/osf.io/uphbj_v1
- 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*. <https://doi.org/10.1037/xhp0001347>
- Zadbood, A., Chen, J., Leong, Y. C., Norman, K. A., & Hasson, U. (2017). How We Transmit Memories to Other Brains: Constructing Shared Neural Representations Via Communication. *Cerebral Cortex*, 27(10), 4988–5000. <https://doi.org/10.1093/cercor/bhx202>
- Zeng, Y., Sauseng, P., & Alamia, A. (2024). Alpha Traveling Waves during Working Memory: Disentangling Bottom-Up Gating and Top-Down Gain Control. *The Journal of Neuroscience*, 44(50), e0532242024. <https://doi.org/10.1523/JNEUROSCI.0532-24.2024>
- Zhan, W., Guo, R., Chang, B., Chen, Y., & Chen, L. (2019). Beijing Daxue CCL Yuliaoku de Yanzhi [The building of the CCL corpus: its design and implementation]. *Yuliaoku Yuyanxue [Corpus Linguistics]*, 6(1), 71–86.
- Zhang, G., Garrett, D. R., & Luck, S. J. (2024a). *Optimal filters for ERP research I: A general approach for selecting filter settings*. November 2023. <https://doi.org/10.1111/psyp.14531>
- Zhang, G., Garrett, D. R., & Luck, S. J. (2024b). *Optimal filters for ERP research II: Recommended settings for seven common ERP components*. January, 1–24. <https://doi.org/10.1111/psyp.14530>
- Zhang, J., Ye, C., Sun, H.-J., Zhou, J., Liang, T., Li, Y., & Liu, Q. (2021). The passive state: A protective mechanism for information in working memory tasks. *Journal of Experimental Psychology: Learning, Memory, and Cognition*. <https://doi.org/10.1037/xlm0001092>
- 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>
- Zhang, X., Chen, C., Zhang, M., Ma, C., Zhang, Y., Wang, H., Guo, Q., Hu, T., Liu, Z., Chang, Y., Hu, K., & Yang, X. (2020). Detection and analysis of MEG signals in occipital region with double-channel OPM sensors. *Journal of Neuroscience Methods*, 346(April), 108948. <https://doi.org/10.1016/j.jneumeth.2020.108948>
- Zhao, C., Hu, L., Wei, L., Wang, C., Li, X., Hu, B., & Zhang, X. (2020). How Do Humans Perform in Multiple Object Tracking With Unstable Features. *Frontiers in Psychology*, 11(July). <https://doi.org/10.3389/fpsyg.2020.01940>
- Zhou, C., Han, M., Liang, Q., Hu, Y. F., & Kuai, S. G. (2019). A social interaction field model accurately identifies static and dynamic social groupings. *Nature Human Behaviour*, 3(8), 847–855. <https://doi.org/10.1038/s41562-019-0618-2>
- Zhou, H., Su, C., Wu, J., Li, J., Lu, X., Gong, L., & Geng, F. (2022). *A domain-general frontoparietal network interacts with domain-preferential intermediate pathways to support*

working memory task. 1–14.

Zhou, K., Luo, H., Zhou, T., Zhuo, Y., & Chen, L. (2010). Topological change disturbs object continuity in attentive tracking. *Proceedings of the National Academy of Sciences of the United States of America*, 107(50), 21920–21924. <https://doi.org/10.1073/pnas.1010919108>

Zhou, L., & Thomas, R. D. (2015). Principal component analysis of the memory load effect in a change detection task. *Vision Research*, 110(Part A), 1–6.
<https://doi.org/10.1016/j.visres.2015.01.028>

Zhu, S. D., Zhang, L. A., & Von Der Heydt, R. (2020). Searching for object pointers in the visual cortex. *Journal of Neurophysiology*, 123(5), 1979–1994.
<https://doi.org/10.1152/JN.00112.2020>

Zylberberg, A., Dehaene, S., Roelfsema, P. R., & Sigman, M. (2011). The human Turing machine: A neural framework for mental programs. *Trends in Cognitive Sciences*, 15(7), 293–300. <https://doi.org/10.1016/j.tics.2011.05.007>