Concurrent visual and motor selection during visual working memory guided action

Freek van Ede[1]*, Sammi R Chekroud[1,2], Mark G Stokes[2], Anna C Nobre[1,2]

[1]*Oxford Centre for Human Brain Activity, Wellcome Centre for Integrative Neuroimaging, Department of Psychiatry, University of Oxford, UK*

[2]*Department of Experimental Psychology, University of Oxford, UK*

*Corresponding author: Dr Freek van Ede (frederik.vanede@ohba.ox.ac.uk)

Abstract | Visual working memory enables us to hold onto past sensations in anticipation that these may become relevant for guiding future actions. Yet, laboratory tasks have treated visual working memories in isolation from their prospective actions and have focused on the mechanisms of memory retention rather than utilisation. To understand how visual memories become utilised for action, we linked individual memory items to particular actions and independently tracked the neural dynamics of visual and motor selection when memories became utilised for action. This revealed concurrent visual-motor selection, engaging appropriate visual and motor brain areas at the same time. Thus we show that items in visual working memory can invoke multiple, item-specific, action plans that can be accessed together with the visual representations that guide them – affording fast and precise memory guided behaviour.

Introduction | Effective behaviour requires detailed sensory information to guide action, but this information is often unavailable to our senses at the time actions become relevant – for example, because visual objects have become occluded or because we have looked away. Visual working memory [1–5] is the core cognitive function that bridges potentially relevant visual sensations to anticipated future actions. Despite this strong conceptual link between visual working memory and motor control, popular laboratory tasks of visual working memory (e.g. [4–6]) tend to consider visual representations in isolation from their prospective actions, while tasks of action preparation (or 'motor' working memory; e.g. [7–9]) tend to neglect the potential contribution of visual memory representations that may guide action. In addition, while the cognitive and neural mechanisms of working-memory retention have received ample investigation, little remains known about the mechanism of working-memory utilisation – i.e., when working memories are actually 'put to work' [10,11]. To understand how visual working memories guide action requires investigating how both visual memories and their corresponding actions become selected to support memory guided behaviour. To this end, we developed a novel laboratory task of working memory in which we linked individual memories to particular actions. Through a carefully balanced task-design, we were able to leverage electroencephalography (EEG) to individuate and independently track human brain activity related to the visual location and the response hand that were uniquely associated with the probed memory item that became relevant for action.

Results | Twenty-five healthy human volunteers performed a working-memory task (**Fig. 1a**) that fused conventional visual and motor working-memory tasks. One of two coloured visual items (tilted bars) was equally likely to become utilised for action after a brief memory delay (randomly drawn between 2 and 2.5 s). A colour change of the central fixation cross (the memory probe) prompted participants to select the colour-matching item in order to reproduce its tilt as accurately as possible. To link visual memory items to specific actions in a controlled laboratory setting, the hand required for responding was directly linked to the tilt of the probed item. Participants pressed a key with their right (left) index finger to initiate a clockwise (counter-clockwise) rotation of a visualised response dial and released the key when the dial reached the desired tilt, terminating the response. The central response dial always started in the vertical position and could be rotated by maximally ± 90

degrees. As a consequence, a leftward (rightward) tilted item could only ever be accurately reported with a left (right) key press. Each trial contained one leftward and one rightward tilted item (each randomly tilted between 20 and 70 degrees) that were randomly allocated to the left and right position on the screen. Item selection (after the memory probe) could thus take place between two visual locations and between two potential response hands.

Participants relied on detailed information of the probed memory item to guide their actions. **Figure 1b** shows response densities, which varied systematically as function of tilt direction (i.e., response hand) and magnitude (i.e., response duration). Reproduction errors (**Fig. 1c**) were on average 14.14 ± 0.84 (M ± SEM) degrees and it took participants on average 755.76 ± 52.29 ms to select the relevant item and to initiate the appropriate action (**Fig. 1d**).

We asked how and when visual representations and their corresponding actions become selected after the memory probe to support memory guided behaviour. We considered three alternative scenarios (**Fig. 1e-g**). (i) The brain may initially focus on only the visual information, and wait for the relevant visual representation to be selected before planning the appropriate action – yielding a serial pattern of visual-then-motor selection. Such a model is implicit in conventional studies of visual working memory in which memory items are deliberately isolated from particular actions during retention [4–6]. (ii) Alternatively, visual representations may be transformed into motor plans soon after encoding and become obsolete, such that the memory probe directly triggers motor selection (**Fig. 1f**). Such a model is implicit in conventional studies of motor working memory, in which actions are instructed by simple sensory cues regarding, for example, the location of a prospective reach or saccade target [7–9]. (iii) Finally, we hypothesized that, when potential actions rely on detailed sensory representations (such as precise visual orientation in our task), visual and motor representations may be held available jointly and thereby afford simultaneous (i.e. parallel, as opposed to serial) sensory and motor selection when an item becomes relevant for action (**Fig. 1g**).

To arbitrate among these scenarios, we capitalised on the high temporal resolution of electrophysiological brain recordings to track the unfolding of visual and motor selection during the post-probe time period of working memory utilisation (sometimes referred to as "output gating"; [11]). We focused on item location and response hand as the visual and motor attributes, because these are particularly tractable in non-invasive human electrophysiology. Because item location and required response hand were orthogonally manipulated across trials (a left/right item equally often required a response with the left or the right hand), we were able to independently characterise the selection of both attributes in the trial-average (thus circumventing any correspondence between our visual and motor selection signatures due to sheer volume conduction / signal mixing). We thus attribute lateralised patterns of neural activity that depend on item location (a purely visual attribute in our task) to visual selection and lateralised patterns that depend on required response hand to motor selection (followed by action implementation).

Two sets of complementary analyses converged on the notion of concurrent visual and motor selection during working-memory utilisation.

We first focused on hypothesized spectral modulations in selected visual and motor electrode clusters (Methods for details). **Figure 2a** shows the time- and frequency-resolved neural modulation relative to the probed item's location in visual electrodes. Visual selection was associated with a marked attenuation of 8-12 Hz alpha oscillations in electrodes contralateral (relative to ipsilateral) to the location of the probed item (**Fig. 2a**; cluster-P < 0.0001; Methods for details). This modulation occurred at posterior electrodes (**Fig. 2b**) and localised to visual and parietal brain areas (**Fig. 2c**). This agrees with previous reports of alpha attenuation during shifts of spatial attention in perception [12], working memory (reviewed in [13]), and long-term memory [14], although we are not aware this has

been demonstrated during working memory utilisation. It is noteworthy that alpha lateralisation reflecting selection of the spatial location of the probed item occurred despite item location not being strictly necessary for task performance (see also [15,16]). This ties in well with recent evidence that spatial location may play a grounding role for working-memory representations [6,17]. Provided that spatial location was a purely visual feature in our task (the required response depended on the probed item's tilt, not location) these data already allow us to reject the 'only motor' scenario in **Figure 1f**.

Analysis of spectral lateralisation relative to the required response hand yielded equivalent signatures of motor selection at central electrodes (**Fig. 2d** and **Fig. 2e**) and motor brain areas (**Fig. 2f**). The relative attenuation of power contralateral to the relevant response hand now also encompassed the expected 13-30 Hz beta band (**Fig. 2d**; cluster-P < 0.0001), in line with e.g. [18–20].

While the quantifications of visual and motor lateralisation were orthogonal and showed clearly distinct spatial maps (see also **Supplementary Fig. 1**), both effects occupied highly similar time ranges. We return to this central observation in more detail later. To gain confidence that both these spectral signatures captured neural processes relevant to selection, we confirmed that they each emerged before response initiation (**Supplementary Figs. 2 and 3**).

In a separate analysis, we applied time-resolved multivariate classification (as in [21,22]) to decode item location and required response hand from the broadband EEG signal (i.e. from the time-varying voltage signal) – combining high sensitivity with excellent temporal resolution [23]. Indeed, because no windowing is required for such time-domain analyses, this approach allowed us to track visual and motor selection at even finer temporal resolution. This further complements the spectral (frequency-domain) analysis by focusing on evoked (as opposed to induced) neural activity. **Figure 3a** shows that we could robustly decode both item location (cluster-P < 0.0001) and response hand (cluster-P < 0.0001) starting around 200 ms after probe onset (well before response initiation; **Fig. 1d;** and equivalent results were observed for a number of decoding metrics; **Supplementary Fig. 4**). To track the spatial-temporal trajectories of neural information linked to item-location and response-hand selection, we also applied this analysis iteratively for subsets of electrodes and source parcels in a searchlight fashion (as in [22,24]). As shown in **Figure 3b**, item-location information started in posterior (putative visual) electrodes and soon spread out more widely. Motor information also spread out with time, but remained most prominent in central (putative motor) electrodes. A source analysis substantiated the respective contributions of visual and motor brain areas for initial item-location and response-hand decoding (**Fig. 3c,d**). This corroborates the independent nature of our visual and motor decoders (given that item location and response hand were orthogonally manipulated). A cross-generalisation analysis, in which decoding of item location generalised between left and right response hand trials and in which decoding of response hand generalised between left and right item location trials (**Supplementary Fig. 5**), further reinforced the unique ability of our design and analysis to measure visual and motor selection effectively and independently. Moreover, we found highly similar decoding of the item location (as well as alpha modulations) in a re-analysis of our prior dataset in which participants performed a similar visual working-memory task, but always responded with the same, dominant hand (**Supplementary Fig. 6**).

Thus, when memoranda in working memory have both visual and motor attributes, utilisation involves selection of both types of attributes, engaging visual and motor brain areas at highly similar intervals after the memory probe. To reveal the temporal relation between visual and motor selection with greater granularity, we reduced the visual and motor spectral lateralisations to simple time courses, and normalised all relevant time courses (for the spectral and the decoding signatures) to their peak value. This confirmed highly overlapping temporal profiles of visual and motor

selection, and this was the case for both types of identified neural signatures (**Fig. 4a**). The results showed no systematic lag indicative of a serial pattern of visual-then-motor selection. If anything, motor selection appeared to start even earlier than visual selection in the spectral signatures (although we note this slight temporal offset disappeared when aligning the data to response initiation; **Supplementary Fig. 3**).

It is conceivable that the observed temporal correspondence at the group-average time courses was merely a coincidence, resulting from averaging slow participants who mainly showed the visual selection signatures with fast participants who mainly showed the motor selection signatures. To rule out this interpretation, we also used a complementary visualisation of the temporal correspondence by calculating cross-correlations between the time courses of visual and motor selection per participant and averaging the resulting coefficients. This confirmed maximal correlations at approximately zero-lag in the vast majority of participants, as well as in their average (**Fig. 4b**). Another possibility is that some items (or some trials) engage a predominant visual memory code, whereas others a predominant motor code. Under this account, one may expect that responses would be faster following probes of motor-coded vs. visually-coded items – and hence that faster trials would appear like the only motor scenario, whereas the slower trials would look more like the visual-then-motor scenario. Yet, we found that the concurrent nature of visual and motor selection was largely invariant to response onset time, with both selection signatures scaling similarly with response times (**Supplementary Fig. 7a,b**). As a bonus, this analysis revealed that faster trials were characterised by neural lateralisation patterns that favoured the subsequently probed item – in line with the notion of a 'spontaneous' prioritisation ("pre-selection") of that item (**Supplementary Fig. 7c**). This was the case for the spontaneous prioritisation of the item's visual location (lower contra vs. ipsilateral alpha power in visual sites; a replication of[16]) – further arguing against the 'only motor' scenario – as well as its associated action (lower contra vs. ipsilateral beta power in motor sites relative to the response hand associated with the probed item) – placing the motor contribution to visual working memory utilisation also into the delay period.

To quantify the timings of visual and motor selection more formally, we applied a jack-knife approach to obtain temporal confidence intervals (as in [22,25]). **Figures 4c** shows the relevant 95% confidence intervals sampled from 10 to 90 % of the identified peak value in each time course. For comparison, the red points depict the predicted timings of motor selection under a strict serial model (cf. **Fig. 1e**) in which motor selection starts when visual selection peaks. Clearly, the observed motor selection signatures occurred much earlier than predicted from this serial model. When expressed in t-values (**Fig. 4d**), the observed motor selection time courses occurred highly significantly earlier than predicted by the serial null-model (red points in **Fig. 4d**; average statistic across all of the 'percentage-of-peak' slices for the spectral data: $t_{avg}(24) = -11.652$, $p_{avg} = 1.705e\text{-}6$; decoding data: $t_{avg}(24) = -10.871$, $p_{avg} = 2.13e\text{-}8$), whereas the timing of motor selection was never significantly later than predicted by the parallel null-model (black points in **Fig. 4d**; average statistic for spectral data: $t_{avg}(24) = -1.68$, $p_{avg} = 0.175$; decoding data: $t_{avg}(24) = 0.493$, $p_{avg} = 0.648$).

Of course, it is impossible to rule out all viable 'in-between' models in which visual and motor selection are only slightly offset in time. However, if present, such delays are minimal compared to the time over which the neural signatures of visual selection evolve. Thus, even if visual and motor selection are not initiated in perfect synchrony, they clearly overlap during their operation, yielding concurrent availability of selected visual and motor attributes to guide performance.

Discussion | While memories inherently regard the past, the purpose of holding detailed sensory information available in memory is to guide adaptive future behaviour [26–30]. In this light, it is surprising that popular laboratory tasks of working memory have tended to consider visual and motor representations in isolation. Although we have gained a vast body of knowledge about

working memory from studies that have focused primarily on visual [2,3] or motor [7,31] contents of working memory, our data make clear that the brain's natural tendency may be to link prospective manual actions to particular sensory representations during working memory. Thus, unlike in perception – in which sensory analysis necessarily precedes action selection –, once sensory information has been encoded into working memory, the brain no longer needs to wait for the relevant sensory representation to be selected before considering the appropriate action. This sensory-motor conceptualisation of working memory may also account for the observation of similar capacity limits for visual working memory [5] and action planning [32].

Previous empirical [8,33–35] and theoretical [36] work have suggested that the brain continuously specifies multiple potential actions in parallel before selecting among them. Our data suggest that such parallel action specification may also occur for the contents of 'visual' working memory and incorporate visual representations. Concurrent availability of both visual and motor memory attributes allows refinement of selected actions by detailed visual memory content, yielding action implementation that is both fast (compared to the visual-then-motor scenario) and precise (compared to the only-motor scenario).

The current work uniquely targeted the selection of memories from their putative stores in visual and motor brain areas. Complementary research has posited a key role for frontal-striatal circuits in controlling the selection of information from working memory [10,11]. Whether sensory and motor attributes of memories are jointly or independently represented in these 'control circuits', and how these circuits interact with the traces in the sensory and motor areas that we studied remain exciting avenues for future research.

The sensory-motor conceptualisation of working memory promoted by the current work complements, and is to be distinguished from, prior work implicating a role for the brain's oculomotor system in visual-spatial working memory [37–40] – a role that is likely mediated by the involvement of this system in covert spatial attention [41–43]. While oculomotor-driven attention mechanisms may also contribute to our visual selection signatures (as these depended on the spatial location of the probed memory item), our study uniquely also targeted the process of guiding a manual action by the memorised shape of the selected memory item (independently of its location). In this light, our data are thus compatible with the concurrent co-activation of two computations that may each be sensory-motor in nature – one dealing with the selection of the relevant visual shape information through its memorised location, and the other dealing with the use of this information for guiding manual action.

In everyday situations, memorised visual information often guides action – such as when navigating to one's bed after turning off the lights, or when changing lanes after having scanned surrounding traffic. To target the essential elements of how visual working memory guides action, we developed a laboratory task with relatively simple visual stimuli and actions. This enabled us to measure visual memory guided action with high precision, and to track the dynamics of visual and motor selection independently in the EEG. As such, we believe our task and results provide an important step toward bridging the literature on visual working memory and motor control. Still, we only tested the relative timing of selecting a narrow set of visual attributes and motor responses. It will be important for future studies to address the generalisability of the current findings to different types of visual stimuli and actions, and to start exploring the links between visual working memory and action in more naturalistic situations.

We finally note that the construct of working memory serves as a central component in many theories of cognitive and brain function – as well as dysfunction. The success and reach of such theories, and of related cognitive therapeutic interventions, will ultimately depend on the validity

and breadth of our understanding of working memory. Our data highlight the importance of considering its fundamentally prospective, goal-oriented, nature for which the efficacy of utilisation may be at least as important as the much more commonly considered capacity of retention.

Accession codes | All data are publically available through the Dryad Digital Repository at: [YET TO INSERT ACCESSION CODE/DOI].

Acknowledgements | This research was funded by a Marie Skłodowska-Curie Fellowship from the European Commission (ACCESS2WM) to F.v.E., a Wellcome Trust Senior Investigator Award (104571/Z/14/Z) and a James S. McDonnell Foundation Understanding Human Cognition Collaborative Award (220020448) to A.C.N., a Medical Research Council Career Development Award (MR/J009024/1) and James S. McDonnell Foundation Scholar Award (220020405) to M.G.S, and by the NIHR Oxford Health Biomedical Research Centre. The Wellcome Centre for Integrative Neuroimaging is supported by core funding from the Wellcome Trust (203139/Z/16/Z).

Author Contributions | F.v.E. designed and programmed the experiment, acquired, analysed and interpreted the data, and drafted and revised the manuscript; S.R.C. acquired and interpreted the data; M.G.S. interpreted the data and drafted and revised the manuscript; A.C.N. designed the experiment, interpreted the data, and drafted and revised the manuscript.

Competing Interests Statement | The authors declare no competing interests.

References |

1. Baddeley, A. Working memory. *Science.* **255,** 556–559 (1992).
2. D'Esposito, M. & Postle, B. R. The cognitive neuroscience of working memory. *Annu. Rev. Psychol.* **66,** 115–142 (2015).
3. Pasternak, T. & Greenlee, M. W. Working memory in primate sensory systems. *Nat. Rev. Neurosci.* **6,** 97–107 (2005).
4. Harrison, S. A. & Tong, F. Decoding reveals the contents of visual working memory in early visual areas. *Nature* **458,** 632–5 (2009).
5. Luck, S. J. & Vogel, E. K. The capacity of visual working memory for features and conjunctions. *Nature* **390,** 279–281 (1997).
6. Schneegans, S. & Bays, P. M. Neural architecture for feature binding in visual working memory. *J. Neurosci.* **37,** 3913–3925 (2017).
7. Snyder, L. H., Batista, A. P. & Andersen, R. A. Coding of intention in the posterior parietal cortex. *Nature* **386,** 167–70 (1997).
8. Cisek, P. & Kalaska, J. F. Neural correlates of reaching decisions in dorsal premotor cortex: Specification of multiple direction choices and final selection of action. *Neuron* **45,** 801–814 (2005).
9. Li, N., Daie, K., Svoboda, K. & Druckmann, S. Robust neuronal dynamics in premotor cortex during motor planning. *Nature* **532,** 459 (2016).
10. Rowe, J. B., Toni, I., Josephs, O., Frackowiak, R. S. J. & Passingham, R. E. The prefrontal cortex: response selection or maintenance within working memory? *Science.* **288,** 1656–1660 (2000).
11. Chatham, C. H., Frank, M. J. & Badre, D. Corticostriatal output gating during selection from working memory. *Neuron* **81,** 930–942 (2014).
12. Worden, M. S., Foxe, J. J., Wang, N. & Simpson, G. V. Anticipatory biasing of visuospatial attention indexed by retinotopically specific alpha-band electroencephalography increases over occipital cortex. *J. Neurosci.* **20(RC63),** 1–6 (2000).
13. van Ede, F. Mnemonic and attentional roles for states of attenuated alpha oscillations in perceptual working memory: a review. *Eur. J. Neurosci.* **48,** 2509–2515 (2018).
14. Waldhauser, G. T., Braun, V. & Hanslmayr, S. Episodic memory retrieval functionally relies on

very rapid reactivation of sensory information. *J. Neurosci.* **36,** 251–260 (2016).
15. Kuo, B.-C., Rao, A., Lepsien, J. & Nobre, A. C. Searching for targets within the spatial layout of visual short-term memory. *J. Neurosci.* **29,** 8032–8038 (2009).
16. van Ede, F., Niklaus, M. & Nobre, A. C. Temporal expectations guide dynamic prioritization in visual working memory through attenuated α oscillations. *J. Neurosci.* **37,** 437–445 (2017).
17. Foster, J. J., Bsales, E. M., Jaffe, R. J. & Awh, E. Alpha-band activity reveals spontaneous representations of spatial position in visual working memory. *Curr. Biol.* **27,** 3216–3223 (2017).
18. Salmelin, R. & Hari, R. Spatiotemporal characteristics of sensorimotor neuromagnetic rhythms related to thumb movement. *Neuroscience* **60,** 537–50 (1994).
19. Neuper, C., Wörtz, M. & Pfurtscheller, G. ERD/ERS patterns reflecting sensorimotor activation and deactivation. *Prog. Brain Res.* **159,** 211–22. (2006).
20. Jenkinson, N. & Brown, P. New insights into the relationship between dopamine, beta oscillations and motor function. *Trends Neurosci.* **34,** 611–8 (2011).
21. Wolff, M. J., Jochim, J., Akyürek, E. G. & Stokes, M. G. Dynamic hidden states underlying working-memory-guided behavior. *Nat. Neurosci.* **20,** 864–871 (2017).
22. van Ede, F., Chekroud, S. R., Stokes, M. G. & Nobre, A. C. Decoding the influence of anticipatory states on visual perception in the presence of temporal distractors. *Nat. Commun.* **9,** 1449 (2018).
23. Stokes, M. G., Wolff, M. J. & Spaak, E. Decoding rich spatial information with high temporal resolution. *Trends Cogn. Sci.* **19,** 636–638 (2015).
24. Kriegeskorte, N., Goebel, R. & Bandettini, P. Information-based functional brain mapping. *Proc. Natl. Acad. Sci. U. S. A.* **103,** 3863–8 (2006).
25. Miller, J., Patterson, T. & Ulrich, R. Jackknife-based method for measuring LRP onset latency differences. *Psychophysiology* **35,** 99–115 (1998).
26. Chun, M. M. & Jiang, Y. Contextual cueing: Implicit learning and memory of visual context guides spatial attention. *Cogn. Psychol.* **36,** 28–71 (1998).
27. Summerfield, J. J., Lepsien, J., Gitelman, D. R., Mesulam, M. M. & Nobre, A. C. Orienting attention based on long-term memory experience. *Neuron* **49,** 905–916 (2006).
28. de Vries, I. E. J., van Driel, J. & Olivers, C. N. L. Posterior α EEG dynamics dissociate current from future goals in working memory-guided visual search. *J. Neurosci.* **37,** 1591–1603 (2017).
29. Rainer, G., Rao, S. C. & Miller, E. K. Prospective coding for objects in primate prefrontal cortex. *J. Neurosci.* **19,** 5493–5505 (1999).
30. Myers, N. E., Stokes, M. G. & Nobre, A. C. Prioritizing information during working memory: beyond sustained internal attention. *Trends Cogn. Sci.* **21,** 449–461 (2017).
31. Svoboda, K. & Li, N. Neural mechanisms of movement planning: motor cortex and beyond. *Curr. Opin. Neurobiol.* **49,** 33–41 (2018).
32. Gallivan, J. P. *et al.* One to four, and nothing more: Nonconscious parallel individuation of objects during action planning. *Psychol. Sci.* **22,** 803–811 (2011).
33. Gallivan, J. P., Logan, L., Wolpert, D. M. & Flanagan, J. R. Parallel specification of competing sensorimotor control policies for alternative action options. *Nat. Neurosci.* **19,** 320–326 (2016).
34. Gallivan, J. P., Barton, K. S., Chapman, C. S., Wolpert, D. M. & Flanagan, R. J. Action plan co-optimization reveals the parallel encoding of competing reach movements. *Nat. Commun.* **6,** 7428 (2015).
35. Gallivan, J. P., Bowman, N. A. R., Chapman, C. S., Wolpert, D. M. & Flanagan, J. R. The sequential encoding of competing action goals involves dynamic restructuring of motor plans in working memory. *J. Neurophysiol.* **115,** 3113–3122 (2016).
36. Cisek, P. Cortical mechanisms of action selection: the affordance competition hypothesis. *Philos. Trans. R. Soc. B Biol. Sci.* **362,** 1585–1599 (2007).

37. Postle, B. R., Idzikowski, C., Della Sala, S., Logie, R. H. & Baddeley, A. D. The selective disruption of spatial working memory by eye movements. *Q. J. Exp. Psychol.* **59,** 100–120 (2006).
38. Theeuwes, J., Belopolsky, A. & Olivers, C. N. L. Interactions between working memory, attention and eye movements. *Acta Psychol.* **132,** 106–114 (2009).
39. Jerde, T. A., Merriam, E. P., Riggall, A. C., Hedges, J. H. & Curtis, C. E. Prioritized maps of space in human frontoparietal cortex. *J. Neurosci.* **32,** 17382–17390 (2012).
40. Merrikhi, Y. *et al.* Spatial working memory alters the efficacy of input to visual cortex. *Nat. Commun.* **8,** 15041 (2017).
41. Deubel, H. & Schneider, W. X. Saccade target selection and object recognition: Evidence for a common attentional mechanism. *Vision Res.* **36,** 1827–1837 (1996).
42. Nobre, A. C. *et al.* Functional localization of the system for visuospatial attention using positron emission tomography. *Brain* **120,** 515–533 (1997).
43. Moore, T., Armstrong, K. M. & Fallah, M. Visuomotor origins of covert spatial attention. *Neuron* **40,** 671–83 (2003).

Figure legends |

Figure 1 | Task, performance, and hypothetical scenarios. a) Participants saw two coloured bars and reproduced the tilt of the colour-matching item after a working memory delay. Bar tilt was directly linked to the required response hand, such that a leftward (rightward) tilted bar required a reproduction response with the left (right) index finger (Methods for details). Each trial contained one leftward and one rightward tilted bar, randomly allocated to the left and right positions on the screen – rendering item location and required response hand orthogonal across trials. **b)** Average response density (proportion of responses) as a function of the reported tilt and the tilt of the probed item. Zero degrees denotes vertical and negative (positive) values denote a leftward (rightward) tilt. **c**) Density of response deviation from the required tilt. Grey lines show individual participants, while the blue line shows the group average. **d)** Average density of response initiation times. Same conventions as in panel **c**. **e-g)** hypothetical patterns of visual and motor selection after the memory probe. Shadings in panels **b-d** represent ± 1 s.e.m, calculated across participants (n=25).

Figure 2 | Spectral signatures of visual and motor selection during working memory utilisation. **a)** Spectral lateralisation (contralateral minus ipsilateral) in selected visual electrodes relative to the memorised location of the probed item. The black outline indicates the significant cluster (two-sided cluster-based permutation test; n=25). Zero permutations yielded a larger cluster than in the observed data, yielding a $P < 0.0001$ (provided 10.000 permutations). **b)** Topography of the difference in 8-12 Hz alpha power in the 400 to 800 ms interval between trials in which the memory probe prompted the selection of the left or right item in memory (left minus right). Note that the probe itself was always central. **c)** Source-level contrast equivalent to the sensor-level contrast in panel **b.** For visualisation, only values were displayed that were at least 25% of the maximum/minimum value. **d)** Same conventions as in panel **a**, except lateralisation was calculated relative to the response hand associated with the probed item and is displayed for the selected motor electrodes. Cluster-$P < 0.0001$, following 10.000 permutations. **Supplementary Fig. 1** for the complementary motor lateralisation in selected visual electrodes, and visual lateralisation in selected motor electrodes. **e-f**) Same conventions as in panels **b** and **c**, except data were contrasted for 13-30 Hz beta power between left and right response hands. The depicted item-location and response-hand contrasts were orthogonal. Results in all panels depict the average across all participants (n=25).

Figure 3 | Time-resolved decoding of visual and motor selection during working memory utilisation. **a)** Time courses of decoding of the visual location and the response hand associated with the probed item. Decoding is expressed as the difference in Mahalanobis distance between matching and non-matching trial classes (non-matching minus matching to yield positive decoding values). All 61 EEG electrodes were included as the multivariate dimensions. Decoding was estimated separately for each time point. Horizontal lines indicate significant temporal clusters (two-sided cluster-based permutation tests; n=25). For both time courses, zero permutations yielded a larger cluster than in the observed data, yielding a $P < 0.0001$ (provided 10.000 permutations). Shading represents ± 1 s.e.m, calculated across participants (n=25). **b)** Topography of visual (item location) and motor (response hand) decoding as a function of time. Decoding topographies were constructed using an iterative 'searchlight' approach, considering each electrode with its immediately adjacent lateral neighbour(s). **c-d)** Source-level equivalents to the data in panel **b** (Methods for details). For visualisation, we only displayed values above 50% of the peak decoding value. Shadings in panel **a** represent ± 1 s.e.m. Results in all panels depict the average across all participants (n=25).

Figure 4 | Parallel, not serial, selection of visual and motor attributes from working memory. **a)** Normalised time courses of the identified spectral and decoding signatures of visual and motor selection. For the spectral data, the visual selection time course was obtained by averaging the item-location lateralisation in selected visual electrodes between 8-12 Hz, while the motor selection time course was obtained by averaging the response-hand lateralisation in selected motor electrodes between 13-30 Hz. To increase visibility, the group-average time courses were normalised as a percentage of their peak value. **b)** cross-correlation coefficients between the time courses of visual and motor selection, evaluated for each participant separately (grey lines) and subsequently averaged (black lines; n=25). **c)** Temporal 95% confidence intervals (Jack-knife-based latency analysis with n=25 participants) of visual and motor selection, estimated at 10 to 90% of the peak value in steps of 10%. The red data points indicate the predicted motor timing under the strict serial model in which motor selection starts when visual selection peaks (cf. **Fig. 1e**). **d)** Jack-knife-based T values of the temporal difference (per percentage-of-peak slice) of the time course of motor selection with that of visual selection (parallel null model) or that of visual selection shifted by its own duration (serial null model). Dashed vertical lines depict critical T boundaries. Shadings in panels **a** and **b** represent ± 1 s.e.m, calculated across participants (n=25).

Methods |

Ethics

This study complied with all relevant ethical regulations and was conducted in accordance with the declaration of Helsinki. Prior to the study, experimental procedures were reviewed and approved by the Central University Research Ethics Committee of the University of Oxford. Each participant provided written consent before participation, and was reimbursed £15/hour.

Participants

Twenty-five healthy human volunteers (11 male; age range 19-36; mean age 25.12 years) participated in the study. No statistical methods were used to pre-determine sample sizes but our sample size is similar to those reported in previous publications from the lab that focused on similar neural signatures (e.g. [16]). All participants had normal or corrected-to-normal vision. Two participants were left handed. Data from all participants were retained in the presented analysis.

Stimuli, task, and procedure

Participants sat in front of a monitor (100-Hz refresh rate) at a viewing distance of approximately 95 cm. Each trial (**Fig. 1a** for a schematic) contained two peripheral oriented bars. One bar was always placed to the left and the other to the right of the central fixation cross, and one bar was tilted leftward and the other rightward. Across trials, bar tilt was orthogonal to bar location – i.e., the left (right) position would equally often contain a leftward or a rightward tilted bar. Bars were centred at a viewing distance of 5.7 degrees visual angle, and were 5.7 degrees in length and 0.8 degrees in width. Bars were randomly tilted between 20 and 70 degrees (thus avoiding all tilts within 20 degrees from vertical and horizontal). Unlike tilt direction, tilt magnitude was drawn independently between both items. In each trial, bars were randomly allocated two unique colours out of a set of four – blue (RGB: 21, 165, 234), orange (RGB: 234, 74, 21), green (RGB: 133, 194, 18), purple (RGB: 197, 21, 234). Colours were drawn independently of bar location and tilt.

Bars were displayed for 250 ms and followed by a working-memory delay (randomly drawn between 2000 and 2500 ms) during which only the fixation cross remained on the screen. After the memory delay, the central fixation cross changed into the colour of either item (the memory probe). Until then, both items were equally likely to be probed. Participants were instructed to reproduce the tilt of the colour-matching item as accurately as possible. For reproduction, participants used the "\" and "/" keys on the keyboard, respectively, using their left and right index finger. Time between probe onset and response initiation was unlimited. Upon response initiation (with either key), a visual response dial appeared on the screen that always started in vertical position (indicated by the north and south handles of the dial, **Fig. 1a**). The response dial had the same diameter as the length of the visual bars and was always presented around fixation. A key press of the right (left) index finger initiated a clockwise (counter-clockwise) rotation of this dial, at the speed of 8 ms per degree (thus requiring 820 ms to bring the dial from vertical to horizontal). Participants released the key when the dial reached the desired tilt. Only one key could be used per response and key release terminated the response (no adjustments could be made). The dial could not be rotated beyond ± 90 degrees, as the response would be terminated by the computer program. As a consequence, a leftward (rightward) tilted bar could only ever be reported with a left (right) key press. Bar tilt was thus directly linked to the action that would be required if that bar would be probed. Visual feedback of the dial was included merely to aid participants' performance in reporting the memorised visual orientation. Dial-feedback was independent of the visual memory attribute that we focused on pertaining to the *memorised location* of the probed item.

Because bar tilt and bar location were independent across trials, the location of each item was orthogonal (across trials) to the location (response hand) of its associated action. This key feature of

our task allowed us to independently track neural activity related to the probed item's memorised location (the visual attribute of interest) and the response hand associated with this item (the motor attribute of interest), while bypassing the contribution of sheer volume conduction / signal mixing.

Participants received feedback immediately after response termination. The fixation cross turned green for 200 ms for reports within 20 degrees from the probed item, and red otherwise. The inter-trial-interval (from feedback offset to encoding onset) was randomly drawn between 500 and 800 ms.

Participants practiced the task for 5-10 minutes until they reported being comfortable with it. They then completed two consecutive sessions of one hour with a 15 minute break in between. Each session contained 10 blocks of 60 trials, yielding 1200 trials per participant. The location and the tilt (response hand) of the probed item were pseudo-randomised at the level of trials. This ensured that each condition (probed item left, tilt left; item left, tilt right; item right, tilt left; item right, tilt right) occurred equally often in each block of 60 trials.

A visual localiser was inserted between blocks during which participants were asked to relax while keeping fixation. Each localiser contained 40 bars (identical to the ones used in the task) that were sequentially presented for 100 ms at an inter-stimulus-interval randomly drawn between 400 and 500 ms. Each localiser stimulus was randomly allocated one of the four colours, randomly tilted, and randomly presented at the left or right item position.

Data collection and analysis were not performed blind to the conditions of the experiments.

Analysis of behavioural data

We quantified accuracy as the average absolute circular deviation between the probed item's tilt and the reported tilt, and response time as the interval between probe onset and response initiation. Only trials in which response initiation times were within 4 SD of the mean were considered. Response densities were quantified in bins of 10 degrees, sampled in steps of 5 degrees from -90 to +90 degrees. We separately considered items whose tilt fell between non-overlapping bins of 10 degrees (i.e., [-70 to -60], [-60 to -50], and so on). Response-time densities were quantified using 50-ms bins, sampled in steps of 25 ms from 0 to 1250 ms.

EEG acquisition and basic processing

Electroencephalography (EEG) was acquired using Synamps amplifiers and Neuroscan acquisition software (Compumedics Neuroscan, North Carolina, USA). We used a 61-channel set-up that followed the international 10-10 system for electrode placement. Data were referenced to the left mastoid during recording and re-referenced offline to the average of both mastoids. The ground was placed on the left upper arm. Two bipolar electrode pairs recorded EOG; one above and below the left eye (vertical EOG) and another lateral of each eye (horizontal EOG). During acquisition, data were filtered between 0.1 and 200 Hz, digitized at 1000 Hz, and stored for offline analysis.

Data were analysed in Matlab (MathWorks, Massachusetts, USA) using a combination of FieldTrip [44] and custom code. Data were down-sampled to 250 Hz and epoched relative to probe onset (from -1500 to +2500 ms) as well as response onset (from -2500 to +1000 ms). Ocular artifacts were removed from the data using independent component analysis (ICA). Relevant ICA components were detected through correlation with the horizontal and vertical EOG. For all sensor-level analysis, we applied a surface Laplacian transform [45] to increase spatial resolution.

We only considered trials in which participants pressed the correct key (which was the case in 92.07 ± 1.11 [M ± SE] % of all trials) and in which response times were within 4 SD of the mean. Remaining

trials with excessive EEG variance were rejected based on visual inspection. After trial removal, it was possible that trials in which item location and response hand were associated with the same or opposite side had become slightly over-represented in the data. To re-balance the data, we finally made sure that item location and required response hand were equally often in the same or the opposite side. Trial numbers were equated by randomly subsampling from the case with more trials. On average, 955 ± 25 (M ± SE) trials (ranging between 710 and 1114) were retained for analysis per participant.

Spectral analysis

Time-frequency analysis was based on a short-time Fourier transform of Hanning-tapered data. We estimated spectral power at frequencies between 2 and 50 Hz in 1-Hz steps, using a fixed 300-ms sliding time window that was advanced over the data in 10-ms steps. To zoom in on lateralised modulations in visual and motor electrodes, we contrasted time-frequency matrices in selected visual and motor electrode clusters between trials in which either the item or the response location was contra vs. ipsilateral to the electrode cluster. We expressed this as a normalised difference (i.e., ((contra-ipsi) / (contra+ipsi)) * 100) and averaged the result between left and right electrode clusters. We did this separately for the visual and motor electrode clusters. To obtain topographical maps of lateralisation, we also calculated separately for each electrode the normalised difference between left vs. right item location as well as left vs. right response hand.

For each participant, we determined four electrode clusters: left visual, right visual, left motor, right motor. Visual clusters were defined by contrasting the neural response induced by left vs. right visual stimuli that were part of a task-free localiser. Motor electrodes were selected based on the neural response locked to all left vs. right button presses. Per cluster, we always selected between two and four electrodes, based on visual inspection of the data at this selection stage. Although the use of participant-specific electrode selections increases sensitivity, our results are not contingent on this selection – equivalent results were obtained when using a generic set of a-priori defined electrodes (PO7 for left visual, PO8 for right visual, C3 for left motor, C4 for right motor).

To reduce time-frequency data to time courses, we averaged over the a-priori-defined 8-12-Hz alpha band for the item-location lateralisation in the visual electrode clusters and we averaged over the a-priori-defined 13-30-Hz beta band for the response-hand lateralisation in the motor electrode clusters.

Multivariate decoding analysis

Multivariate decoding was based on the broadband (0.1 to 30 Hz) evoked responses for which we applied two additional pre-processing steps that are conventional in the analysis of evoked activity: we subtracted a trial-specific 250-ms pre-probe baseline and we removed high-frequency noise by applying a low-pass filter with a 30-Hz cut-off.

Decoding was evaluated separately for each time sample and was based on the multivariate Mahalanobis distance metric in which electrodes serve as dimensions (as in [21,22]). Decoding relied on a leave-one-out procedure. For each trial, we calculated the Mahalanobis distance between that trial and the average of all remaining trials whose class was either matching or non-matching with the trial under investigation. Classes were defined once by matching/non-matching item location (yielding item-location decoding) and once by matching/non-matching response hand (yielding response-hand decoding). If the multivariate neural pattern contains information regarding the class under consideration, then the multivariate distances should be smaller to the average of the matching class compared to the non-matching class. To express decoding as a positive value, we therefore subtracted the non-matching from the matching distances and averaged this metric across trials.

To maximize sensitivity, our main decoding analysis was based on all 61 EEG electrodes. We additionally performed this analysis iteratively for subsets of electrodes to obtain decoding topographies (as in [22,24]). In each iteration, we centered our 'searchlight' on a different electrode that we considered together with its immediately adjacent lateral neighbour(s) – yielding subsets of 2 electrodes for all 'outer' electrodes, and subsets of 3 for all 'inner' electrodes.

To increase sensitivity and visualization, we lightly smoothed the trial-averaged decoding time courses for each participant using as Gaussian kernel with a standard deviation of 20 ms. We confirmed that this step was not essential and that qualitatively similar (albeit slightly more noisy) results were obtained when no smoothing was applied.

Source analysis

We placed a grid with 1-cm^3 spacing inside the generic MNI T1 template brain and used a boundary-element volume-conduction model [46] to describe how activity in each grid point projected to the electrodes positioned according to the generic 10-10 system. Before source analysis, data were re-referenced to a common average reference.

Spectral power was localized using a frequency-domain beamformer [47]. For each grid point, we calculated the normalised difference in power between trials in which the item/response hand was on the left vs. right – in the same way as we had done at the sensor-level. Separate analyses were run for the a-priori-defined 8-12 Hz alpha band and the 13-30 Hz beta band.

To evaluate decoding at the source level, we applied a time-domain beamformer [48] to obtain three spatial filters associated with each grid point. We then used the anatomical automatic labeling (AAL) atlas to allocate every spatial filter to its corresponding source parcel. To reduce dimensionality, for each parcel, we entered all allocated spatial filters to a singular value decomposition and retained the five components (spatial filters) with the largest singular value. These components were used to re-construct five virtual channels per parcel, the time courses of which were entered into our multivariate decoding analysis – yielding two decoding time courses per parcel (one for item location, one for response hand). For plotting, we averaged decoding within a desired time window and placed the resulting value in all grid points that belonged to that parcel.

The attribution of our selection signatures to 'visual' and 'motor' was based on the *information* that was considered in our experimental contrasts and decoder (item location, response hand), not *where* this information localised to in the brain. Source reconstructions were included with the primary purpose to evaluate the 'plausibility' [49] of these signatures.

Statistical analysis and latency quantification

Statistical analysis involved two steps: (1) evaluating the identified neural signatures of visual and motor selection and (2) quantifying their temporal relationship.

For step 1 we used a cluster-based permutation approach [50] that is ideally suited for evaluating the reliability of neural patterns at multiple neighbouring data points – as in our case along the dimensions of time and, for the spectral analysis, also frequency. This approach effectively circumvents the multiple-comparisons problem by evaluating clusters in the observed group-level data against a single permutation distribution of the largest clusters that are found after random permutations (or sign-flipping) of the trial-average data at the participant-level. We used 10,000 permutations and used Fieldtrip's default cluster-settings (grouping adjacent same-signed data points that were significant in a mass univariate t-test at a two-sided alpha level of 0.05, and defining cluster-size as the sum of all t-values in a cluster). The p-value for each cluster in the non-

permuted data is calculated as the proportion of permutations for which the size of the largest cluster is larger than the size of the considered cluster in the non-permuted data. When zero permutations yield a larger cluster (as was the case for all our analyses), this Monte Carlo p value is thus smaller than 1/N-permutations (in our case $P < 0.0001$). We applied this approach to the time-frequency maps and to the decoding time courses. Topographical and source analyses served only to verify the plausibility of the identified patterns [49] and were not subjected to further statistical evaluation.

Step 2 involved two analyses. First, we calculated cross-correlation coefficients (using the xcov function in Matlab) between the identified time courses of visual and motor selection. We did this separately for each participant and averaged the resulting coefficients. The main purpose of this complementary visualisation of the data was to rule out that the temporal relations observed at the group-level may not be representative because different participants may drive the timing of the different time courses. Second, we used a jack-knife approach to obtain temporal confidence intervals (as in [22,25]). To increase transparency and avoid the arbitrary selection of a particular aspect of each time course (its onset, peak, midway point, and so on), we always considered several 'slices' of each time course, ranging from 10% of the peak value to 90% of the peak value in steps of 10%. For the jack-knife quantification, we iteratively removed one participant from the participant pool and, for each slice, compared the time at which that slice-value first occurred to the time that was observed when all participants were included. The jack-knife-based estimate of the temporal standard error then allowed us to obtain 95% confidence intervals under the student's t-distribution [25]. We also obtained confidence intervals for the temporal offset between the time courses of visual and motor selection (motor minus visual). Comparing this offset to 0 entailed a test against the parallel null model. A test against the serial null model was provided by the comparison of this offset to the predicted temporal shift if motor selection would start when visual selection peaked. This shift was determined by the start-to-peak (i.e., 10-100%) duration of the visual selection time course. Effectively, we thus compared the motor selection time course once to the visual selection time course as observed, and once to this time course shifted by its own duration.

All reported measures of spread involve ± 1 SEM, calculated across participants (n=25). All inferences were two-sided at an alpha level of 0.05 (0.025 per side). Data distributions were assumed to be normal but this was not formally tested. We report a single experiment with 25 participants that was not repeated.

Please also see our "Life Sciences Reporting Summary" for complementary methods information.

Data Availability Statement | All data are publically available through the Dryad Digital Repository at: [YET TO INSERT ACCESSION CODE/DOI].

Code Availability Statement | Code will be made available by the authors upon request.

Methods-only References |

44. Oostenveld, R., Fries, P., Maris, E. & Schoffelen, J.-M. FieldTrip: Open source software for advanced analysis of MEG, EEG and invasive electrophysiological data. *Intell. Neurosci.* **2011,** 1:1--1:9 (2011).
45. Perrin, F., Pernier, J., Bertrand, O. & Echallier, J. F. Spherical splines for scalp potential and current density mapping. *Clin. Neurophysiol.* **72,** 184–187 (2018).
46. Oostendorp, T. F. & Oosterom, A. van. Source parameter estimation in inhomogeneous volume conductors of arbitrary shape. *IEEE Trans. Biomed. Eng.* **36,** 382–391 (1989).
47. Gross, J. *et al.* Dynamic imaging of coherent sources: Studying neural interactions in the human brain. *Proc. Natl. Acad. Sci.* **98,** 694–699 (2001).

48. Van Veen, B. D. & Buckley, K. M. Beamforming: a versatile approach to spatial filtering. *IEEE ASSP Mag.* **5,** 4–24 (1988).
49. van Ede, F. & Maris, E. Physiological plausibility can increase reproducibility in cognitive neuroscience. *Trends Cogn. Sci.* **20,** 567–569 (2016).
50. Maris, E. & Oostenveld, R. Nonparametric statistical testing of EEG- and MEG-data. *J. Neurosci. Methods* **164,** 177–90 (2007).