Received: 16 March 2023 | Revised: 21 March 2024 | Accepted: 6 June 2024

DOI: 10.1111/psyp.14636

PSYCHOPHYSIOLOGY SPR WILEY

ORIGINAL ARTICLE

Decoding load or selection in visuospatial working memory?

Miriam Tortajada[1] | **Johannes J. Fahrenfort**[2] | **Alejandro Sandoval-Lentisco**[1] | **Víctor Martínez-Pérez**[1] | **Lucía B. Palmero**[1] | **Alejandro Castillo**[1] | **Luis J. Fuentes**[1] | **Guillermo Campoy**[1] | **Christian N. L. Olivers**[2]

[1]Departamento de Psicología Básica y Metodología, Facultad de Psicología, Universidad de Murcia, Murcia, Spain

[2]Department of Experimental and Applied Psychology, Institute for Brain and Behavior Amsterdam, Vrije Universiteit Amsterdam, Amsterdam, The Netherlands

Correspondence
Miriam Tortajada, Departamento de Psicología Básica y Metodología, Facultad de Psicología, Universidad de Murcia, Murcia, Spain.
Email: miriam.tortajada@um.es

Christian N. L. Olivers, Department of Experimental and Applied Psychology, Institute for Brain and Behavior Amsterdam, Vrije Universiteit Amsterdam, Amsterdam, The Netherlands.
Email: c.n.l.olivers@vu.nl

Funding information
Ministerio de Ciencia e Innovación, Grant/Award Number: FPU17/00427, FPU18/00288, FPU19/06016 and FPU19/06017; Agencia Estatal de Investigación, Grant/Award Number: PID2021-125408NB-I00; Nederlandse Organisatie voor Wetenschappelijk Onderzoek, Grant/Award Number: 453-16-002

Abstract

Flexible updating of information in Visual Working Memory (VWM) is crucial to deal with its limited capacity. Previous research has shown that the removal of no longer relevant information takes some time to complete. Here, we sought to study the time course of such removal by tracking the accompanying drop in load through behavioral and neurophysiological measures. In the first experimental session, participants completed a visuospatial retro-cue task in which the Cue-Target Interval (CTI) was manipulated. The performance revealed that it takes about half a second to make full use of the retro-cue. In a second session, we sought to study the dynamics of load-related electroencephalographic (EEG) signals to track the removal of information. We applied Multivariate Pattern Analysis (MVPA) to EEG data from the same task. Right after encoding, results replicated previous research using MVPA to decode load. However, especially after the retro-cue, results suggested that classifiers were mainly sensitive to a selection component, and not so much to load per se. Additionally, visual cue variations, as well as eye movements that accompany load manipulations can also contribute to decoding. These findings advise caution when using MVPA to decode VWM load, as classifiers may be sensitive to confounding operations.

KEYWORDS

attentional selection, working memory load, visual working memory, multivariate pattern analysis, electroencephalography

1 | INTRODUCTION

Visual working memory (VWM) is the ability to maintain relevant visual information for an ongoing task. The amount of information that we can keep active in VWM is generally thought to be limited to about three or four items (Cowan, 2010; Vogel et al., 2001), but even below this capacity limit, performance tends to decline as load increases. Importantly, retrospectively cueing an item within VWM as relevant can at least partially save it from such detrimental effects, as compared to information that is not cued (Astle et al., 2012; Günseli et al., 2015; Kuo et al., 2012; Shepherdson et al., 2018; Souza & Oberauer, 2016; van Moorselaar et al., 2015). While there is no clear consensus yet about the exact underlying mechanisms of these retro-cueing benefits, most accounts assume some form of attentional selection process within VWM that increases the robustness of the selected item against decay or interference (Lepsien & Nobre, 2006; Myers et al., 2017; Souza et al., 2014; Souza & Oberauer, 2016). Here we were interested in the dynamics of this process: How long does it take for cue-induced benefits to occur in VWM? And can we then track its underlying mechanisms using electrophysiological measures (specifically EEG)?

One way to explain retro-cueing benefits is that the cue typically allows for irrelevant items to be removed from WM, thus reducing any interference such items may have on, and freeing memory capacity for, the target information (Lewis-Peacock et al., 2018; Makovski, 2012; Williams et al., 2013). In other words, cueing part of the information in VWM as relevant (and as a consequence the rest as irrelevant) effectively allows for a reduction of memory *load*, where load refers to the number of items to be retained. The aim of the current study was to track this load reduction across time. To this end, participants completed a visuospatial probe recognition task in which they were asked to remember a number of colored disks, one of which would be probed at the end of the trial. There were three main conditions (see Figure 1). In the 2- and 4-load baseline conditions, respectively, two and four colored items had to be maintained during the entire trial, as was signaled by a pre-cue appearing prior to the stimulus display. These two conditions served as a comparison for the third condition, in which a retro-cue was introduced during the maintenance of the items in VWM, and which signaled which half of the information would be relevant – as the other half would not be tested. We refer to this as the 4/2-load condition because four items were to be initially encoded and afterwards two remained relevant and two could be dropped. The 4/2-load condition was thus meant to induce removal of half of the information from VWM.

The study consisted of two components. First, in a behavioral version of the task, we systematically varied what can be referred to as the Cue-Test Interval (CTI), which is the interval between the retro-cue and the memory probe.

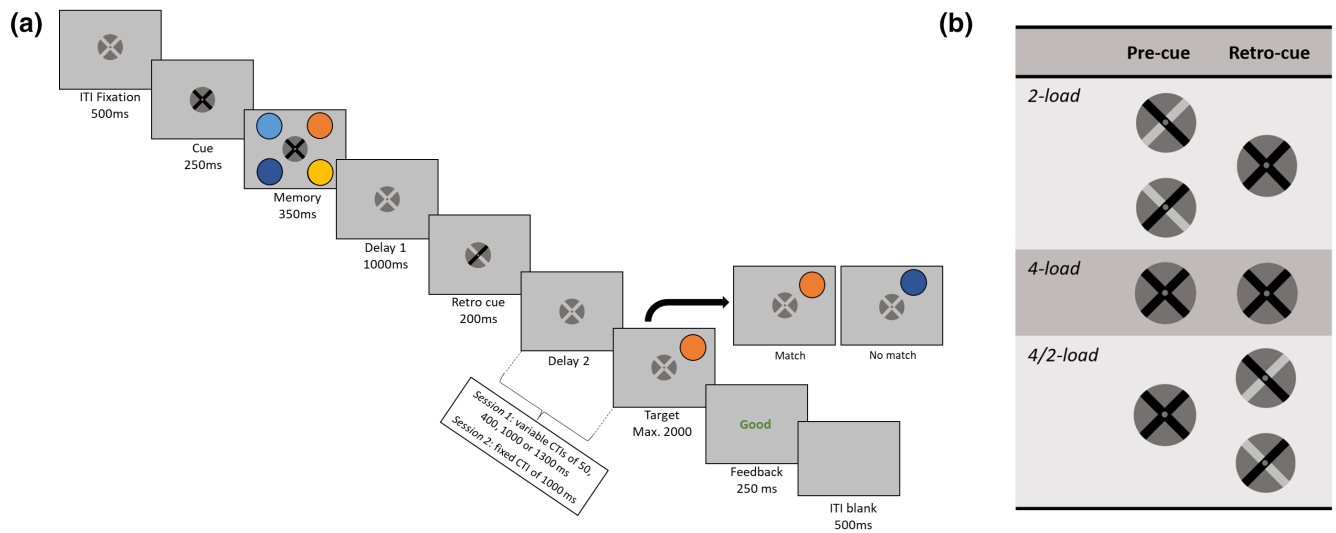

FIGURE 1 (a) Task procedure for a 4/2-load trial. The task was a visuospatial working memory task that required delayed colored item recognition. First, participants saw a pre-cue, followed by four colored disks in four positions. The pre-cue pointed to the disks that had to be encoded. In this present example, all disks had to be encoded and maintained in memory. In the case of the 2-load condition, the uncued disks could be ignored. Then, after delay 1 (1000 ms) a retro-cue signaled the disks that remained relevant. This retro-cue was irrelevant for the 2- and 4-load conditions, but was informative for the 4/2-load condition. After a second delay (variable in the first session; fixed to 1000 ms in second session), a target disk appeared in one of the relevant locations and participants had to press a button reporting whether it matched or not the retained items. intetrial interval (ITI), cue-target interval (CTI). (b) Summary of possible cue combinations for each trial type.

With time after cue, we would then expect performance to improve, with performance initially being comparable to high-load baseline, while with the removal of items from memory it should become comparable to low-load baseline levels. Earlier work by Oberauer (2018) using this procedure showed that indeed, for short CTIs, cued performance still behaved like the high-load condition, but it resembled low-load performance as CTI increased. These and other results have suggested that it takes at least 500 to 1000 ms for the retro-cue to have full effect (see Souza & Oberauer, 2016, for a review). However, so far, these studies have used verbal stimuli. There have been studies that have looked at the time course of retro-cueing effects in VWM (Gressmann & Janczyk, 2016; Pertzov et al., 2013; Shepherdson et al., 2018; Souza et al., 2014; van Moorselaar et al., 2015), but these did not assess how performance transitions from resembling high to resembling low-load baselines, for example, because such conditions were not included. Here we were specifically interested in how selection within visual working memory reduces load across time.

Second, we sought to uncover the electrophysiological counterpart of these dynamics, using EEG. There have been electrophysiological studies of the time course of the retro-cue effect in VWM (Kuo et al., 2012; Schneider et al., 2016), but these made use of univariate measures involving event-related components that required averaging across extended time windows, precluding a detailed time course assessment. In the current experiment, we wanted to track the removal of contents from VWM using highly time-resolved multivariate decoding methods. Recent studies have suggested that multivariate pattern analyses (MVPA) of the EEG signal provides an index of visual working memory load (Adam et al., 2020; Thyer et al., 2022). Specifically, these studies showed that MVPA-based decoding analyses could distinguish between levels of memory load, capturing differences as little as one item. Compared to univariate VWM load measures, such as contralateral delay activity (McCollough et al., 2007; Vogel & Machizawa, 2004) or negative slow wave (Fukuda et al., 2015), MVPA provides the advantages of being more sensitive and temporarily precise as it does not necessarily rely on the timing of a specific event-related potential measured from at most a few electrodes. In addition, load decoding with MVPA promises to be more universally applicable, as it does not depend on lateralized stimuli. Moreover, it has shown generalization across several factors, including item complexity and the type of information retained (Adam et al., 2020; Thyer et al., 2022). Thus the method promises to be both more sensitive and more versatile. We, therefore, applied this method in a second session, where we did not manipulate CTI, but provided sufficient post-cue time and then instead assessed whether multivariate signal traces the assumed reduction in VWM load. Specifically, we hypothesized that if the uncued information was indeed removed from VWM, the load-related multivariate signal should first resemble the high-load baseline, but over time start to resemble more and more the low-load baseline.

To foreshadow the findings, behaviorally we found a clear increase in the retro-cueing benefit with time, suggesting a relatively rapid shedding of load within about half a second. This extends earlier findings showing a similar time course to that observed for verbal information (Oberauer, 2018). However, the EEG findings indicated that the MVPA method was sensitive, not only to a load component but also to selection mechanisms. That is, whenever observers had to consolidate (after a pre-cue) or keep (after a retro-cue) a certain number of items in VWM, this entailed both a retention component (i.e. the more items selected, the more items to be maintained), and a selection component (i.e. select the items to consolidate, or select the items to keep in memory). We argue that it makes MVPA analyses susceptible to confounding interpretations, especially given the frequent co-occurrence of selection and maintenance requirements in standard VWM tasks. This is especially the case after the retro-cue, where the selection component precluded successful tracking of load reduction in VWM. Given that these components tend to correlate in VWM experiments in general, our study serves as a warning that it is important to clearly define what is meant by *load*. We will return to this in the General Discussion, but in the meantime we will use the term load as referring to the number of items retained in memory.

2 | METHOD

Data and scripts from the two sessions are available at OSF (https://osf.io/fh3wa/?view_only=ecb3bf32279e429cbf0c1b369ed68a0d).

2.1 | Participants

Fifty-six undergraduate students (mean age = 20.5, SD = 3.6, 50 females) from the Faculty of Psychology of the University of Murcia completed the first experimental session and were informed that they would be contacted again to complete a second experimental session. From these participants, those who performed close to or below chance in any of the three experimental conditions (i.e. under 0.60) were excluded from the analyses and were not contacted for the second session. From the remaining 52, 10 participants did not respond to the

invitation for the second session. Forty-two participants assisted in the EEG session. Three of them could not complete the task due to technical issues. Thus, a final sample of 39 participants completed the EEG session. All participants reported normal or corrected-to-normal vision and signed a written consent at the beginning of every session. Participants received course credits for their participation. The study was approved by the Ethics Committee of the University of Murcia and was conducted according to the ethical standards of the 1964 Declaration of Helsinki.

2.2 | Apparatus and stimuli

The tasks were programmed in E-Prime 3 and were performed in individual sound-attenuated booths where participants were seated and responded using a five buttons Chronos® device (Psychology Software Tools, Inc, 2016).

Figure 1 illustrates the task procedure. The task began with a fixation point of 500 ms which consisted of a dark gray colored circle with a dark gray point in the middle, inside an outline cross. The cross was designed to minimize eye movements (the ABC shape in Thaler et al., 2013) and was rotated 45 degrees from the original to accommodate the purpose of this experiment. The fixation point was followed by a cue of 250 ms duration in black. The cue consisted of the previous fixation point plus the filling of one or two of the arms of the cross Next, it appeared together with four colored disks for 350 ms. The cue indicated the disks that had to be encoded into memory, with a left diagonal cue indicating that the top left and bottom right items had to be encoded, and a right diagonal cue indicating the top right and bottom left items (2-load condition), and a fully filled cue indicated that all four items had to be encoded (4-load and 4/2-load conditions). This way we fully matched the initial perceptual information in all load conditions except for the cue. Furthermore, in the 2-load condition, we intentionally had participants recall items arranged diagonally, that is, both to the left and right of fixation, thus avoiding any lateralization effects on EEG and eye movements. Next, the fixation point was presented during a first delay of 1000 ms while participants had to keep in memory the indicated colored disks. Delay 1 was followed by a 200 ms retro-cue that differed between conditions. In the 4/2-load condition, the retro-cue was one of the diagonals, indicating that only two disks of the previously presented colored items should be retained, while the remaining items could be forgotten, as they would never be tested. In both the 2- and 4-load conditions, a non-informative fully filled cross was presented to indicate that what had to be encoded also had to be retain. Looking at the results, we observe a larger area under the ROC curve, A, and faster RTs in the 2-load condition (see Figure 2). This shows that participants correctly understood these instructions.

After the retro-cue there was a second delay after which a target probe was presented until the response, with a maximum of 2000 ms. In the behavioral session, the CTI varied between 50 and 1300 ms, while in the EEG session it lasted 1000 ms (see *General procedure*). The target consisted of a colored disk placed in one of two (2-load) or four (4 and 4/2-load) possible locations. Participants had to decide whether the previously memorized colored item at that position matched the target or not. On mismatch trials, the target disk was filled in with one color of the other item(s) that should have also been encoded, thus preventing participants from responding simply on the basis of familiarity.

Stimuli were presented on a 23-inch flatscreen LED monitor (LG 23MP68VQ-P) with a resolution of

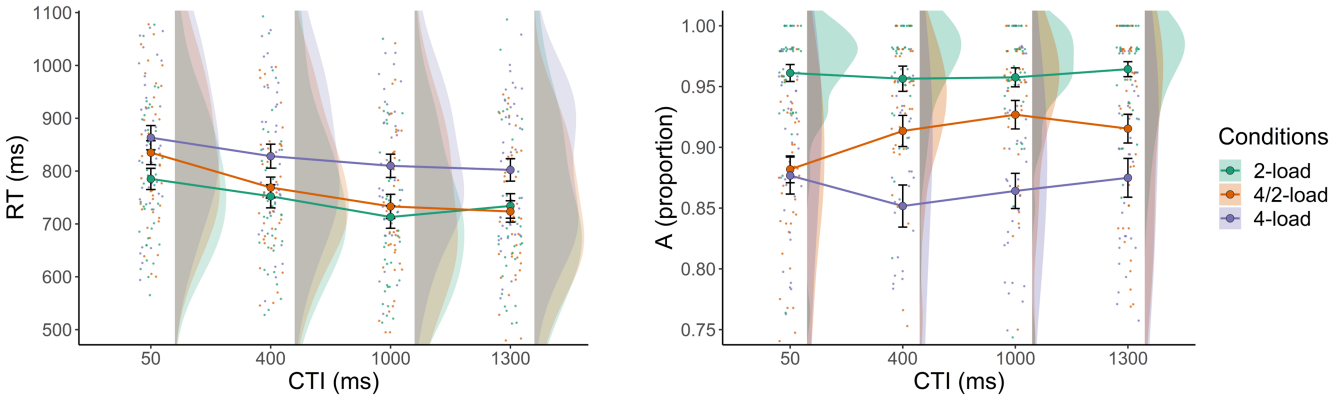

FIGURE 2 Reaction Times (RT) and mean A for the time course during the second delay for participants who completed both sessions, $N = 39$. Small colored points represent the mean of each participant in each condition. Points with black border represent the mean of each trial type for each interval. Error bars represent standard error of the mean. The individual distributions for all the trial types × CTI combinations are represented in the split violin plots.

1920 × 1080 pixels on a gray background, RGB (192, 192, 192). Participants were sitting 70 cm away from the screen. The fixation point was made out of a 1.23° diameter circle, a cross inside the circle (arms were 0.18° thick) and an additional circle (0.12° of diameter) in the intersection of the cross (see Figure 1), based on Thaler et al. (2013) recommendations to avoid eye movements during fixation. The circle of the fixation point was dark gray, and the arrows were the same color as the background. The distance of the memory disks from the fixation point was 1.23°. Disks had a radius of 0.6° and were equally distanced from each other. The color pool comprised 12 colors and it was obtained from de Vries et al. (2017). Colors were firstly determined in DKL color space and later converted to RGB. They were chosen to have the same contrast and luminance, differing only in hue. The 12 colors were discrete colors from an imaginary wheel in which consecutive colors were more similar than colors at the other extreme of the wheel (see methods in de Vries et al., 2017, for a detailed description of color extraction). Four non-consecutive colors were then randomly chosen for the memory set of each trial.

2.3 | General procedure

All participants completed two experimental sessions. At the beginning of the first session, they received instructions and practiced each condition separately for 17 trials. Additionally, they performed one extra practice block of 18 trials with all the three conditions mixed. Subsequently, they completed 17 experimental blocks of 18 trials each with a total of 306 trials, 102 in each condition. For the experimental blocks, the three conditions were always mixed. Importantly, in this first session, the CTI was manipulated to examine the time course of performance as a function of load reduction during the second delay. Four CTIs were chosen: 50, 400, 1000 and 1300 ms, with a mean of 25.5 and a minimum of 21 experimental trials for each combination of CTI and condition per participant.

For the second session (i.e. the EEG session), participants were given the same instructions and administered three short practice blocks of 10 trials, one for each condition. In this session, they completed 900 experimental trials, 300 of each condition, presented in 45 experimental blocks of 20 randomly shuffled trials. The CTI had a fixed duration of 1000 ms. Before beginning the practice trials, participants were explicitly encouraged to look at the fixation point, to avoid eye movements or blinking, and to keep their eyes as still as possible during the experiment. At the end of the practice, the light was dimmed, the door was closed, and EEG was recorded while they completed the experimental trials.

2.4 | Data recording and preprocessing

EEG data were recorded using BrainVision Recorder (Brain Products, 2020b) with a 64-channel ActiCap (Brain Products, 2020a) setup at 1000 Hz following the 10–20 system with active online referencing to the right mastoid. Impedances of the Ag/AgCl electrodes were kept below 5 kΩ. Bilateral horizontal EOG electrodes were placed on the outer canthi, together with two vertical EOG electrodes above and below the left eye. All preprocessing steps and analyses were performed in MATLAB, version R2020a (MATLAB, 2020). Preprocessing was performed using custom code in conjunction with EEGLAB (Delorme & Makeig, 2004). For the MVPA analysis, the Amsterdam Decoding and Modeling toolbox (ADAM; Fahrenfort et al., 2018) was used.

EEG raw data were first imported into EEGLAB (v2021.0) using the NE EEGLAB NIC plugin and re-referenced to the average of the left and right mastoids. The data were then downsampled to 200 Hz, followed by high-pass filter of 0.01 Hz using the pop_eegfiltnew function. This filter was chosen because of the artifacts that can occur when less conservative filters are applied to the data when performing multivariate classification, especially when trial durations are long as in this case (van Driel et al., 2021). Epochs from −0.45 to 2.8 s were extracted, locked at the beginning of the first cue until the end of the second delay. A baseline correction was applied, which spanned the 250 ms prior to the first cue. Independent Component Analysis (ICA) was then run with the compute_ICs_new function of the ADAM toolbox using the pop_runica function of EEGLAB. Eye blink components were removed using the ADJUST plugin and components were visually inspected to ensure that only the blink components were removed from the data. Errors and no responses were excluded from the analyses.

2.5 | Behavioral data analysis

Behavioral data were preprocessed in Rstudio (RStudio Team, 2022) and analyzed with JASP version 0.16.2.0 (JASP Team, 2022) for Reaction Times (RT) and A, a non-parametric measure of sensitivity that takes into account hits and false alarms (Zhang & Mueller, 2005). This measure is derived from signal detection theory (Green & Swets, 1966), and resolves some of the problems associated with d' and its non-parametric counterpart A' (Pollack & Norman, 1964). The R code to compute A is in OSF, and the formula is:

$$A = \begin{cases} \frac{3}{4} + \frac{H-F}{4} - F(1-H) & \text{if } F \leq 0.5 \leq H; \\ \frac{3}{4} + \frac{H-F}{4} - \frac{F}{4H} & \text{if } F \leq H < 0.5; \\ \frac{3}{4} + \frac{H-F}{4} - \frac{1-H}{4(1-F)} & \text{if } 0.5 < F \leq H. \end{cases} \quad (1)$$

Where, H is the hit rate and F is the false alarm rate for a given participant and condition (Zhang & Mueller, 2005). Participants with an accuracy of less than 0.6 (the proportion of trials in which a correct response was given) in any of the three main conditions in the first session were excluded from all analyses and were not contacted for the second session. When Mauchly's sphericity test reached statistical significance, the Greenhouse–Geisser correction was applied. For post hoc tests, the Holm-Bonferroni correction was used. For the data from both sessions, trials with errors and non-responses were removed from the RT analysis. This accounted for 13.25% and 14.86% of the data from the first and second session, respectively. We also excluded trials with responses below 250 ms, above 1900 ms, or 3 SD from the participant's mean. Based on these criteria, 0.27% and 0.25% of the data from the first and second session, respectively, were removed.

For the behavioral analysis of the first session, we analyzed the time course of the different conditions during the second delay. Data from the participants who completed both sessions were analyzed. For the final sample ($N = 39$), two repeated-measures ANOVAs were conducted with Condition (3 levels: retro-cue, low load, and high load) and CTI (4 levels: 50, 400, 1000, and 1300 ms) as within-participant factors and RTs and A as dependent variables. In Supplementary Materials we present analyses including all participants who successfully completed the first experimental session (52 participants). The main results were replicated for the total sample.

For the second session, the CTI was set at 1000 ms. Here, RTs and A were entered into one-way ANOVAs with Condition (3 levels: retro-cue, low load, and high load) as the main factor.

All plots were made in R (R Core Team, 2021) with RStudio (RStudio Team, 2022) using the ggplot2 package (Wickham, 2016) and the Raincloud tool (Allen et al., 2021).

2.6 | Multivariate pattern analysis (MVPA)

We analyzed data from the 39 participants who completed the EEG session. To assess the effects of the different conditions on eye movements, we performed separate analyses of the 59 EEG electrodes (excluding EOG) and the four EOG electrodes. Analyses on EOG data were performed on VEOG and HEOG channels, after subtraction of vertical (i.e. lower minus upper ocular electrode) and horizontal (i.e. right minus left ocular electrode) activity. MVPA analyses were performed with the Amsterdam Decoding and Modeling toolbox (ADAM; Fahrenfort et al., 2018). All analyses used Linear Discriminant Analysis (LDA) to predict the different conditions from the EEG data. Analyses were performed from the first cue until the target was presented, including the first delay (600–1600 ms epoch) and the second delay (1800–2800 ms epoch). Therefore, the entire epoch lasted 2.8 s. The data were downsampled to 40 Hz, resulting in 70 time points of interest. In addition, between-class and within-class balancing was applied prior to the analyses. Within-class balancing was applied by ensuring that the cue directionality of each condition was equally represented within each stimulus class. Small between-class imbalances were corrected to ensure that all classes were equally represented (e.g. the same number of trials in the 2- and 4-load conditions). In the current data, the mean percentage of oversampled trials was 6.03%, 0.36%, and 17% for the 4/2-load, 2-load, and 4-load conditions, respectively. Main analyses were repeated downsampling the number of trials used for decoding in each condition to that of the condition with the lowest number of trials, for each participant. Results are similar to the oversampling correction (see Figures S3 and S4).

In analyses where two classes were compared (e.g. 2-load versus 4-load), a 10-fold cross-validation method was used, where all trials were randomly distributed in the 10 folds. Within each participant, the model was trained on 90% of the trials at a specific time point and then tested on the remaining 10% of the trials (fold) at the same time point. This procedure was repeated 10 times, until each fold was tested once. The classifier performance was obtained by averaging the performance of all folds. This process was repeated for each time point in the epoch and for each participant. To estimate performance, we used the Area Under the Curve (AUC) of a Receiver Operating Characteristic (ROC), which is the area covered when plotting the cumulative probabilities of a class being classified as the class it belongs to (i.e. the true positive rate) against the cumulative probabilities of being classified as another class (i.e. the false positive rate). The AUC shows how well the classes are distinguished by the model and it goes from 0.5 (classification at theoretical chance level) to 1 (perfect classification), regardless of the number of classes in the analysis. Unlike binary classification accuracy, the AUC takes into account the confidence to classify

each individual case, that is, the distance from the decision boundary. This allows each case to be weighted according to the confidence with which it has been classified. To check that that theoretical chance level was appropriate, we calculated empirical chance level for the 2-load versus 4-load classification analysis. The AUC derived from this analysis provides an estimation of a sample-size dependent chance level classification AUC (Combrisson & Jerbi, 2015). This was calculated by decoding two classes but shuffling the labels of the conditions. We performed 10 iterations of this analysis and, on average, it performed almost exactly at the theoretical chance level (0.5010 vs 0.5). The variation in empirical chance across participants for the peak point of decoding based on these iterations was relatively low (min: 0.481; max: 0.531; M: 0.503; SD: 0.014). Thus, to reduce computational costs, the theoretical chance level was chosen for the analyses (see Figure S2).

The AUC was computed for each time point for each participant. To obtain group level results, a *t*-test was performed at each epoch time point between each participant's AUC values and the chance level (0.5). To deal with multiple comparisons in EEG data (large number of time points), the cluster-based random permutation testing (Maris & Oostenveld, 2007) was used. This method computes the probability of finding an observed cluster size (i.e. significant contiguous *t*-tests) under random permutation. The size of a cluster is determined by the sum of the *t*-values of that cluster. The significance threshold we used was $p < .05$, both for the individual *t*-tests and for the cluster-based analysis (for a detailed description of cluster-based random permutation testing see Maris & Oostenveld, 2007). Three types of analyses were conducted. First, several two-classes decoding analyses were performed. These analyses aim to show whether two classes can be distinguished based on the multivariate EEG data at different time points. Second, we also performed temporal generalization analyses, in which training and testing were performed on all possible combinations of epoch time points, generating what is known as a temporal generalization matrix. This matrix contains the performance for all possible combinations of training and test times and provides information about the stability/dynamics of brain activity over time (King & Dehaene, 2014). Finally, to assess the time course of the retro-cue effect on load, we trained the LDA algorithm in the two load conditions (2-load versus 4-load) and tested this classifier in the third condition (4/2-load). This analysis was used to determine whether the 4/2-load condition could be classified differently as one or the other condition at the time points of the delays. All significant latencies were relative to the onset of the initial cue.

TABLE 1 Descriptive statistics of behavioral data from participants who completed both experimental sessions, $N = 39$. Reaction Times (RT) and mean A (A) split by trial type (Condition) and cue-target interval (CTI).

Condition	CTI (ms)	RT (ms)		A (proportion)	
		Mean	SD	Mean	SD
2-load	50	785	125.1	0.961	0.044
	400	752	136.8	0.956	0.065
	1000	713	133.0	0.958	0.049
	1300	734	145.0	0.964	0.038
4/2-load	50	835	141.4	0.882	0.069
	400	769	123.2	0.913	0.080
	1000	733	143.4	0.927	0.073
	1300	724	126.4	0.915	0.074
4-load	50	863	143.2	0.877	0.095
	400	828	141.8	0.852	0.108
	1000	810	137.5	0.864	0.090
	1300	802	132.4	0.875	0.100

3 | RESULTS

Our predictions were as follows. At the behavioral level we expected performance in the 4/2 load condition to, over time, transition from resembling the 4-load baseline to the 2-load baseline. At the electrophysiological level we expected something similar. First, when applying MVPA to the maintenance delays of the 2- and 4-load baseline conditions, we expected to replicate Adam et al. (2020) findings, with VWM load being decodable from the multivariate EEG signal. Second, to investigate the time course of the removal stage, we then used these baselines as a comparison for the 4/2-load condition, where a retro-cue signaled only half of the information as relevant. We hypothesized that if the uncued information was indeed removed from VWM, the load-related signal should first resemble the high-load baseline, but over time start to resemble more and more the low-load baseline.

3.1 | Behavioral time course (session 1)

Table 1 shows the descriptive statistics for the RTs and A (Zhang & Mueller, 2005) in the first experimental session including only participants who completed both sessions ($N = 39$).[1] During this session, the time course of performance was assessed as a function of CTI. For the RTs, we

[1]Analyses with the whole sample that completed the first experimental ($N = 52$) session are in Supplementary Materials and the main interactions and post hoc tests show the same results.

observed a main effect of CTI, $F_{(2.25,85.63)} = 58.745$, $p < .001$, $\eta_p^2 = .607$, with overall RTs being faster with longer CTIs and stabilizing at the 1000 ms CTI. We also observed a main effect of Condition $F_{(2,76)} = 48.846$, $p < .001$, $\eta_p^2 = .562$, where the RTs were faster for the 2-load condition ($M = 746$, $SD = 130$) than for the 4/2-load and the 4-load conditions ($M = 758$, $SD = 130$ and $M = 762$, $SD = 127$, respectively). As expected, the Condition × CTI interaction reached statistical significance, $F_{(6,228)} = 3.564$, $p = .002$, $\eta_p^2 = .086$, showing different time courses for the three conditions. In the short 50 ms CTI, post hoc tests showed that the 4/2-load condition differed significantly from the 2-load condition, $t_{(38)} = 3.904$, $p = .004$, $d = 0.366$, but not from the 4-load condition, $t_{(38)} = 2.222$, $p = .464$, $d = 0.208$. However, from the second CTI onwards, the tendency reversed until the longest CTI, with statistically significant differences between the 4/2-load condition and the 4-load condition (CTI 2: $t_{(38)} = 4.657$, $p < .001$, $d = 0.436$; CTI 3: $t_{(38)} = 6.033$, $p < .001$, $d = 0.565$; CTI 4: $t_{(38)} = 6.174$, $p < .001$, $d = 0.579$), but not the 2-load condition, (CTI 2: $t_{(38)} = 1.301$, $p = 1$, $d = 0.122$; CTI 3: $t_{(38)} = 1.572$, $p = 1$, $d = 0.147$; CTI 4: $t_{(38)} = 0.823$, $p = 1$, $d = 0.077$). Thus, the development of RTs during delay 2 showed how the 4/2-load condition gradually distanced itself from the 4-load condition and ended up being comparable to the 2-load condition. This has been interpreted as the removal of information from WM (e.g. Souza et al., 2014), and the results here suggest it takes about half second.

For A values (sensitivity measure), a repeated-measures ANOVA showed a main effect of Condition, $F_{(1.68,63.77)} = 47.427$, $p < .001$, $\eta_p^2 = .555$, with the highest mean A for the 2-load condition ($M = 0.96$, $SD = 0.05$), lower for the 4/2-load condition ($M = 0.91$, $SD = 0.08$), and the lowest performance for the 4-load condition ($M = 0.87$, $SD = 0.1$). However, the main effect of CTI did not reach significance, showing that overall performance did not change across the second delay, $F_{(3,114)} = 1.287$, $p = .282$, $\eta_p^2 = .033$. Similar to the RT analysis and in line with what was expected, a significant CTI × Condition interaction was observed, $F_{(6,228)} = 3.214$, $p = .005$, $\eta_p^2 = .078$. post hoc comparisons showed a similar pattern to the one observed with RTs: the 4/2-load condition and the 4-load condition were equivalent at the shortest CTI ($t_{(38)} = .375$, $p = 1$, $d = 0.067$) and differed in the second and third CTIs ($t_{(38)} = 4.517$, $p < .001$, $d = 0.804$ and $t_{(38)} = 4.588$, $p < .001$, $d = 0.817$, respectively) but not in the last CTI ($t_{(38)} = 2.960$, $p = .102$, $d = 0.527$), and the reverse was true for the 4/2-load condition and the 2-load condition, which differed at the first delay ($t_{(38)} = 2.960$, $p < .001$, $d = 1.033$), became equal in the second and third interval ($t_{(38)} = 3.144$, $p = .062$, $d = 0.560$ and $t_{(38)} = 2.248$, $p = .536$, $d = 0.400$, respectively) and differed in the last one ($t_{(38)} = 3.578$, $p < .016$, $d = 0.637$). Thus, we observed a similar pattern of results, that is, a general improvement in performance for the 4/2-load condition with longer CTIs (see time course of RTs and A in Figure 2).

3.2 | Behavioral retro-cue effect (session 2)

In the second session, participants completed the same task but with a fixed 1000 ms CTI. Mirroring the first session results, the main effect of Condition was significant for the RTs, $F_{(1.64,62.39)} = 108.072$, $p < .001$, $\eta_p^2 = .740$, and A, $F_{(2,76)} = 117.744$, $p < .001$, $\eta_p^2 = .756$. For the RTs, post hoc comparisons showed that the 4-load condition ($M = 704.75$, $SD = 127.32$) was significantly slower than both the 2-load condition ($M = 629.68$, $SD = 109.17$), $t_{(38)} = 11.848$, $p < .001$, $d = 0.661$, and the 4/2-load condition ($M = 619.44$, $SD = 103.05$), $t_{(38)} = 13.463$, $p < .001$, $d = 0.751$, but the last two were equivalent, $t_{(38)} = 1.615$, $p = .111$, $d = 0.090$. For A values, post hoc comparisons showed a higher A for the 2-load condition ($M = 0.95$, $SD = 0.04$) than for the 4/2-load condition ($M = 0.90$, $SD = 0.06$), $t_{(38)} = 6.527$, $p < .001$, $d = 0.803$, and higher for the 4/2-load condition than for the 4-load condition ($M = 0.85$, $SD = 0.06$), $t_{(38)} = 8.764$, $p < .001$, $d = 1.078$ (see Figure 3).

3.3 | Can the change in VWM load be decoded from the EEG data?

First, to determine whether different levels of VWM load could be distinguished in the EEG data, MVPA was performed on two classes: 2-load and 4-load. The classifier was trained at each epoch time point to dissociate 4- from 2-load and then it was tested at the same time points, using k-folding (see Section 2). The EOG electrodes were excluded from this analysis and the raw data from the remaining 59 EEG electrodes were used. This analysis showed significant decoding from the beginning of the epoch to the end of the first delay (85–1535 ms). These results replicate the results of Adam et al. (2020). However, the two conditions became indistinguishable from the appearance of the second cue (the retro-cue) until the end of the epoch (see Figure 4a). Since the second cue was uninformative in these conditions (it should not alter the load), the pattern for this second delay was unexpected, as one would expect to still be able to decode the load even at the second delay. To further study the dynamics of these conditions, a temporal generalization analysis was performed. In this analysis, the algorithm was trained at each time point and tested at all other time points of the epoch. This method was then repeated for all the possible train-test combinations of time points. This allowed us to

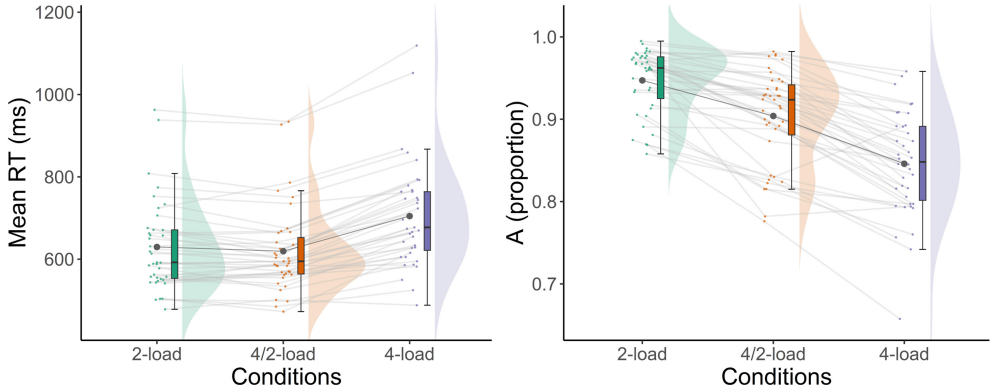

FIGURE 3 Mean Reaction Times (RT) and mean A in EEG session. Small colored points represent the mean of each participant in each condition, joined by a light gray line. Black points represent the mean of each trial type (conditions). Box plots of each condition are represented. The individual distributions for all the trial types × CTI combinations are represented in the split violin plots.

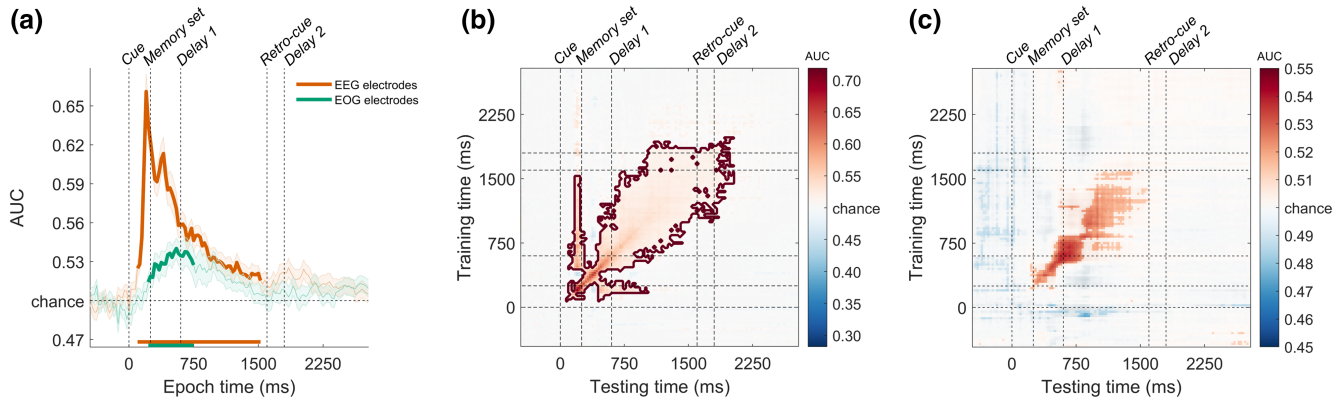

FIGURE 4 Decoding of load from electroencephalography (EEG) and electrooculogram (EOG) signal. AUC, Area Under the Curve. Dotted lines correspond with the onset of the following task moments, from left to right: initial cue, memory set, first delay, retro-cue, and second delay. (a) Diagonal decoding of 2-load versus 4-load conditions from EEG (orange line) and EOG (green line) electrodes. Bold lines show *p*-values that survived multiple comparisons corrections and shaded area surrounding the line is the standard error. (b) Temporal generalization matrix for 2-load vs 4-load decoding on EEG electrodes. (c) Temporal generalization matrix for 2-load vs 4-load decoding on EOG electrodes. In both (b) and (c), saturated colors reflect uncorrected $p < .05$ decoding, whereas areas circumscribed by a dark bold line highlight clusters that survive cluster-based permutation testing at $p < .05$. Above chance decoding is colored in red color and below chance decoding in blue.

assess the stability of the pattern. In addition, it allowed us to see whether training at a fixed time point (e.g. at peak activity) improved classification for the remaining time windows. Figure 4b showed a moderate generalization of decoding within the first delay period. However, at the time of the second cue and within the second subsequent delay period, the signal remained weak regardless of the training time point.

The results observed so far allow for multiple interpretations. On the one hand, the visuospatial information in WM may simply decay, and more so for the 4-load condition, which would then reduce the distinction between the two load conditions (Ricker et al., 2014, 2016). However, behavioral performance renders this unlikely, as a clear difference in performance between the 2-load and 4-load conditions was observed in both the behavioral and the EEG sessions. Another possibility is that decoding was decreased due to the greater distance from the pre-trial baselining period, which could increase noise in the data. However, further analyses showed significant decoding later in the epoch (see Figures 5, 6, and 7), making this unlikely. A third possibility is that some processes triggered by the retro-cue—and other than maintenance—mask the load manipulation decoding. In the present design, trials were randomly presented. In the 4/2-load condition, participants had to select some disks as relevant and drop the rest. In contrast, in the 2- and 4-load conditions—the ones of the present analysis—the second cue instruction was the same in both cases (i.e. a full cross), and indicated that nothing had changed (observers simply had to retain what

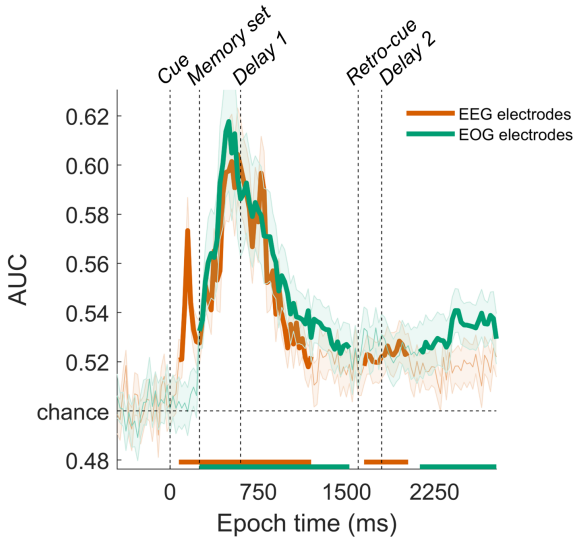

FIGURE 5 Diagonal decoding of the direction of the initial cue in the 2-load condition for EEG (orange) and EOG electrodes (green). AUC, Area Under the Curve. Bold lines show *p*-values that survived multiple comparisons corrections and shaded area surrounding the line is the standard error.

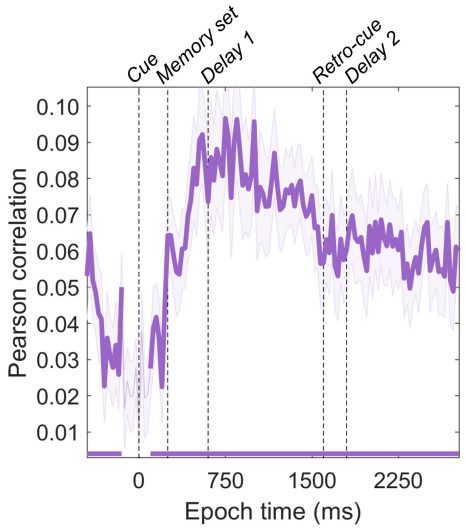

FIGURE 6 Correlation of EEG and EOG decoding for 2- vs 4-load. Bold lines show $p < .05$ values for Pearson correlations after cluster correction. Shaded area surrounding the line is the standard error.

they already had in memory). Then, it might be that a selection (vs non-selection) component is prevailing over the maintenance at the time of the retro-cue. If it was the case that the retro-cue was driving an attentional selection component that was superimposed on the load decoding, a similar case could apply to the pre-cue. In other words, in addition to load, the selection processes linked to the load manipulation, and triggered by the pre-cue, could in principle also contribute to the decoding during the first delay.

To test this possibility further, and given the close relationship between attentional selection and eye movements (van Ede et al., 2019; Zhao et al., 2012), we repeated the same analyses on the EOG electrodes. To isolate eye movements from brain activity, subtraction for VEOG (i.e. right minus left ocular electrode) and HEOG (i.e. lower minus upper ocular electrode) were computed. Thus, all the EOG analyses were performed on the obtained HEOG and VEOG subtracted channels. The idea behind these analyses was that, if 2- and 4-load conditions could be decoded from the signal coming from the ocular electrodes after the first cue, this would support the idea that attentional selection was also playing a role in the previous decoding analyses. The EOG analysis showed a significant decoding of the 2- and 4-load conditions during the memory set and the beginning of the first delay (0.235–0.760 ms; see Figure 4a). The temporal generalization was then also repeated in EOG. Although we observed a similar descriptive pattern as for the 59 EEG electrodes, there was no significant decoding after cluster correction (Figure 4c). This lack of significant decoding after cluster-based correction and the lower AUC values of the present analyses could be explained by the smaller number of variables included (4 EOG channels versus 59 EEG channels above).

To further test the role of selection, we additionally decoded the *directionality* of the initial cue of the 2-load condition. For this, two classes were made based on the diagonal direction of the encoding cue, that is, whether it was tilted to the left or to the right. Figure 5 shows diagonal decoding performed on EEG and EOG electrodes. In the EEG electrodes, the analysis showed significant decoding until the end of the first delay (85–1210 ms) and again during the second cue and the beginning of the second delay (1660–2035 ms). When the classification was performed at the EOG electrodes, significant decoding extended from memory encoding to the end of the trial (260–2785 ms), with a gap around the retro-cue. Again, these analyses showed a strong contribution of eye movements, closely linked to the selection required in the 2-load condition.

Finally, using the *adam_correlate_CONF_stats* function, we computed Spearman correlations at each time point between EEG and EOG classifier confidence scores (i.e. resulting from the distance to the decision boundary) of 2- vs 4-load results across trials for each participant, and tested the Fisher-transformed correlations against zero at the group level. The confidence scores in this analysis represent the single trial evidential support for the classification that the classifier provides at any given time point (see Figure 6). Thus, this analysis tested whether both decoding analyses captured similar information at any given time point.

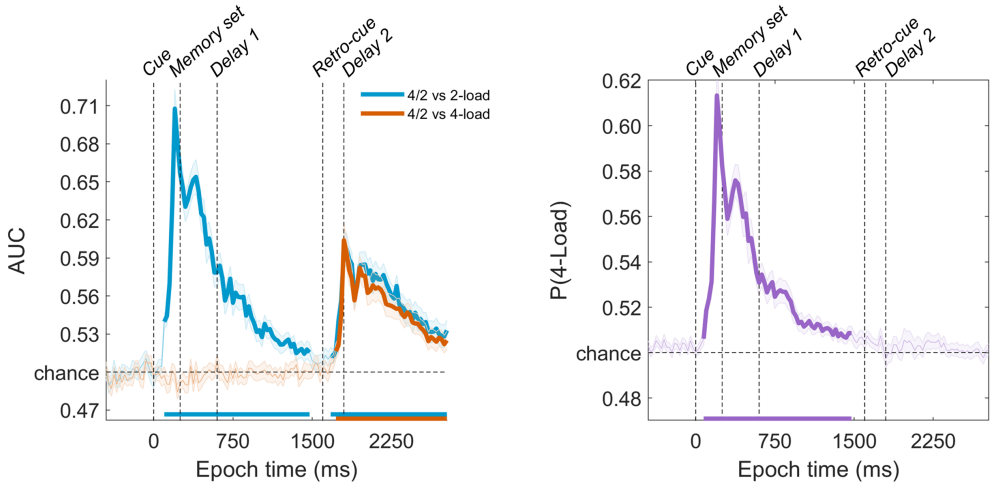

FIGURE 7 (a) Diagonal decoding 4/2-load vs 2- and 4-load conditions (blue and orange, respectively). AUC, Area Under the Curve. Bold lines show *p*-values that survived multiple comparisons corrections and shaded area surrounding the line is the standard error. (b) Training the 4- vs 2-load conditions and testing the 4/2-load condition. When classification is above chance, the 4/2-load condition is classified as 4-load and when it is below chance, the 4/2-load condition is classified as 2-load.

Results showed a significant positive correlation during the complete epoch. However, the correlation is numerically small, capturing relatively little shared variance. If we take the EOG analyses as evidence of selection, this analysis supports a contribution of selection to the 2- vs 4-load decoding. The correlation is higher during the first delay, congruent with the decoding results of Figure 4a, and where we expect differences in selection in these conditions. However, given the weak correlation, we can conclude that the decoding analysis cannot be driven by this selection component alone and that, in line with previous experiments, load is also being decoded (Adam et al., 2020; Thyer et al., 2022).

Apart from load and selection, there are two additional factors that could be partly contributing to the present significant decoding in the EEG channels: eye movements and the visual differences in the cue. Regarding eye movements, we interpreted here the EOG decoding as evidence for selection. However, we did not expect eye movements to be the only contribution to the significant decoding of the analyses run with the EEG channels because, in some cases, they may not co-occur with selection (Liu et al., 2022). Robustness analyses removing eye movements showed that decoding remained significant when we removed them before running the 4- vs 2-load decoding analysis (see ***Figure S3***). Thus, we can confirm that eye movements information cannot explain by itself the present analysis results. This is also congruent with the previous correlational analyses (Figure 6), that showed a small shared variance between EEG and EOG electrodes. Regarding any visual differences between different precues, earlier work showed that any sensory modulations would occur early, during the first 300 ms after the stimuli presentation (e.g. Jongen et al., 2007; Luck, 2006; Quentin et al., 2019). However, other decoding work showed that differential signals can be observed beyond that (Dijkstra et al., 2018; Noah et al., 2023). Thus, we cannot discard the possibility that these perceptual differences of the precues (and not only the attentional modulation produced by them) are partially driving the delay decoding.

Overall, we replicated previous work decoding visual WM load. However, we argue that differences in attentional selection of relevant items (as partially reflected in the overt eye movements) may also partially contribute to the observed decoding.

3.4 | Can we decode the drop in load after the retro-cue?

The second and main goal of the present work was to test the time course with which item information was removed from the VWM. For this purpose, we wanted to compared the 4/2-load condition with, on the one hand, the 2-load condition and, on the other hand, the 4-load condition (Figure 7a). If from the current data the load had been decoded, we would expect (1) the 4/2-load condition would not be distinguishable from the 4-load condition during the first delay, prior to the cue, while it would be distinguishable from the 2-load condition; (2) the 4/2-load would be similar to the 4-load condition and would be distinguishable from the 2-load condition at the beginning of the second delay and after the cue. As the second delay develops, this pattern should reverse, as the 4/2 load condition should begin to resemble more like the 2-load and less to the 4-load condition.

Thus, in a first step, we trained two linear discriminant classifiers at each epoch time point: one classifier was trained to discriminate between the 4/2-load and 2-load conditions, while the other was trained to discriminate between the 4/2-load and 4-load condition. These multivariate analyses showed significant decoding during the first delay between the 4/2-load condition and the 2-load condition (110–1485 ms), but not between the 4/2-load and the 4-load conditions. In other words, prior to the second cue, the 4/2-load condition resembled the 4-load condition more than the 2-load condition, as would be expected. Analyses of the second delay revealed above chance decoding between the 4/2-load and 2-load conditions (1685–2785 ms), as well as between the 4/2-load and 4-load conditions (1735–2785 ms) during virtually the same time windows. That is, the 4/2-load condition could be distinguished from both baseline conditions during the whole second delay and there was no sign of a transition from high to low load. This is congruent with previous studies that showed sustained decoding of an attentional selection component after a retro-cue (Quentin et al., 2019). For the sake of completeness, we performed an additional analysis in which the algorithm was trained on the baseline conditions (4-load vs. 2-load), and then tested in the 4/2-load condition at each time point, the result of which is fully in line with what would be predicted given the results in Figure 4a. This analysis allowed us to check at each time point whether the 4/2-load trials were classified as 4-load or 2-load. As in previous analyses, the 4/2-load condition could be classified as the 4-load condition during the first delay (85–1485 ms), but could not be classified as either category during the second delay (Figure 7b). These results confirm that the load reduction hypothesis could not be tested, given the lack of decoding of the 2-load and 4-load conditions during the second delay. Although we cannot discard the possibility that load was partly contributing to this decoding, the present results for the second delay are congruent with the multivariate analysis capturing the effects of cue-induced selection, as only the retro-cue condition required further selection for the second maintenance period.

To further support the conclusion that the second delay results were driven by the selection of information (this time, within the VWM), we trained linear classifiers to decode the directionality of the retro-cue (i.e. right- or left-tilted cue). This analysis was performed independently on EEG and EOG electrodes. The classifier found significant information about the directionality of the retro-cue at EEG electrodes (2410–2785 ms) and also at EOG electrodes (1985–2785 ms) (Figure 8). Since the retro-cue disappeared from the screen at 1800 ms, the fact that significant decoding was detected especially towards the end of the delay showed that participants could use eye movements

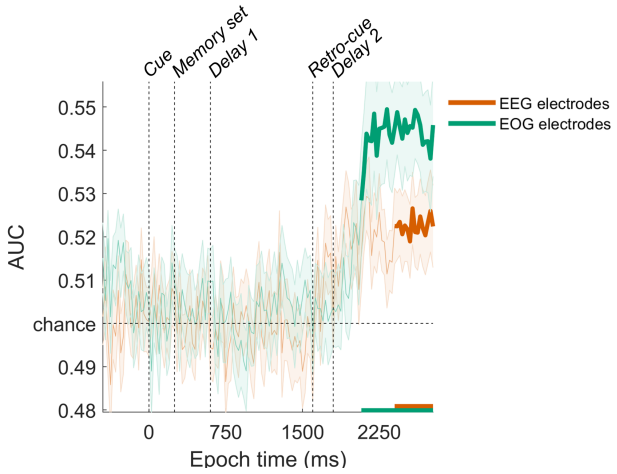

FIGURE 8 Diagonal decoding of the directionality of the retro-cue in the 4/2-load condition for EEG (orange) and EOG electrodes (green). AUC, Area Under the Curve. Bold lines show *p*-values that survived multiple comparisons corrections and shaded area surrounding the line is the standard error.

to support the maintenance of the relevant items active in VWM (van Ede & Nobre, 2023). The present results confirm that it is not a reduction of the VWM load, but the selection of relevant information within the VWM that underlies the decoding of the second delay results.

3.5 | Correlation of classifier performance and VWM capacity

To test whether participants with higher working memory capacity might show both better maintenance and better selection of items, we ran two additional correlational analyses. First, we correlated a measure of VWM capacity with the AUC from the MVPA analysis to classify 2- and 4-load trials (see Figure 4a). We know from previous work that decoding of 2- vs 4-load reflects load decoding (Adam et al., 2020; Thyer et al., 2022), and we also know from previous analyses that it might partially reflect the selection of items driven by the initial cue (see Section 3). To provide a measure of visuospatial working memory capacity, we calculated the estimated number of items remembered for each participant using the Cowan's K formula (Cowan, 2001; Rouder et al., 2008): $K=(H-FA)\times N$, where K is the number of items remembered, H and FA are the hit and false alarm rates, and N is the number of item presents to be remembered. As for the AUC measure, the first delay decoding was chosen because (1) it was the only time at which the classifier could differentiate conditions, and (2) to be as equivalent as possible to Adam et al. (2020). The Pearson correlation of the 39 participants who completed both experimental sessions showed a significant positive

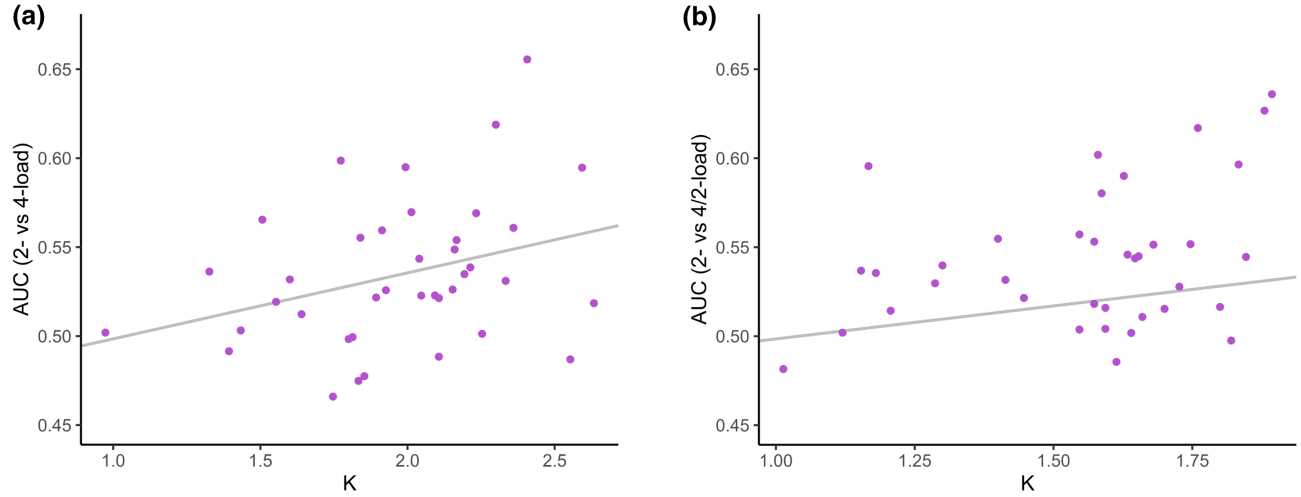

FIGURE 9 (a) Correlation of decoding for 2- vs 4-load and K during the first delay. (b) Correlation of decoding 2- vs 4/2-load and K during the second delay. Each point represents a single participant. The gray line represents the regression line of the correlation.

trend between the two variables, $r = .329$, $p = .041$, 95% CI = [0.584, 0.015], depicted in Figure 9a.

Additionally, we also performed a correlation where we explicitly tested whether decoding of selection would also be correlated with working memory capacity (K). For that, we used the second delay AUC of decoding of 2- and 4/2-load conditions and mean K of these same two conditions. The reason for choosing the second delay of these two conditions is because the same amount of load is relevant by the end of the delay (seen in K), but one of them has required a selection within WM of this information, the 4/2-load condition. Therefore, this analysis should largely restrict to selection. The Pearson correlation showed a significant positive trend, $r = .420$, $p = .008$, 95% CI = [0.649, 0.120], depicted in Figure 9b. These results showed that working memory capacity does not correlate only with load maintenance, but also with successful items selection.

4 | DISCUSSION

During the past decade, the cognitive neurosciences have seen an upsurge in the application of Multivariate Pattern Analysis (MVPA) to data from electro- and magnetoencephalography (EEG and MEG, respectively). Contrary to univariate methods, MVPA allows for the decoding or identification of different representations or states by considering multidimensional patterns of sensor activity, which makes it a highly sensitive technique (Grootswagers et al., 2017; Haxby et al., 2014). This has made MVPA a particularly popular tool in the field of working memory research, where it is used to track item-related representations and different memory states during delays, when the stimulus is absent (e.g. Bae & Luck, 2018; Bocincova & Johnson, 2019; King et al., 2016; LaRocque et al., 2013; Rose et al., 2016; Trübutschek et al., 2017; Wolff et al., 2015, 2017). However, the increased sensitivity and complexity comes with a price tag, as it is not always evident what the source is of the information that is being used to successfully decode. For instance, it has been shown that artifacts introduced by regular cleaning steps such as high-pass filtering can lead to spurious decoding (van Driel et al., 2021), as can unintended eye movements (Mostert et al., 2018; Quax et al., 2019). This may lead to potential confounds that can hinder or, in the worst case, invalidate the conclusions of a study, in that the multivariate patterns may not capture the mental state that they were intended to capture. The present study serves as another case in point.

We attempted to unravel the time course of the reduction in VWM load after observers have been cued which information needs to be retained, while other information can be dropped. To this aim, participants completed a visuospatial working memory task involving delayed recognition of colored stimuli in which VWM load was manipulated in three conditions: two or four items maintained throughout the trial (low- and high-load conditions, respectively), or two retro-cued items out of four initially encoded items, allowing a transition from high to low load. Two main results were obtained in the current study. First, the time course of behavioral performance (during the pre-EEG session) suggests that, following the cue, it takes about half a second to exclude irrelevant information that may affect behavior. Specifically, we varied the CTI from 50 to 1300 ms, and found that a retro-cue that told participants that only two of the four items would need to be retained caused RTs to drop from close to load 4 levels to close to load 2 levels within about 500 ms. To our knowledge, this is the first time that the time course

of visual working memory has been tracked. A similar pattern was previously reported by Oberauer (2018) using verbal stimuli, who observed a time course suggesting that participants needed around 1 s to drop the irrelevant information. Such timing differences could reflect not only the type of information (verbal versus visuospatial), but also the amount of information selected/removed, which was three out of six items in Oberauer (2018) and two out of four items in the present design. Future studies should take into account the variations in this time course by modifying the amount of information maintained and selected.

In a subsequent step, we sought to track the same time course of load reduction using EEG measurements. First, we replicated previous results that found that the MVPA analyses were sensitive to load (Adam et al., 2020). However, we found that decoding can also be sensitive to selection processes during the same epochs where one might expect load to change. Several characteristics of our results support this idea. First, during the time windows in which we observed a reliable load effect in the EEG, we also observed a reliable load effect in the EOG electrodes, suggesting the involvement of eye movements. Moreover, these measures were positively correlated, albeit weakly. This is consistent with the close link between selection of items within VWM and eye movements (van Ede et al., 2019). Additionally, the pre-cue directionality was also decoded from both EEG and EOG electrodes, supporting the enrollment of selection processes that were to some extent captured by the classifier. Second, during the second delay we predicted a transition in the 4/2-load retro-cue condition from being classified as more similar to the high-load condition (load four) at the beginning of the delay, to being classified as more similar to the low-load condition (load two) towards the end of the delay. In contrast, the 4/2-load retro-cue condition remained distinct from both baseline conditions throughout the delay and, if anything, the two baseline conditions were more similar, as indicated by weak classification performance. What distinguishes the retro-cue condition from the two baseline conditions is the selection component. Third, during the same time window we were able to distinguish the directionality of the retro-cue in both EEG and EOG electrodes, which allowed us to confirm that there is a general selection mechanism behind both cues. This is congruent with the conceptualization that proposes that selecting content maintained in working memory relies on processes similar to those involved when attending to perceptual information (Gazzaley & Nobre, 2012; Kiyonaga & Egner, 2013; Panichello & Buschman, 2021; van Ede & Nobre, 2023). Below we elaborate on this argument.

4.1 | Load effects: Retention versus selection

For our argument, it is important to clearly define what we mean by load and selection. Note that load as such is an independent variable, which is then assumed to affect certain mental processes. It is probably fair to say that when they think of load, researchers tend to think of *retention* (or maintenance), specifically the number of items that are held in some active state in working memory. Indeed, studies that have used MVPA to decode load also appear to interpret load this way (Adam et al., 2020; Feldmann-Wüstefeld, 2021; Thyer et al., 2022). This is also the definition of load that we adopted in the current work. From this conceptualization, retention is a stable and sustained process, which is reflected in neural indicators such as CDA (Vogel & Machizawa, 2004), negative slow wave (Fukuda et al., 2015), or load decoding temporal generalization (Adam et al., 2020). Note that more dynamic maintenance processes have also been proposed (Miller et al., 2018; Stokes, 2015). However, even those dynamic trajectories are stable and reproducible.

On the other side, we defined selection as a relatively transient, item-specific operation that transforms an item into another state for further processing. Examples of transition operations are separating items from distractors (i.e. input gating; Chatham & Badre, 2015; Frank et al., 2001; Nir-Cohen et al., 2020), encoding or consolidating information (i.e. transforming them into a maintenance state; Ricker et al., 2018; Woodman & Vogel, 2005), refreshing or updating relevant items (e.g. Kessler & Meiran, 2008; Lemaire et al., 2018), dropping items after a retro-cue raising or lowering the representational state (e.g. Kruijne et al., 2021; Oberauer, 2018), and retrieving items for response (i.e. output gating; Chatham et al., 2014). Many of these operations are assumed to involve some form of attention, in the sense of prioritizing or enhancing some representations over others. While these operations—and, therefore, also the selection process—are assumed to be transient (Woodman & Luck, 1999, 2003), this does not imply that they would not occur and be measurable throughout longer periods into the delay. First, such selection processes may last longer than typically assumed especially when they concern "internal selection", within memory (e.g. Oberauer, 2018; Quentin et al., 2019; Ricker et al., 2018; see also the time course after the retro-cue in our behavioral results). Second, even if the processes are transient and short-lived, they may occur at a delay, or occur repeatedly within working memory, especially when *multiple* items need to be selected for the next operation (e.g. Lemaire et al., 2018; Vogel et al., 2006; Vogel & Luck, 2002).

As a last remark, it is noteworthy that some researchers have argued that maintenance itself is a repeated attentional serial selection process to refresh items and raise their strength (Camos et al., 2018; Kiyonaga & Egner, 2013; Lepsien & Nobre, 2006; Olivers, 2008; Rac-Lubashevsky & Frank, 2021; Souza et al., 2015; Souza & Oberauer, 2017). If so, then any measured load effects would essentially be selection effects and vice versa. However, we would then have expected decoding during the second delay period in our study to be sensitive to the remaining load, which is not what we observed. Thus, to what extent selection and maintenance overlap or are independent processes is still a matter of debate.

We note that positive correlations between load and selection exist for most if not all of the experiments reanalyzed in Adam et al. (2020). Many of these experiments involved an initial cue pointing to the half of the display containing the set to be memorized. The size of this set was then varied to induce different loads. This then also involves different numbers of items to be *selected* for encoding and memory consolidation. Some other recent works trying to decode load also used experimental designs where load depended on selection operations, since only some colors or geometric figures needed to be maintained among other stimuli that had to be ignored (Feldmann-Wüstefeld, 2021; Thyer et al., 2022). Even when displays do not include spatial or feature cues, simply varying set sizes still implies varying the number of items that need to be selected and consolidated from a display. The same argument also goes for another piece of evidence reported by Adam et al. (2020) in support of load decoding, which was a positive correlation between classifier performance and individual working memory capacity. However, this correlation could also reflect more effective selection mechanisms, as previous work has shown that VWM capacity may actually reflect attentional filtering efficiency (Fukuda & Vogel, 2009; Vogel et al., 2005). We performed two additional analyses of our data where we correlated, for each participant, a measure of the mean number of items in memory (K) with the mean AUC of two decoding analyses that reflect the initial load manipulation (with the potential of perceptual selection) or selection from VWM (i.e. first delay 2- versus 4-load and second delay 2- versus 4/2-load, respectively). We too observed significant positive correlations for both analyses (see Figure 9), supporting the idea that high-capacity individuals may also be better selectors in addition to better loaders. Our point that decoding may at least in part reflect selection mechanisms does not in any way exclude the possibility that retention load was being decoded from those experiments. What we call for is for future experimental designs to focus on the dissociation between these two processes.

One might argue that Adam et al. (2020) already controlled for attention effects in their Experiment 3, where they compared a VWM task in which a number of lateralized colored squares had to be remembered to a sustained spatial attention task, in which the locations indicated by the same number of lateralized colored squares had to be monitored for the appearance of an unrelated visual target (a small line segment). Adam and colleagues found an effect of the number of colored squares in the VWM condition that lasted throughout the 1300 ms delay period. In the attention condition, the number of squares could also be decoded, but for about half that time period, around 700 ms. Adam et al., therefore, concluded that load decoding is caused by more than attention alone. However, this still leaves a decent, relatively long-lasting attention effect. Finally, it is also noteworthy that Adam et al. also observed strong cross-task decoding between the VWM and the attention task for those same 700 ms, suggesting shared mechanisms for at least the first half of the delay period.

A number of additional issues remain. First, it is possible that we failed to decode retention load because VWM activity fell "silent" during the first delay (Oberauer & Awh, 2022; Stokes, 2015). We cannot exclude this possibility, but we note that this delay was not that long (1000 ms), and others showed significant decoding with similar or longer delays (e.g. Feldmann-Wüstefeld, 2021; Thyer et al., 2022). Second, within the current design the 4/2 retro-cue may create an additional overall task demand, and the classifier may pick up on that rather than the selection operation per se. This could be addressed in future work by parametrically varying the number of items to be retained as indicated by the retro-cue. This would also help to further dissociate the relative contributions of selection versus retention processes. Third, we observed a contribution of eye movements to the decoding (see Figure S3). While such eye movements are consistent with selective attention operations (but may also play a role in maintenance), future studies should include additional eye-tracking measures to remove eye movements more accurately (Quax et al., 2019). Fourth, although we tried to minimize the visual differences of the cues, we observed an early peak in decoding with the appearance of the pre-cue, before the memory delay was presented (see Figures 5 and 7), and thus suggestive of differential perceptual signals. While for univariate EEG studies, these perceptual signals tend to emerge early (<300 ms; Jongen et al., 2007; Luck, 2006), the case might be different for multivariate analyses, where decoding of such signals may last further into the epoch (e.g. Dijkstra et al., 2018; Noah et al., 2023; but see Goddard et al., 2022; Quentin et al., 2019 for shorter perceptual decoding). Also using the retro-cue paradigm, Quentin et al. (2019) compared

the decoding of two cues: one that was simply perceived with one that required selecting certain stimuli for posterior retrieval. They found that the cue that implied selection showed a longer significant decoding (~1.5 s) than the cue that was merely perceived (~0.5 s). Even though we have no "cue perception" condition to make this direct comparison in the present data, our sustained decoding results would be in line with the sustained decoding of selection after the cue observed by Quentin et al. (2019). However, as stated before, the contribution of perceptual differences to decoding may vary between experiments. This will need to be taken into account in future studies attempting to differentiate between working memory load and working memory operations.

4.2 | Conclusion

The original goal of the present study was to track the time course of changes in visual working memory (VWM) load when part of the stored information in VWM is no longer relevant. While the behavioral performance suggested a gradual drop in load, we failed to observe an EEG correlate of this load reduction after the retro-cue. Instead, the multivariate analyses of EEG and EOG electrodes suggested an important role of attentional selection to flexibly update the relevance of VWM content. The present results are in line with previous work supporting the role of attention as a mechanism for information selection and prioritization within VWM (e.g. Astle et al., 2012; Murray et al., 2013; Panichello & Buschman, 2021; Serin & Günseli, 2022). Finally, the current results advise caution when using MVPA to track VWM retention per se, as other mechanisms may also contribute to the classification outcome.

AUTHOR CONTRIBUTIONS

Miriam Tortajada: Conceptualization; data curation; formal analysis; investigation; methodology; software; visualization; writing – original draft; writing – review and editing. **Johannes J. Fahrenfort:** Conceptualization; formal analysis; methodology; resources; software; supervision; visualization; writing – original draft; writing – review and editing. **Alejandro Sandoval-Lentisco:** Conceptualization; investigation; methodology; writing – review and editing. **Víctor Martínez-Pérez:** Conceptualization; investigation. **Lucía B. Palmero:** Conceptualization; investigation. **Alejandro Castillo:** Software. **Luis J. Fuentes:** Conceptualization; funding acquisition; supervision; writing – review and editing. **Guillermo Campoy:** Conceptualization; formal analysis; methodology; software; supervision. **Christian N. L. Olivers:** Conceptualization; methodology; project administration; resources; supervision; visualization; writing – original draft; writing – review and editing.

ACKNOWLEDGMENTS

This study was supported by grants PID2021-125408NB-I00 funded by MCIN/AEI/10.13039/501100011033 and by "ERDF A way of making Europe", and by the Spanish Ministry of Science, Innovation and Universities (predoctoral grants: FPU19/06017, FPU19/06016, FPU17/00427 and FPU18/00288). C.N.L.O. and J.J.F. were partly funded by the Dutch Organization for Scientific Research (NWO; grant 453-16-002).

CONFLICT OF INTEREST STATEMENT

The authors have no conflict of interest to declare.

DATA AVAILABILITY STATEMENT

Data and scripts from the two sessions are available at OSF (https://osf.io/fh3wa/?view_only=ecb3bf32279e429cbf0c1b369ed68a0d).

ETHICS STATEMENT

The study was approved by the Ethics Committee of the University of Murcia and was conducted according to the ethical standards of the 1964 Declaration of Helsinki.

ORCID

Miriam Tortajada https://orcid.org/0000-0002-3871-8283
Johannes J. Fahrenfort https://orcid.org/0000-0002-9025-3436
Alejandro Sandoval-Lentisco https://orcid.org/0000-0002-7876-0101
Víctor Martínez-Pérez https://orcid.org/0000-0003-0358-346X
Lucía B. Palmero https://orcid.org/0000-0001-5693-6500
Alejandro Castillo https://orcid.org/0000-0002-1482-7688
Luis J. Fuentes https://orcid.org/0000-0002-3203-9976
Guillermo Campoy https://orcid.org/0000-0002-3532-9178
Christian N. L. Olivers https://orcid.org/0000-0001-7470-5378

REFERENCES

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), e13691. https://doi.org/10.1111/psyp.13691

Allen, M., Poggiali, D., Whitaker, K., Marshall, T., van Langen, J., & Kievit, R. (2021). Raincloud plots: A multi-platform tool for

robust data visualization. *Wellcome Open Research*, *4*(63), 63. https://doi.org/10.12688/wellcomeopenres.15191.2

Astle, D. E., Summerfield, J., Griffin, I., & Nobre, A. C. (2012). Orienting attention to locations in mental representations. *Attention, Perception, & Psychophysics*, *74*(1), 146–162. https://doi.org/10.3758/s13414-011-0218-3

Bae, G.-Y., & Luck, S. J. (2018). Dissociable decoding of spatial attention and working memory from EEG oscillations and sustained potentials. *Journal of Neuroscience*, *38*(2), 409–422. https://doi.org/10.1523/JNEUROSCI.2860-17.2017

Bocincova, A., & Johnson, J. S. (2019). The time course of encoding and maintenance of task-relevant versus irrelevant object features in working memory. *Cortex*, *111*, 196–209. https://doi.org/10.1016/j.cortex.2018.10.013

Brain Products. (2020a). *actiCAP snap [apparatus]*. [Software]. Brain Products GmbH.

Brain Products. (2020b). *BrainVision recorder (version 1.23.0003) [software]*. [Software]. Brain Products GmbH.

Camos, V., Johnson, M., Loaiza, V., Portrat, S., Souza, A., & Vergauwe, E. (2018). What is attentional refreshing in working memory? *Annals of the New York Academy of Sciences*, *1424*(1), 19–32. https://doi.org/10.1111/nyas.13616

Chatham, C. H., & Badre, D. (2015). Multiple gates on working memory. *Current Opinion in Behavioral Sciences*, *1*, 23–31. https://doi.org/10.1016/j.cobeha.2014.08.001

Chatham, C. H., Frank, M. J., & Badre, D. (2014). Corticostriatal output gating during selection from working memory. *Neuron*, *81*(4), 930–942. https://doi.org/10.1016/j.neuron.2014.01.002

Combrisson, E., & Jerbi, K. (2015). Exceeding chance level by chance: The caveat of theoretical chance levels in brain signal classification and statistical assessment of decoding accuracy. *Journal of Neuroscience Methods*, *250*, 126–136. https://doi.org/10.1016/j.jneumeth.2015.01.010

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. (2010). The magical mystery four: How is working memory capacity limited, and why? *Current Directions in Psychological Science*, *19*(1), 51–57. https://doi.org/10.1177/0963721409359277

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

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

Dijkstra, N., Mostert, P., Lange, F. P. d., Bosch, S., & van Gerven, M. A. (2018). Differential temporal dynamics during visual imagery and perception. *eLife*, *7*, e33904. https://doi.org/10.7554/eLife.33904

Fahrenfort, J. J., van Driel, J., van Gaal, S., & Olivers, C. N. L. (2018). From ERPs to MVPA using the Amsterdam decoding and modeling toolbox (ADAM). *Frontiers in Neuroscience*, *12*, 351586. https://doi.org/10.3389/fnins.2018.00368

Feldmann-Wüstefeld, T. (2021). Neural measures of working memory in a bilateral change detection task. *Psychophysiology*, *58*(1), e13683. https://doi.org/10.1111/psyp.13683

Frank, M. J., Loughry, B., & O'Reilly, R. C. (2001). Interactions between frontal cortex and basal ganglia in working memory: A computational model. *Cognitive, Affective, & Behavioral Neuroscience*, *1*(2), 137–160. https://doi.org/10.3758/CABN.1.2.137

Fukuda, K., Mance, I., & Vogel, E. K. (2015). α power modulation and event-related slow wave provide dissociable correlates of visual working memory. *The Journal of Neuroscience*, *35*(41), 14009–14016. https://doi.org/10.1523/JNEUROSCI.5003-14.2015

Fukuda, K., & Vogel, E. K. (2009). Human variation in overriding attentional capture. *Journal of Neuroscience*, *29*(27), 8726–8733. https://doi.org/10.1523/JNEUROSCI.2145-09.2009

Gazzaley, A., & Nobre, A. C. (2012). Top-down modulation: Bridging selective attention and working memory. *Trends in Cognitive Sciences*, *16*(2), 129–135. https://doi.org/10.1016/j.tics.2011.11.014

Goddard, E., Carlson, T. A., & Woolgar, A. (2022). Spatial and feature-selective attention have distinct, interacting effects on population-level tuning. *Journal of Cognitive Neuroscience*, *34*(2), 290–312. https://doi.org/10.1162/jocn_a_01796

Green, D. M., & Swets, J. A. (1966). *Signal detection theory and psychophysics*. John Wiley.

Gressmann, M., & Janczyk, M. (2016). The (un)clear effects of invalid retro-cues. *Frontiers in Psychology*, *7*, 167519. https://doi.org/10.3389/fpsyg.2016.00244

Grootswagers, T., Wardle, S. G., & Carlson, T. A. (2017). Decoding dynamic brain patterns from evoked responses: A tutorial on multivariate pattern analysis applied to time series neuroimaging data. *Journal of Cognitive Neuroscience*, *29*(4), 677–697. https://doi.org/10.1162/jocn_a_01068

Günseli, E., van Moorselaar, D., Meeter, M., & Olivers, C. N. L. (2015). The reliability of retro-cues determines the fate of noncued visual working memory representations. *Psychonomic Bulletin & Review*, *22*(5), 1334–1341. https://doi.org/10.3758/s13423-014-0796-x

Haxby, J. V., Connolly, A. C., & Guntupalli, J. S. (2014). Decoding neural representational spaces using multivariate pattern analysis. *Annual Review of Neuroscience*, *37*, 435–456. https://doi.org/10.1146/annurev-neuro-062012-170325

JASP Team. (2022). *JASP (Version 0.16.2.0) [Computer software]* [Software]. https://jasp-stats.org/

Jongen, E. M. M., Smulders, F. T. Y., & Van der Heiden, J. S. H. (2007). Lateralized ERP components related to spatial orienting: Discriminating the direction of attention from processing sensory aspects of the cue. *Psychophysiology*, *44*(6), 968–986. https://doi.org/10.1111/j.1469-8986.2007.00557.x

Kessler, Y., & Meiran, N. (2008). Two dissociable updating processes in working memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *34*(6), 1339–1348. https://doi.org/10.1037/a0013078

King, J. R., & Dehaene, S. (2014). Characterizing the dynamics of mental representations: The temporal generalization method. *Trends in Cognitive Sciences*, *18*(4), 203–210. https://doi.org/10.1016/j.tics.2014.01.002

King, J. R., Pescetelli, N., & Dehaene, S. (2016). Brain mechanisms underlying the brief maintenance of seen and unseen sensory information. *Neuron*, *92*(5), 1122–1134. https://doi.org/10.1016/j.neuron.2016.10.051

Kiyonaga, A., & Egner, T. (2013). Working memory as internal attention: Toward an integrative account of internal and external

selection processes. *Psychonomic Bulletin & Review*, *20*(2), 228–242. https://doi.org/10.3758/s13423-012-0359-y

Kruijne, W., Bohte, S. M., Roelfsema, P. R., & Olivers, C. N. L. (2021). Flexible working memory through selective gating and attentional tagging. *Neural Computation*, *33*(1), 1–40. https://doi.org/10.1162/neco_a_01339

Kuo, B. C., Stokes, M. G., & Nobre, A. C. (2012). Attention modulates maintenance of representations in visual short-term memory. *Journal of Cognitive Neuroscience*, *24*(1), 51–60. https://doi.org/10.1162/jocn_a_00087

LaRocque, J. J., Lewis-Peacock, J. A., Drysdale, A. T., Oberauer, K., & Postle, B. R. (2013). Decoding attended information in short-term memory: An EEG study. *Journal of Cognitive Neuroscience*, *25*(1), 127–142. https://doi.org/10.1162/jocn_a_00305

Lemaire, B., Pageot, A., Plancher, G., & Portrat, S. (2018). What is the time course of working memory attentional refreshing? *Psychonomic Bulletin & Review*, *25*(1), 370–385. https://doi.org/10.3758/s13423-017-1282-z

Lepsien, J., & Nobre, A. C. (2006). Cognitive control of attention in the human brain: Insights from orienting attention to mental representations. *Brain Research*, *1105*(1), 20–31. https://doi.org/10.1016/j.brainres.2006.03.033

Lewis-Peacock, J. A., Kessler, Y., & Oberauer, K. (2018). The removal of information from working memory. *Annals of the New York Academy of Sciences*, *1424*(1), 33–44. https://doi.org/10.1111/nyas.13714

Liu, B., Nobre, A. C., & van Ede, F. (2022). Functional but not obligatory link between microsaccades and neural modulation by covert spatial attention. *Nature Communications*, *13*(1), 3503. https://doi.org/10.1038/s41467-022-31217-3

Luck, S. J. (2006). The operation of attention—Millisecond by millisecond—Over the first half second. In *The first half second: The microgenesis and temporal dynamics of unconscious and conscious visual processes* (pp. 187–206). MIT Press.

Makovski, T. (2012). Are multiple visual short-term memory storages necessary to explain the retro-cue effect? *Psychonomic Bulletin & Review*, *19*(3), 470–476. https://doi.org/10.3758/s13423-012-0235-9

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

MATLAB. (2020). *9.8.0.1323502 (R2020a)*. The MathWorks Inc.

McCollough, A. W., Machizawa, M. G., & Vogel, E. K. (2007). Electrophysiological measures of maintaining representations in visual working memory. *Cortex*, *43*(1), 77–94. https://doi.org/10.1016/S0010-9452(08)70447-7

Miller, E. K., Lundqvist, M., & Bastos, A. M. (2018). Working memory 2.0. *Neuron*, *100*(2), 463–475. https://doi.org/10.1016/j.neuron.2018.09.023

Mostert, P., Albers, A. M., Brinkman, L., Todorova, L., Kok, P., & de Lange, F. P. (2018). Eye movement-related confounds in neural decoding of visual working memory representations. *eNeuro*, *5*(4), ENEURO.0401-17.2018. https://doi.org/10.1523/ENEURO.0401-17.2018

Murray, A. M., Nobre, A. C., Clark, I. A., Cravo, A. M., & Stokes, M. G. (2013). Attention restores discrete items to visual short-term memory. *Psychological Science*, *24*(4), 550–556. https://doi.org/10.1177/0956797612457782

Myers, N. E., Stokes, M. G., & Nobre, A. C. (2017). Prioritizing information during working memory: Beyond sustained internal attention. *Trends in Cognitive Sciences*, *21*(6), 449–461. https://doi.org/10.1016/j.tics.2017.03.010

Nir-Cohen, G., Kessler, Y., & Egner, T. (2020). Neural substrates of working memory updating. *Journal of Cognitive Neuroscience*, *32*(12), 2285–2302. https://doi.org/10.1162/jocn_a_01625

Noah, S., Meyyappan, S., Ding, M., & Mangun, G. R. (2023). Time courses of attended and ignored object representations. *Journal of Cognitive Neuroscience*, *35*(4), 645–658. https://doi.org/10.1162/jocn_a_01972

Oberauer, K. (2018). Removal of irrelevant information from working memory: Sometimes fast, sometimes slow, and sometimes not at all. *Annals of the New York Academy of Sciences*, *1424*(1), 239–255.

Oberauer, K., & Awh, E. (2022). Is there an activity-silent working memory? *Journal of Cognitive Neuroscience*, *34*(12), 2360–2374. https://doi.org/10.1162/jocn_a_01917

Olivers, C. N. L. (2008). Interactions between visual working memory and visual attention. *Frontiers in Bioscience: a Journal and Virtual Library*, *13*, 1182–1191. https://doi.org/10.2741/2754

Panichello, M. F., & Buschman, T. J. (2021). Shared mechanisms underlie the control of working memory and attention. *Nature*, *592*(7855), 601–605. https://doi.org/10.1038/s41586-021-03390-w

Pertzov, Y., Bays, P. M., Joseph, S., & Husain, M. (2013). Rapid forgetting prevented by retrospective attention cues. *Journal of Experimental Psychology: Human Perception and Performance*, *39*, 1224–1231. https://doi.org/10.1037/a0030947

Pollack, I., & Norman, D. A. (1964). A non-parametric analysis of recognition experiments. *Psychonomic Science*, *1*(1), 125–126. https://doi.org/10.3758/BF03342823

Psychology Software Tools, Inc. (2016). *E-Prime 3.0*. [Software]. https://support.pstnet.com/

Quax, S. C., Dijkstra, N., van Staveren, M. J., Bosch, S. E., & van Gerven, M. A. J. (2019). Eye movements explain decodability during perception and cued attention in MEG. *NeuroImage*, *195*, 444–453. https://doi.org/10.1016/j.neuroimage.2019.03.069

Quentin, R., King, J.-R., Sallard, E., Fishman, N., Thompson, R., Buch, E. R., & Cohen, L. G. (2019). Differential brain mechanisms of selection and maintenance of information during working memory. *Journal of Neuroscience*, *39*(19), 3728–3740. https://doi.org/10.1523/JNEUROSCI.2764-18.2019

R Core Team. (2021). R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing. https://www.R-project.org/

Rac-Lubashevsky, R., & Frank, M. J. (2021). Analogous computations in working memory input, output and motor gating: Electrophysiological and computational modeling evidence. *PLoS Computational Biology*, *17*(6), e1008971. https://doi.org/10.1371/journal.pcbi.1008971

Ricker, T. J., Nieuwenstein, M. R., Bayliss, D. M., & Barrouillet, P. (2018). Working memory consolidation: Insights from studies on attention and working memory. *Annals of the New York Academy of Sciences*, *1424*(1), 8–18. https://doi.org/10.1111/nyas.13633

Ricker, T. J., Spiegel, L. R., & Cowan, N. (2014). Time-based loss in visual short-term memory is from trace decay, not temporal distinctiveness. *Journal of Experimental Psychology. Learning,*

Memory, and Cognition, *40*(6), 1510–1523. https://doi.org/10.1037/xlm0000018

Ricker, T. J., Vergauwe, E., & Cowan, N. (2016). Decay theory of immediate memory: From Brown (1958) to today (2014). *Quarterly Journal of Experimental Psychology (2006)*, *69*(10), 1969–1995. https://doi.org/10.1080/17470218.2014.914546

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., Cowan, N., Zwilling, C. E., Morey, C. C., & Pratte, M. S. (2008). An assessment of fixed-capacity models of visual working memory. *Proceedings of the National Academy of Sciences*, *105*(16), 5975–5979. https://doi.org/10.1073/pnas.0711295105

RStudio Team. (2022). *RStudio: Integrated development environment for R* (2022.02.3) [Software]. RStudio, Inc. http://www.rstudio.com/

Schneider, D., Mertes, C., & Wascher, E. (2016). The time course of visuo-spatial working memory updating revealed by a retro-cuing paradigm. *Scientific Reports*, *6*(1), 1. https://doi.org/10.1038/srep21442

Serin, F., & Günseli, E. (2022). Internal attention is the only retroactive mechanism for controlling precision in working memory. *Attention, Perception, & Psychophysics*, *85*, 1375–1386. https://doi.org/10.3758/s13414-022-02628-7

Shepherdson, P., Oberauer, K., & Souza, A. S. (2018). Working memory load and the retro-cue effect: A diffusion model account. *Journal of Experimental Psychology: Human Perception and Performance*, *44*, 286–310. https://doi.org/10.1037/xhp0000448

Souza, A. S., & Oberauer, K. (2016). In search of the focus of attention in working memory: 13 years of the retro-cue effect. *Attention, Perception, & Psychophysics*, *78*(7), 1839–1860. https://doi.org/10.3758/s13414-016-1108-5

Souza, A. S., & Oberauer, K. (2017). The contributions of visual and central attention to visual working memory. *Attention, Perception, & Psychophysics*, *79*(7), 1897–1916. https://doi.org/10.3758/s13414-017-1357-y

Souza, A. S., Rerko, L., & Oberauer, K. (2014). Unloading and reloading working memory: Attending to one item frees capacity. *Journal of Experimental Psychology: Human Perception and Performance*, *40*(3), 1237–1256. https://doi.org/10.1037/a0036331

Souza, A. S., Rerko, L., & Oberauer, K. (2015). Refreshing memory traces: Thinking of an item improves retrieval from visual working memory. *Annals of the New York Academy of Sciences*, *1339*(1), 20–31. https://doi.org/10.1111/nyas.12603

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

Thaler, L., Schütz, A. C., Goodale, M. A., & Gegenfurtner, K. R. (2013). What is the best fixation target? The effect of target shape on stability of fixational eye movements. *Vision Research*, *76*, 31–42. https://doi.org/10.1016/j.visres.2012.10.012

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

Trübutschek, D., Marti, S., Ojeda, A., King, J.-R., Mi, Y., Tsodyks, M., & Dehaene, S. (2017). A theory of working memory without consciousness or sustained activity. *eLife*, *6*, e23871. https://doi.org/10.7554/eLife.23871

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*, 109080. https://doi.org/10.1016/j.jneumeth.2021.109080

van Ede, F., Chekroud, S. R., & Nobre, A. C. (2019). Human gaze tracks attentional focusing in memorized visual space. *Nature Human Behaviour*, *3*(5), 462–470. https://doi.org/10.1038/s41562-019-0549-y

van Ede, F., & Nobre, A. C. (2023). Turning attention inside out: How working memory serves behavior. *Annual Review of Psychology*, *74*, 137–165. https://doi.org/10.2139/ssrn.4082572

van Moorselaar, D., Gunseli, E., Theeuwes, J., & Olivers, C. N. L. (2015). The time course of protecting a visual memory representation from perceptual interference. *Frontiers in Human Neuroscience*, *8*, 1053. https://doi.org/10.3389/fnhum.2014.01053

Vogel, E. K., & Luck, S. J. (2002). Delayed working memory consolidation during the attentional blink. *Psychonomic Bulletin & Review*, *9*(4), 739–743. https://doi.org/10.3758/BF03196329

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

Vogel, E. K., McCollough, A. W., & Machizawa, M. G. (2005). Neural measures reveal individual differences in controlling access to working memory. *Nature*, *438(7067), Article 7067*, 500–503. https://doi.org/10.1038/nature04171

Vogel, E. K., Woodman, G. F., & Luck, S. J. (2001). Storage of features, conjunctions and objects in visual working memory. *Journal of Experimental Psychology. Human Perception and Performance*, *27*(1), 92–114. https://doi.org/10.1037//0096-1523.27.1.92

Vogel, E. K., Woodman, G. F., & Luck, S. J. (2006). The time course of consolidation in visual working memory. *Journal of Experimental Psychology: Human Perception and Performance*, *32*(6), 1436–1451. https://doi.org/10.1037/0096-1523.32.6.1436

Wickham, H. (2016). *ggplot2: Elegant graphics for data analysis*. Springer-Verlag. https://ggplot2.tidyverse.org

Williams, M., Hong, S. W., Kang, M.-S., Carlisle, N. B., & Woodman, G. F. (2013). The benefit of forgetting. *Psychonomic Bulletin & Review*, *20*(2), 348–355. https://doi.org/10.3758/s13423-012-0354-3

Wolff, M. J., Ding, J., Myers, N., & Stokes, M. (2015). Revealing hidden states in visual working memory using electroencephalography. *Frontiers in Systems Neuroscience*, *9*, 123. https://doi.org/10.3389/fnsys.2015.00123

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

Woodman, G. F., & Luck, S. J. (1999). Electrophysiological measurement of rapid shifts of attention during visual search. *Nature*, *400*(6747), 867–869. https://doi.org/10.1038/23698

Woodman, G. F., & Luck, S. J. (2003). Serial deployment of attention during visual search. *Journal of Experimental Psychology.*

Human Perception and Performance, *29*(1), 121–138. https://doi.org/10.1037//0096-1523.29.1.121

Woodman, G. F., & Vogel, E. K. (2005). Fractionating working memory: Consolidation and maintenance are independent processes. *Psychological Science*, *16*(2), 106–113. https://doi.org/10.1111/j.0956-7976.2005.00790.x

Zhang, J., & Mueller, S. T. (2005). A note on ROC analysis and nonparametric estimate of sensitivity. *Psychometrika*, *70*(1), 203–212. https://www.doi.org/10.1007/s11336-003-1119-8

Zhao, M., Gersch, T. M., Schnitzer, B. S., Dosher, B. A., & Kowler, E. (2012). Eye movements and attention: The role of pre-saccadic shifts of attention in perception, memory and the control of saccades. *Vision Research*, *74*, 40–60. https://doi.org/10.1016/j.visres.2012.06.017

SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.

Figure S1. Reaction Times (RT) and mean A for the time course during the second delay for all participants who completed the first session, $N = 52$. Small colored points represent the mean of each participant in each condition. Points with black border represent the mean of each trial type for each interval. Error bars represent standard error of the mean. The individual distributions for all the trial types × CTI combinations are represented in the split violin plots.

Figure S2. Classifier decoding performance randomizing labels, showing empirical chance level. AUC, Area Under the Curve. Bold lines show *p*-values that survived multiple comparison corrections and the shaded area surrounding the line is the standard error.

Figure S3. Reanalysis after different EEG preprocessing of decoding of load from electroencephalography (EEG) and electrooculogram (EOG) signal. AUC, Area Under the Curve. (a) Diagonal decoding of 2-load versus 4-load conditions from EEG (orange line) and EOG (green line) electrodes. Bold lines show *p*-values that survived multiple comparisons corrections and shaded area surrounding the line is the standard error. (b) Temporal generalization matrix for 2-load vs 4-load decoding on EEG electrodes. (c) Temporal generalization matrix for 2-load vs 4-load decoding on EOG electrodes. In both (b and c), saturated colors reflect uncorrected $p < .05$ decoding, whereas areas circumscribed by a dark bold line highlight clusters that survive cluster-based permutation testing at $p < .05$. Above chance decoding is colored in red color and below chance decoding in blue.

Figure S4. Reanalysis after preprocessing modifications. (a) Diagonal decoding 4/2-load vs 2- and 4-load conditions (blue and orange, respectively). AUC, Area Under the Curve. Bold lines show *p*-values that survived multiple comparisons corrections and shaded area surrounding the line is the standard error. (b) Training the 4- vs 2-load conditions and testing the 4/2-load condition. When classification is above chance, the 4/2-load condition is classified as 4-load and when it is below chance, the 4/2-load condition is classified as 2-load.

Table S1. Descriptive statistics of behavioral data from all participants who completed the first session, $N = 52$.

How to cite this article: Tortajada, M., Fahrenfort, J. J., Sandoval-Lentisco, A., Martínez-Pérez, V., Palmero, L. B., Castillo, A., Fuentes, L. J., Campoy, G., & Olivers, C. N. L. (2024). Decoding load or selection in visuospatial working memory? *Psychophysiology*, *00*, e14636. https://doi.org/10.1111/psyp.14636