Behavioral/Cognitive

Induced Alpha Rhythms Track the Content and Quality of Visual Working Memory Representations with High Temporal Precision

David E. Anderson,[1,2] John T. Serences,[3,4] Edward K. Vogel,[1,2] and Edward Awh[1,2]

[1]Department of Psychology and [2]Institute of Neuroscience, University of Oregon, Eugene, Oregon 97403, [3]Department of Psychology and [4]Neuroscience Graduate Program, University of California, San Diego, San Diego, California 92110

Past work has suggested that neuronal oscillations coordinate the cellular assemblies that represent items in working memory (WM). In line with this hypothesis, we show that the spatial distribution of power in the alpha frequency band (8–12 Hz) can be used to decode the content and quality of the representations stored in visual WM. We acquired EEG data during an orientation WM task, and used a forward encoding model of orientation selectivity to reconstruct orientation-specific response profiles (termed channel tuning functions, or CTFs) that tracked the orientation of the memorandum during both encoding and delay periods of the trial. Critically, these EEG-based CTFs were robust predictors of both between- and within-subject differences in mnemonic precision, showing that EEG-based CTFs provide a sensitive measure of the quality of sensory population codes. Experiments 2 and 3 established that these EEG-based CTFs are contingent on the voluntary storage goals of the observer. When observers were given a postsample cue to store or drop the memorandum, the resulting CTF was sustained in the "store" condition and rapidly eliminated following the "drop" cue. When observers were instructed to store one of two simultaneously presented stimuli, only the stored item was represented in a sustained fashion throughout the delay period. These findings suggest that the oscillatory activity in the alpha frequency band plays a central role in the active storage of information in visual WM, and demonstrate a powerful approach for tracking the precision of on-line memories with high temporal resolution.

Key words: **attention; decoding; EEG; oscillatory activity; tuning function; working memory**

Introduction

An emerging hypothesis is that the cellular assemblies that represent items in working memory (WM) may be coordinated via the synchronization of neural activity in the alpha frequency band (8–12 Hz; Jensen et al., 2002; Sauseng et al., 2009; Palva et al., 2010). For example, using EEG and MEG data, Palva et al. (2010) demonstrated that interareal synchronization in a network anchored in the intraparietal sulcus–a cortical region commonly associated with the maintenance of on-line representations (Todd and Marois, 2004; Vogel and Machizawa, 2004; Xu and Chun, 2005)–predicted WM capacity. This study suggests a direct links between alpha band activity and the number of items stored in WM. Nonetheless, the hypothesis that WM storage is mediated by oscillatory activity in the alpha band makes another clear, and thus far untested, prediction. If the cellular assemblies representing specific items are synchronized in the alpha frequency band, then alpha oscillations should also track the content and quality of the stored information.

A growing body of evidence supports the hypothesis that stimulus-specific representations are held in WM via the recruitment of the same sensory regions that encoded the memoranda (Awh and Jonides, 2001; Jonides et al., 2005; Postle, 2006). For example, multiple studies have used fMRI to show that the spatial pattern of activity in visual cortex enables decoding of the specific feature values stored in WM (Harrison and Tong, 2009; Serences et al., 2009a), and predicts individual differences in WM precision (Emrich et al., 2013; Ester et al., 2013). However, while fMRI studies show that the content and quality of WM representations are tracked by activity in sensory cortex, the sluggish temporal resolution of this method is a significant limitation. Thus, we turned to EEG as an approach for tracking WM contents with high temporal precision, and for assessing the role of alpha oscillations in WM.

Recently, Garcia et al. (2013) used the spatial distribution of evoked EEG power and a forward encoding model (FEM) to reconstruct orientation-selective channel tuning functions (CTFs) that predicted target discrimination performance. Here, we show that a similar analytic approach can be used to track the content of on-line memory by focusing on the spatial distribution of oscillatory power in frequency bands previously implicated in WM storage. We measured the spatial distribution of

Received Jan. 21, 2014; revised April 2, 2014; accepted April 23, 2014.

Author contributions: D.E.A. and E.A. designed research; D.E.A. performed research; D.E.A. and J.T.S. analyzed data; D.E.A., J.T.S., E.K.V., and E.A. wrote the paper.

This work is supported by National Institute of Mental Health R01-MH087214 (E.A. and E.K.V.).

The authors declare no competing financial interests.

Correspondence should be addressed to either David Anderson or Edward Awh, Department of Psychology, 1227 University of Oregon, Eugene, OR, 97403, E-mail: danders3@uoregon.edu or awh@uoregon.edu.

DOI:10.1523/JNEUROSCI.0293-14.2014

Copyright © 2014 the authors 0270-6474/14/347587-13$15.00/0

oscillatory power across a range of frequency bands while subjects were engaged in an orientation WM task. Although both evoked and induced activity across a range of low-frequency bands tracked the sample orientation during the initial encoding period, only induced activity in the alpha band enabled sustained tracking of the stored orientation during the memory delay period. Moreover, CTFs measured during the delay period were robust predictors of both within- and between-subject differences in memory quality. Thus, these findings provide new evidence linking alpha oscillations with the content and quality of on-line memory representations.

Materials and Methods

Participants. A total of 65 (26, 10, and 29 in Experiments 1, 2, and 3, respectively) undergraduates at the University of Oregon completed the experiment for monetary compensation ($10 per hour). All participants (47% female) self-reported normal or corrected-to-normal visual acuity and gave informed consent according to procedures approved by the University of Oregon Institutional Review Board.

Stimuli displays. Stimuli were generated in MATLAB (MathWorks) using the Psychophysics Toolbox extension (Brainard, 1997; Pelli, 1997) and presented on a 17 inch flat cathode ray tube computer screen (refresh rate of 120 Hz). Viewing distances were ~100 cm. Stimuli were rendered against a gray screen (RGB = [125 125 125]), and participants fixated on a central black dot that subtended 0.3° × 0.3° of visual angle.

Participants were instructed to remember the orientation of a solid bar spanning the diameter of a centrally presented ring (see Fig. 1*A*). The length (diameter of ring) and width (thickness of ring) of the bar were 7.1 and 0.6° of visual angle. In Experiments 1 and 2, the ring and oriented bar were rendered in black; in Experiment 3, two bars (one target, one distractor) were rendered in isoluminant blue (RGB = [0 150 200]) and green (RGB = [0 162 0]), color assignment was counterbalanced across subjects, and the ring was rendered in dark gray (RGB = [100 100 100]).

For each trial, stimulus orientation was randomly sampled from one of eight orientation bins spanning 0–157.5° in steps of 22.5°, with jitter added (±11.25°) to cover all 180° of unique orientation space and preclude categorical coding of the orientations. Orientation bins were balanced within each block. In Experiment 3, orientation differences between the target and distractor stimulus were fully counterbalanced within each block.

At the end of each trial, participants were cued by a change in fixation color to recall the orientation of the target item by using a mouse to click on a point along the perimeter of the ring corresponding to the target orientation.

Procedures. All experiments took ~2 h to complete, and were composed of 15 blocks of 64 trials each. The events in a single trial of Experiment 1 went as follows (Fig. 1*A*). First, subjects were prompted to press the spacebar to initiate the trial. Once the trial was initiated, a 1500 ms blank interval was followed by the presentation of the sample stimulus for 250 ms. A1250 ms delay period followed the offset of the sample stimulus. After the delay, fixation color changed to white, cueing subjects to make an unspeeded response by using a mouse to click the perimeter of the ring to indicate the orientation of the sample stimulus. The trial events in Experiments 2 and 3 were identical with those in Experiment 1, with the following exceptions. In Experiment 2, subjects were given a postcue (change in fixation color to blue or green, counterbalanced across subjects) to either store or drop the sample stimulus; no probe display appeared after drop trials. In Experiment 3, half of the subjects were instructed to always remember the blue target stimulus and ignore the green stimulus, while the other half of subjects stored the green stimuli; orientation differences between the target and distractor stimulus were fully counterbalanced within each block, and target color mapping was counterbalanced across subjects.

Modeling response error distributions. Response error values were defined by the difference between the subjects' response and the orientation of the sample stimulus (ranging from −90 to 90°). Maximum-likelihood estimation was used to fit a von Mises distribution (the circular analog of a Gaussian distribution) to the distribution of response error values. Two parameters were estimated: μ, the mean of the von Mises distribution, and s.d., the dispersion of the von Mises distribution (smaller values correspond to higher mnemonic precision).

EEG acquisition and analysis. EEGs were recorded using our standard recording and analysis procedures, including rejection of trials contaminated by blocking, blinks, or large (>1°) eye movements (Vogel et al., 1998; McCollough et al., 2007). We recorded from 22 tin electrodes mounted in an elastic cap (Electro-Cap International) using the International 10/20 System. 10/20 sites F3, FZ, F4, T3, C3, CZ, C4, T4, P3, PZ, P4, T5, T6, O1, and O2 were used along with five nonstandard sites: OL midway between T5 and O1; OR midway between T6 and O2; PO3 midway between P3 and OL; PO4 midway between P4 and OR; and POz midway between PO3 and PO4. All sites were recoded with a left-mastoid reference, and the data were re-referenced off-line to the algebraic average of the left and right mastoids. Horizontal electro-oculogram (EOG) was recorded from electrodes placed ~1 cm to the left and right of the external canthi of each eye to measure horizontal eye movements. The EEG and EOG were amplified with an SA Instrumentation amplifier with a bandpass of 0.01– 80 Hz and were digitized at 250 Hz in LabVIEW 6.1 running on a PC. To detect blinks, vertical EOG was recorded from an electrode mounted beneath the left eye and referenced to the left mastoid. Any trials containing either a blink or eye movement were excluded from further analysis. Subjects with trial rejection rates >20% were excluded from the sample (2, 0, and 4 subjects in Experiments 1, 2, and 3, respectively).

All analyses were performed using MATLAB in conjunction with the Signal Processing and EEGLAB Toolboxes. The raw EEG signal was bandpass filtered into a range of frequency bands using a two-way least-squares finite impulse response filter (eegfilt.m from EEGLAB toolbox; Delorme and Makeig, 2004). This filtering method uses a zero-phase forward and reverse operation, which ensures that phase values are not distorted, as can occur with forward-only filtering methods. A Hilbert Transform (MATLAB Signal Processing Toolbox) was then applied to the frequency-specific filtered waveforms to extract instantaneous power values. The Hilbert Transform produces the complex analytic signal, $z(t)$, of the filtered EEG, $f(t)$, where $z(t) = f(t) + i\tilde{f}(t) = A(t)e^{i\varphi(t)}$, from which instantaneous amplitude, $A(t)$, was extracted; $\tilde{f}(t)$ is the Hilbert Transform of $f(t)$ and $i = \sqrt{-1}$. Power was estimated at each time point using the following MATLAB syntax:

```
abs(hilbert(eegfilt(data,Fs,f1,f2))).^2,
where data is a 2D matrix of raw EEG (# of
trials × # of samples), Fs is the sampling
frequency (250 Hz), f1 is the lower bound of the
filtered frequency band, and f2 is the upper bound
of the filtered frequency band.
```

We examined feature-specific neural activity in the spatial distribution of both evoked and induced power across electrodes. Induced power was estimated by squaring the absolute value of the complex analytic signal obtained from the Hilbert transform for each trial, then averaging across all trials in each orientation bin within each block of trials. Evoked power was estimated by averaging the complex analytic signal obtained from the Hilbert transform across all trials in each orientation bin within each block, then squaring the absolute value of the averaged complex analytic signal associated with each orientation bin; this isolated oscillatory activity was phase locked to the onset of the sample stimulus. Thus, the output of the EEG analysis was comprised of an $n^*b \times m \times s$ matrix of both evoked and induced power values, where n is the number of orientation bins, b is the number of blocks, m is the number of electrodes, and s is the number of samples.

Multivariate pattern analysis. To determine whether patterns of oscillatory activity across the human scalp contained information about the remembered orientation, we used a linear classifier to determine if an unconstrained linear model could decode the content of working memory. Our reasoning was as follows: if a simple linear classifier fails to categorize the stored orientation with above-chance classification accuracy, then the FEM of orientation selectivity, which uses a constrained basis set–rather than a set of linear functions with unconstrained parameters–during model training, should similarly fail to decode content-specific information. Conversely, if

the linear classifier provides above-chance classification accuracy while the FEM fails to decode the content of subjective awareness, then we would conclude that the spatial distribution of oscillatory activity contains content-specific information, but the nature of this content-related neural signature is not well defined by our a priori model of feature selectivity.

To this end, we first filtered the raw EEG data into one of 27 frequency bands (4–30 Hz, in increments of 1 Hz), computed evoked and induced power, sorted power estimates from all trials into one of eight orientation bins, and computed mean power estimates for each orientation bin in each block (see above, EEG acquisition and analysis). Data from all but one block were then used to construct a "training" dataset that was used to train a linear discriminant function to discriminate between different stimulus conditions. The trained classifier was then used to decode the orientation label of each orientation bin within the remaining block (the "test" set). Data could be assigned to one of eight orientation labels, so chance discrimination accuracy was 12.5%. This analysis was iterated using a hold-one-out cross-validation procedure until data from every block had served as the test set. This analysis was completed for each frequency band and time point. Classification accuracies were then averaged across each test set, yielding a single classification accuracy value.

To evaluate the robustness of the magnitude of classification values obtained from this routine, we used a randomization routine that generates a null distribution of classification values for each frequency and time point by shuffling the labels of the orientation bin during training. This procedure allows us to approximate the significance of our empirically observed decoding accuracy against a hypothetical null distribution (see Randomization analyses, below).

Forward encoding model. To examine whether patterns of frequency-specific neural activity across electrodes contained information about the content and quality of the working memory representations, we used an FEM of orientation selectivity. Our approach is similar to one described by previous methods (Brouwer and Heeger, 2009, 2011; Serences and Saproo, 2012). Briefly, this model assumes that the pattern of responses across each electrode samples from a large number of orientation-selective neurons and that the response of any given electrode is proportional to the summed responses of all neurons contributing to the response of that electrode. Thus, one can characterize the orientation selectivity of a given electrode as a weighted sum of N orientation channels, each with an idealized tuning curve. We modeled the responses of each electrode using a basis set of eight half-sinusoids (one per sample angle) raised to the seventh power.

To estimate orientation selectivity with high temporal precision, we first bandpass filtered and independently estimated instantaneous evoked and induced power for frequency f (see above, EEG acquisition and analysis) for each orientation bin in each block across all 20 electrodes, then sorted the data into training and test sets. Each training set had 112 observations (14 blocks in training set × 8 orientation bins). Critically, every block sampled from each orientation bin equally. Similarly, the test set had eight observations (1 block in test set × 8 orientation bins). This analysis was completed for each frequency (e.g., α) and each time sample (e.g., 400 ms) time locked to the onset of the sample stimulus.

In the first phase of the analysis, data from the training set were used to estimate weights on the hypothetical orientation channels separately for each electrode. Briefly, let m be the number of electrodes, k be the number of hypothetical orientation channels (C_1, $k \times n_1$) composed of half-sinusoidal functions as the basis set discussed above, and n_1 and n_2 be the number of observations in the training and test sets, respectively. Let B_1 ($m \times n_1$) be the training set and B_2 ($m \times n_2$) be the test set (note that the structure of the training and test sets are identical across multivariate pattern analysis (MVPA) and FEM procedures). The training data in B_1 were then mapped onto the matrix of channel outputs in C_1 by the channel weight matrix (W, $m \times k$), which was estimated with GLM of the form:

$$B_1 = WC_1.$$

The channel weight matrix can be derived via least-squares estimation as follows:

$$W = B_1C_1^T(C_1C_1^T)^{-1}.$$

In the second phase of the analysis, channel responses (C_2, $k \times n_2$) were estimated based on the test data in B_2 with the weight matrix:

$$C_2 = (W^TW)^{-1}W^TB_2.$$

This process was then repeated by holding each block out in turn until each block had served as the test set. Following the procedure above, each estimated channel response function was circularly shifted so that the channel aligned with the stimulus that evoked the response was positioned in the center of orientation space, thereby aligning the estimated channel responses to a common center (i.e., 0° on the x-axis of the figures).

The channel weight matrix of the FEM computes the linear transformation of observed neural responses into the hypothetical space of the basis function (C_1). Therefore, because the parameters of the idealized tuning curve determine the shape of the basis set, the measurement unit of channel responses obtained in C_2 is arbitrary.

In Experiment 2, the training steps of the forward encoding procedure outlined above were collapsed across store and drop trials, because any idiosyncratic differences (other than orientation selectivity) in the spatial distribution of alpha power between storage conditions would have contributed independently to the estimation of channel weights; therefore, parsing each condition and performing the forward encoding procedure in isolation for each stimulus condition would have biased estimates of channel weights, thus obscuring the conclusion that differences in CTFs were attributable to systematic differences in feature selectivity. Following the training procedure, each estimated channel response function in C_2 was sorted based on condition, producing separate channel response profiles for store and drop conditions. In Experiment 3, the steps of the forward encoding procedure were identical to that outlined above, and the estimation of the channel response functions for the distractor item were estimated separately. This procedure would not introduce any bias in the channel weight matrix because orientation differences between target and distractor values were fully counterbalanced within each block.

Finally, to characterize the amplitude and dispersion of these functions, each participant's response profile was fitted with a Gaussian function of the form:

$$f(x) = ae^{\frac{(x-\mu)^2}{2\sigma^2}} + b,$$

where μ is the mean, σ is the dispersion (i.e., SD) of the distribution, and a and b are scaling factors that correspond to the amplitude and baseline, respectively. Our Gaussian fitting routine was completed in two steps. First, we created a set of coarsely defined seed Gaussian functions, each created from the parametric variation of all four fitting parameters, and calculated the sum of squared error (SSE) between the reconstructed CTFs and each seed Gaussian function. The seed parameters of the Gaussian providing the lowest SSE after the coarse analytic routine were fed into a least-squares nonlinear curve-fitting routine in MATLAB to further optimize the fit, with the following [lower, upper] boundary limits on each parameter: μ [−22.5°, 22.5°], σ [0 π], a [0 *Inf*]; no constraints were placed on b. a was taken as an estimate of response amplitude, and σ was taken as an estimate of dispersion.

Modeling of CTFs allowed us to decompose observed channel response profiles into discrete metrics of tuning amplitude and dispersion, each corresponding to multiplicative increases in channel responses relative to baseline activity and proportional differences in channel responses between neighboring orientation channels relative to baseline activity, respectively. Here, we will broadly refer to observed increases in on-channel activity coupled with decreases in off-channel activity (i.e., graded tuning profiles) as orientation-selectivity. We will then independently examine the contribution of tuning amplitude and dispersion to observed changes in orientation selectivity of CTFs.

Randomization analyses. We assessed the prediction that the spatial distribution of oscillatory power (across electrodes) contains content-specific information across a range of frequencies and time points. Given that the null distribution of our measures of classification accuracy and channel response functions may not be normal, we implemented a Monte Carlo randomization procedure to compute the deviation of observed ANOVA statistics from surrogate null distributions in the following manner. Note that the derived null distribution emerged from the analysis of surrogate time series data that share statistical properties with the original data. The initial processing steps of the randomization analyses were identical to those described in the EEG acquisition and analysis section, and data were partitioned into training and test sets (see above, FEM). Next, we used a shuffling procedure that allowed us to infer the chance distribution for observing above-chance decoding accuracy (MVPA) or robust orientation-selectivity (FEM). We shuffled the orientation labels (training labels in Multivariate pattern analysis; C_1, see above, FEM)–meaning that orientation labels were random with respect to the observed response in each electrode–during the estimation of channel weights (FEM) and training of the linear classifier (MVPA), and estimated classification accuracies and channel response functions normally from the surrogate channel weights. Finally, the surrogate data were submitted to an ANOVA, yielding a surrogate *F* statistic. Thus, this procedure effectively randomly samples test statistics associated with decoding accuracy and orientation selectivity from a hypothetical null distribution. To generate hypothetical null distributions, the randomization procedure outlined above was repeated 1000 times.

To determine the significance of the observed data against the surrogate null distribution, we calculated the probability of a surrogate data series producing an *F* statistic greater than or equal to the empirically observed *F* statistic (i.e., $F_{\text{null}} \geq F_{\text{obs}}$). Thus, for each effect, we report both the *F* statistic from the ANOVA on the observed data and the probability of observing type I error. For within- and between-subject analyses based on mnemonic precision, we report only results from the omnibus test, with degrees of freedom Greenhouse–Geisser corrected for violations of sphericity; these effects directly follow reported effects of channel modulations and storage requirements obtained from the randomization analyses.

Mutual information analysis. To examine which electrodes provided the most information about the stored feature values, we estimated electrode-based mutual information (MI) between the observed set of neural responses in each electrode and the presented set of orientation stimuli. Electrodes demonstrating sufficient heterogeneity (or selectivity) of neural responses–coupled with highly predictable mappings between stimulus values and induced oscillatory power–contribute more robustly to the selectivity of CTFs, leading to a higher estimate of MI. Thus, we were able to estimate MI as a function of electrode position for both evoked and induced oscillatory activity. Separate analyses were performed for the encoding (0–250 ms) and delay (250–1500) epochs of the trial.

MI is an information theoretic measure of general interdependence between random variables (Shannon, 1948). Here, we measured MI between the observed set of neural responses (X) and the set of stimulus orientations (Y). Broadly, the mutual information $I(X;Y)$ is defined as follows:

$$I(X;\, Y) = H(X) - H(X|Y),$$

where $H(X)$ is the entropy, or uncertainty, of a random variable, and $H(X|Y)$ is the conditional entropy between variables X and Y. If X (or Y) is a discrete variable, then $H(X)$ is defined as follows:

$$H(X) = -\sum_{X} p(X) log_2\, p(X),$$

and $H(X|Y)$ is defined as follows:

$$H(X|Y) = -\sum_{X,Y} p(X,Y) log_2\, p(X|Y)$$

where $p(X)$ represents the marginal probability distribution of X, $p(X,Y)$ represents the joint marginal probability distribution of X and Y, and $p(X|Y)$ represents the conditional marginal probability distribution of X and Y. The distribution of continuous oscillatory power (X) with kurtosis κ was discretized into a histogram of j bins as follows: $log_2 N + 1 + log_2(1 + \kappa\sqrt{N/6})$.

MI quantifies the reduction of uncertainty about the stimulus that can be gained from observation of a single trial of the neural response in units of bits. Thus, MI is zero when the stimulus and response are statistically independent quantities, indicating that no knowledge about the stimulus can be gained by observing the response.

Results

Experiment 1: decoding the content and quality of working memory representations

Time-frequency representation of evoked and induced power

We examined time-frequency representations (TFRs) of evoked and induced oscillatory power across a broad range of frequencies (4–50 Hz) by estimating proportional power modulations (percentage change) relative to prestimulus baseline activity (−500 to −100 ms) for each time point in the trial. Here, an increase in oscillatory power corresponds to the phase synchronization (or temporal alignment) of recorded neural ensembles; conversely, a decrease in oscillatory power corresponds to the phase desynchronization (or temporal disalignment) or recorded neural ensembles. Oscillatory activity was examined in posterior electrodes (O1, O2, OL, OR, PO3, PO4, POz, P3, P4, and Pz).

TFRs of evoked power (Fig. 1*B*) revealed strong synchronization of oscillatory activity between 8 and 40 Hz during stimulus encoding (0–250 ms). The magnitude of these stimulus-evoked power modulations were most pronounced in the alpha (8–2 Hz) and low beta (13–20 Hz) bands (Fig. 1*B*, black box; *C*). Conversely, we failed to observe sustained modulations of evoked oscillatory activity during stimulus maintenance (250–1500 ms).

TFRs of induced power (Fig. 1*D*) revealed strong synchronization theta (4–7 Hz) oscillatory activity during stimulus encoding. In contrast to the pattern of evoked activity, we observed sustained, delay-specific desynchronization of induced oscillatory activity in the theta, alpha (8–12 Hz), and low beta (13–20 Hz) frequency bands (Fig. 1*D*, black box; *E*).

Thus, we have demonstrated: (1) synchronized cortical activity in induced theta, evoked alpha, and evoked low beta oscillatory bands during stimulus encoding and (2) sustained desynchronized cortical activity in induced theta, induced alpha, and induced low beta oscillatory bands during stimulus maintenance.

Determining the frequency bands that contain stimulus-specific information

Our first goal was to identify the range of neural oscillatory frequencies that contained information about the memorandum in the spatial distribution of evoked (stimulus-driven) and induced (endogenous) power. Using frequency-tagged stimuli, Garcia et al. (2013) showed that the spatial distribution of EEG responses in the tagged frequency band could be used to decode the orientation of attended items. Motivated by the hypothesis that active storage in WM is mediated by synchronized neural activity in specific frequency bands, we sought to determine whether the spatial distribution of oscillatory power would also enable the decoding of stored feature values in the absence of external visual stimulation. Thus, we used a linear classifier across a broad range of frequencies (4–30 Hz) to determine which frequency bands contained information about the orientation of the memorandum. After identifying the informative frequency bands, we implemented an FEM of orientation selectivity to reconstruct

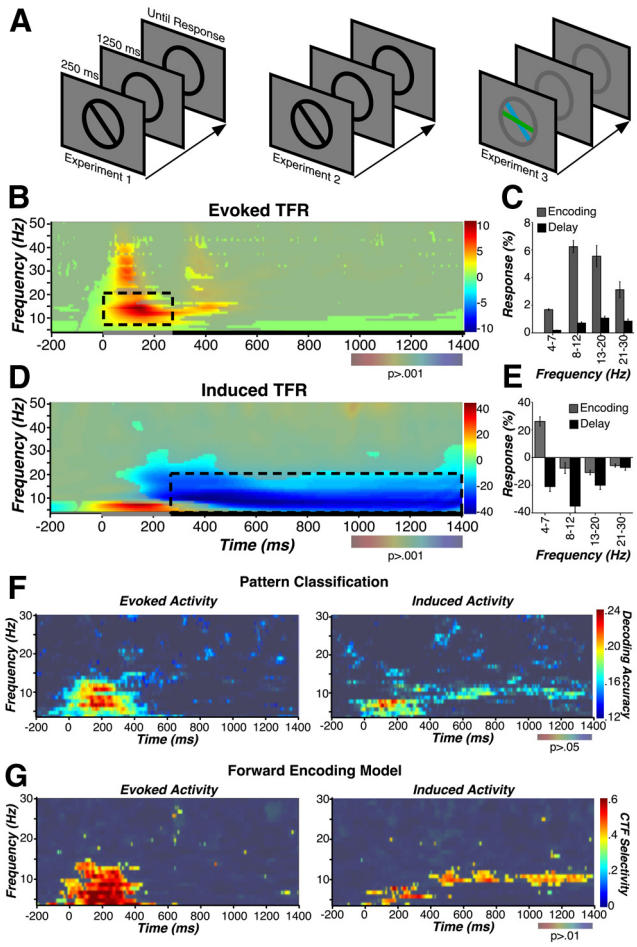

Figure 1. ***A***, Subjects maintained fixation and were instructed to remember the orientation of a sample stimulus; after a short delay, subjects were instructed to indicate the orientation of the sample stimulus by clicking on the perimeter of the rim. In Experiment 2, a change in fixation color (blue or green) immediately after sample offset indicated that the sample stimulus should be stored or dropped. In Experiment 3, subjects were instructed to remember the blue sample stimulus and ignore the green sample stimulus. ***B–E***, Oscillatory activity was normalized relative to a prestimulus baseline by calculating proportional changes in task-related oscillatory power. TFRs were collapsed across posterior electrodes (see Results). ***B***, TFRs revealed significant increases in evoked oscillatory power during stimulus encoding in low-frequency ranges. ***D***, TFRs revealed significant decreases in induced oscillatory power during stimulus maintenance. ***C***, ***E***, Average oscillatory power observed in evoked and induced activity, respectively, across frequency bands and task epochs. ***F***, Linear classifier performance based on orientation-selective patterns in the spatial distribution of evoked or induced oscillatory power across a broad range of individual frequencies (4–30 Hz in increments of 1 Hz) and time points. During encoding (0–250 ms), the spatial pattern of evoked and induced activity reliably decoded the content of WM in the theta (4–7 Hz) and alpha (8–12 Hz). During maintenance (250–1500 ms), only induced activity in the alpha band reliably decoded WM representations. ***G***, An FEM was used to generate orientation-selective CTFs based on the same information used for linear classification. CTF selectivity revealed a similar decoding profile as that observed in the classification routine.

orientation-specific response profiles, or CTFs. As the results will show, induced–but not evoked–activity from the same frequency bands that were found to contain stimulus-specific information via the linear classifier also revealed sustained orientation-selective CTFs during the delay period.

The results of the classification routine are shown in Figure 1*F*. During the initial encoding period (0–250 ms), the spatial distribution of evoked (left) and induced (right) power in the theta (4–7 Hz) and alpha (8–12 Hz) bands enabled reliable above-chance classification of the stored orientation (evoked theta: $t_{(23)} = 7.77$, $p < 0.001$; evoked alpha: $t_{(23)} = 4.64$, $p < 0.001$; induced theta: $t_{(23)} = 5.17$, $p < 0.001$; induced alpha: $t_{(23)} = 2.64$, $p < 0.05$). In contrast, during the delay period (250–1500 ms), only induced activity in the alpha band enabled above-chance classification ($t_{(23)} = 7.97$, $p < 0.001$); classification accuracy was statistically equivalent to chance in the theta (4–7 Hz; $t_{(23)} = 0.26$, $p = 0.80$), low beta (13–20 Hz, $t_{(23)} = 0.76$, $p = 0.46$), and high beta (21–30 Hz; $t_{(23)} = 0.06$, $p = 0.95$) frequency range. Evoked activity did not enable sustained above-chance classification in any frequency band (theta: $t_{(23)} = 0.07$, $p = 0.95$; alpha: $t_{(23)} = 1.3$, $p = 0.20$; low beta: $t_{(23)} = 0.18$, $p = 0.86$; high beta: $t_{(23)} = 0.31$, $p = 0.76$).

Modeling orientation-selective CTFs

Above-chance decoding accuracy in the alpha frequency band motivated the use of an FEM of orientation selectivity to reconstruct population-level, feature-selective tuning functions. Given that the classification routine demonstrated that stimulus information could be decoded from the spatial distribution of alpha activity, we sought to further quantify the nature of the orientation-specific information carried by low-frequency oscillatory activity. Specifically, a linear classifier is quite flexible in estimating categorical boundaries between stimulus conditions, providing a weaker understanding of how or why the contents of subjective awareness can be decoded from the spatial pattern of oscillatory activity. In contrast, the FEM uses a basis set of hypothetical, orientation-selective tuning functions to reduce the dimensionality of the data, and provides a quantifiable pattern of channel response estimates that can be used to estimate characteristics of the underlying population code, such as response amplitude and selectivity (dispersion). Thus, we used the FEM to more precisely characterize and quantify the feature-selective activity in the frequency bands of interest.

The results of the forward encoding analysis are shown in Figure 1*G*. Here, we used the slope of the channel response by orientation channel functions–collapsed across like channels (e.g., ±22.5 degrees) at each time point and frequency to quantify CTF channel modulation. Increases in CTF slope from zero (no change in channel response across orientation channels) to larger values indicate increases in the degree of channel modulation in feature-selective CTFs. CTFs based on evoked power (Fig. 1*G*, left) showed a strong increase in channel modulation during the initial encoding period (0–250 ms poststimulus) in the alpha ($F = 32.19$, $p = 0$) and theta ($F = 57.81$, $p = 0$) frequency bands, but not in the low beta ($F = 2.13$, $p = 0.09$) or high beta ($F = 0.84$, $p = 0.56$) frequency bands (Fig. 2*B*). Thus, evoked activity generated orderly CTFs only in the theta and alpha bands, and only early in the trial. We observed virtually no evidence of delay-specific (250–1500 ms) evoked CTFs in any frequency band (theta: $F = 0.99$, $p = 0.33$; alpha: $F = 1.57$, $p = 0.21$; low beta: $F = 1.20$, $p = 0.35$; high beta: $F = 1.27$, $p = 0.27$). To summarize, the same frequency bands and time points that enabled above-chance classification based on evoked power also enabled the reconstruction of orientation-specific CTFs.

Next, we examined which frequency bands would support the reconstruction of CTFs from the spatial distribution of induced power. Increased CTF channel modulation was evident throughout both the encoding and delay period of the trial in the alpha frequency band (Fig. 1*G*, right). The shapes of these delay-specific induced CTFs are depicted in Figure 2*D*. While we observed significant channel modulation during the delay period in the alpha frequency band ($F = 46.21$, $p = 0$), we failed to observe significant channel modulation in the theta ($F = 0.94$, $p = 0.36$),

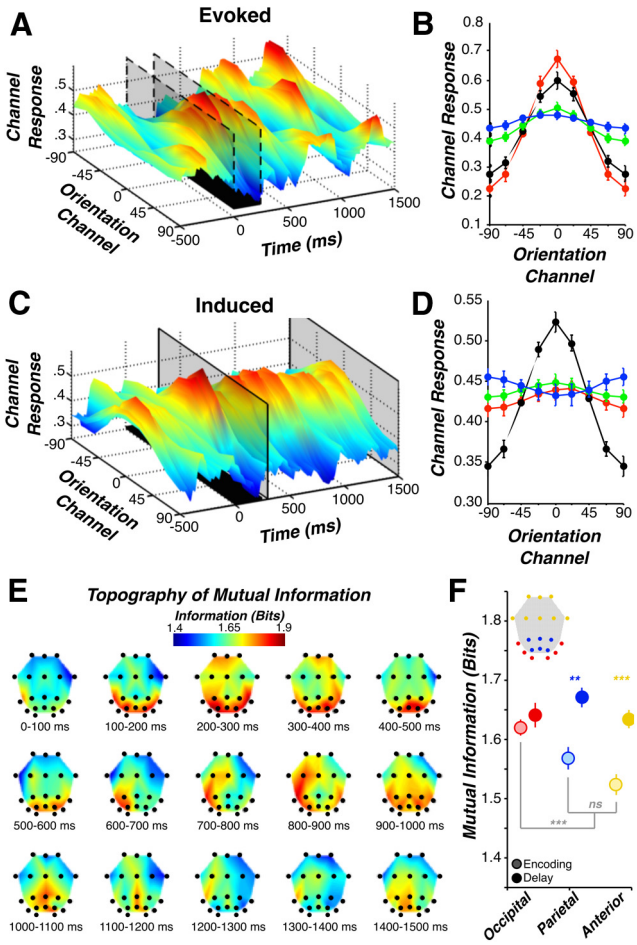

Figure 2. Orientation selectivity in oscillatory alpha activity. ***A***, Evoked alpha CTFs revealed a clear CTF during encoding (0–250 ms; boundaries of gray shaded window) that quickly dissipated after stimulus offset. ***B***, Average evoked CTFs during stimulus encoding for alpha activity (black), as well as theta (red), low beta (green), and high beta (blue) activity for comparison. ***C***, Induced alpha CTFs revealed a clear CTF during stimulus encoding and maintenance (250–1500 ms; boundaries gray shaded window). ***D***, Average induced CTFs during stimulus maintenance for alpha activity (black), as well as theta (red), low beta (green), and high beta (blue) activity for comparison. ***E***, Scalp topography of MI between presented stimulus orientations and induced alpha power across 100 ms bins. ***F***, For comparison between putative cortical regions, electrodes were sorted into occipital (red), parietal (blue), and anterior (yellow) topographic regions, and average MI was calculated for each region during encoding and delay epochs. Error bars represent 95% confidence interval.

low beta ($F = 1.29$, $p = 0.27$), and high beta ($F = 0.76$, $p = 0.46$) frequency bands. Thus, as we observed with the analysis of evoked activity, the same frequency band and time points that enable above-chance classification based on induced power also enabled the reconstruction of orientation-specific CTFs.

Visualizing CTF time course

Our initial analyses indicated that only the alpha frequency band enabled the reconstruction of orderly CTFs during the delay period of the trial. To visualize the time course of the CTFs derived from alpha band activity, we bandpass filtered in the alpha band (8–12 Hz) and constructed CTFs for each time point in the trial (Fig. 2*A*, *C*). The CTF based on evoked power emerged during the initial encoding period, and dissipated 250 ms after sample offset (Fig. 2*A*); accordingly, randomization analyses of the evoked CTF revealed significant channel modulation during the 0–250 ms encoding epoch ($F = 4.47$, $p = 0.018$), but not during the ensuing delay period (250–1500 ms; $F = 0.36$, $p = 0.25$). Thus, evoked alpha power carries content-related information only during stimulus encoding, suggesting that evoked CTFs reflect the early bottom-up, stimulus-driven response to the memorandum.

Next, we examined feature-selective population response profiles constructed from the spatial pattern of induced alpha activity (Fig. 2*C*). Consistent with our initial analysis of induced power, we observed sustained feature-selective tuning functions that spanned both the initial encoding period (0–250 ms) as well as the delay period (250–1500 ms). We estimated the average delay-specific CTF during the delay period (indicated by the boundaries of the gray shaded region in Fig. 2*C*), and observed significant channel modulation during this maintenance epoch ($F = 10.42$, $p = 0.001$). Thus, while both evoked and induced alpha activity track the sample orientation during the initial encoding period, only induced alpha activity tracks the content of WM throughout the delay.

Topography of orientation-selectivity during encoding and maintenance

The FEM estimates the channel-weighted linear sum of neural responses to each orientation stimulus across all electrodes, thus obscuring any topographic information regarding orientation selectivity. To examine which electrodes provided the most information about the stored feature values, we estimated electrode-based MI–measured in bits–between observed set of neural responses in each electrode and the presented set of orientation stimuli (see Materials and Methods, Mutual information analysis).

MI analyses of induced oscillatory alpha power revealed nonzero interdependence between the presented orientation stimulus and the observed neural response in each electrode during both stimulus encoding and maintenance (Fig. 2*E*), in line with the above-chance decoding that was possible in this frequency band. Interestingly, we observed nonuniformity in the magnitude of MI in orientation selectivity across broadly defined topographic sites (Fig. 2*F*). During stimulus encoding (light circles), MI was larger in occipital (red) electrodes than posterior (blue) and more anterior (yellow) electrodes ($t_{(23)} = 3.01$, $p = 0.006$), whereas MI was statistically indistinguishable between parietal and anterior electrodes ($t_{(23)} = 1.40$, $p = 0.17$). Conversely, delay-specific MI (dark circles) was fairly uniform across electrode sites ($p > 0.22$).

Induced alpha activity tracks the quality of working memory representations

Thus far, we have demonstrated that induced oscillatory activity specific to the alpha frequency band enables sustained tracking of the contents of working memory. Next we examined whether these CTFs also tracked variations in memory quality. Indeed, a median split based on WM precision–the dispersion of response errors for each observer–revealed systematic changes in the selectivity of CTFs based on induced activity from the delay period. This led to an interaction between channel activity and group (Fig. 3*D*; $F_{(1,22)} = 6.40$, $p < 0.05$) and a significant correlation between CTF dispersion and WM precision (Fig. 3*E*; $R^2 = 0.51$, $p < 0.001$). CTF amplitude did not predict WM precision (Fig. 3*F*; $R^2 = 0.04$, $p = 0.35$; cf., Ester et al., 2013). A similar analysis of the evoked alpha band activity shows no significant interaction between channel activity and group (Fig. 3*A*; $F_{(1,22)} = 1.43$, $p = 0.25$), and no correlation between CTF dispersion and WM precision (Fig. 3*B*; $R^2 = 0.01$, $p = 0.64$). Thus, CTF dispersion based on induced, but not evoked, EEG activity is a strong predictor of individual differences in WM precision. Moreover, CTF dispersion tracked within-subject variations in WM precision. A median split of trials based on response error magnitude showed

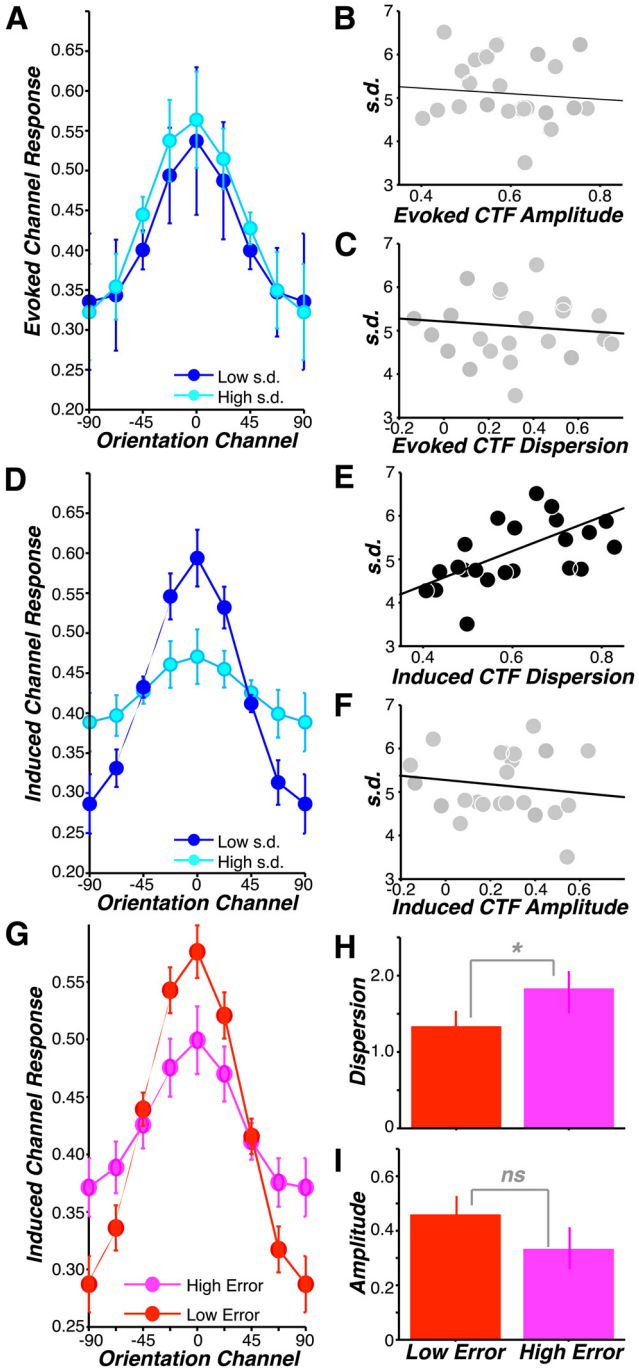

Figure 3. Between- and within-subject links with mnemonic precision. ***A***, ***D***, A median-split analysis based on between-subject differences in mnemonic precision (s.d.) was performed on evoked and induced CTFs. No group difference was observed in evoked CTFs (***A***), and no apparent link was observed between mnemonic precision and CTF dispersion (***B***; $R^2 = 0.01$, $p = 0.64$) or amplitude (***C***; $R^2 = 0.07$, $p = 0.22$). ***D***, A clear loss of selectivity was observed for low precision (dark blue) relative to high precision (light blue) subjects in induced CTFs ($p < 0.05$). A strong link was observed between mnemonic precision and induced CTF dispersion (***E***; $R^2 = 0.51$, $p < 0.001$), but not amplitude (***F***; $R^2 = 0.04$, $p = 0.35$). ***G***, A median-split analysis based on within-subject differences in response error. CTFs were generated from trials grouped into high error (low precision; pink line) or low error (high precision; red line) bins. A clear loss of selectivity was observed in high error trials. The observed loss of selectivity was attributable to a significant difference in CTF dispersion (***H***; $p < 0.05$), but not amplitude (***I***; $p = 0.18$).

greater CTF selectivity in low error trials compared with high error trials (Fig. 3*G*). This led to a significant interaction between error bin and response channel ($F_{(2.11,48.65)} = 6.33$, $p < 0.01$). Next, we fitted low and high error CTFs with a Gaussian function to examine whether the observed error-related difference in selectivity was due to changes in response amplitude or dispersion. Consistent with the pattern observed in the between-subject analysis, we observed larger dispersion estimates in low–relative to high–error trials (Fig. 3*H*; $t_{(23)} = 2.35$, $p < 0.05$), whereas we observed no difference in amplitude (Fig. 3*I*; $t_{(23)} = 1.4$, $p = 0.18$). Thus, the spatial distribution of induced alpha power predicts both within- and between-subject variations in WM precision.

Testing alternative explanations of the link between EEG activity and the contents of WM

Our working hypothesis is that these EEG-based CTFs reflect endogenous activity generated by the synchronization of feature-selective population codes. We considered two alternative explanations: (1) there is some inherent circularity in the forward encoding modeling analysis that forces any pattern of data to exhibit feature selectivity and (2) the spatial distribution of alpha activity is correlated with the magnitude of the EOG signal, reflecting minor changes in gaze direction yoked to the sample orientation.

To address the first concern, we first compared delay-specific CTFs obtained from Experiment 1 to a time window in the observed data that should demonstrate zero feature selectivity. Specifically, we measured CTFs observed during the baseline period, in which the observed baseline oscillatory activity should vary randomly with respect to the trial-specific orientation channel (i.e., observers have not been presented with the sample stimulus). If the analytic routine of the FEM forces a pattern of data to exhibit feature selectivity, then we should observe significant channel modulation during both baseline and delay periods. Conversely, if the FEM successfully recovers feature-selective tuning functions only when the pattern of data exhibits robust feature selectivity, then we should only observe significant channel modulation in the delay period of the CTFs reported in Experiment 1. Consistent with the latter prediction, we failed to observe significant channel modulation in the baseline (Fig. 4*A*; $F_{(1.22,30.41)} = 0.37$, $p = 0.59$) data, and baseline channel activity was significantly reduced compared with delay-specific channel activity ($F_{(1,23)} = 4.03$, $p < 0.05$). Also, we note that our randomization procedure, comparing observed F values against a null F distribution derived from surrogate data, also disconfirmed this alternative explanation by showing statistically significant differences between observed orientation selectivity and surrogate orientation selectivity (Fig. 4*B*). Importantly, the shape of each surrogate F distribution (estimated from baseline, encoding, and maintenance periods) is nearly identical, highlighting the robustness of this metric. Thus, these findings cannot be explained by inherent biases in the analytic routine.

Examining the possible role of eye movements

Eye movements generate electrical signals that influence EEG recordings. Our subjects were instructed to hold fixation and horizontal EOG was used to reject trials with detectable eye movements. Nevertheless, if subtle differences in eye position covaried with the stored angle, this could provide an alternative explanation of the links we observed between EEG activity and WM storage. Thus, to determine whether there were orientation-specific eye movements, we examined the amplitude of baselined horizontal EOG across different orientation channels (Fig. 4*C*) to examine whether EOG amplitude covaries with orientation channel. Under the predictions of this alternative explanation, we

would expect maximal EOG for orientations presented at 90 degrees (horizontal), whereas we would expect minimal EOG for orientations presented at 0 degrees (vertical). In contrast to the predictions of this alternative explanation, we observed no significant modulation of EOG amplitude across orientation channels ($F_{(1.67,41.75)} = 0.99$, $p = 0.48$). To further test this alternative account, we applied a linear classifier to horizontal EOG and vertical EOG amplitudes; this analysis was done separately for the horizontal and vertical channels as well as with both channels combined. If observers were moving their eyes in a stereotyped fashion across orientations, then the linear classifier would accurately predict the correct stimulus label (orientation) more frequently than chance (.125; 1 of 8 orientations). In contrast to the predictions of this alternative explanation, we failed, across all three EOG data matrices, to reliably observe classification accuracy above chance during any part of the trial (Fig. 4*D*). Critically, during the delay period, we failed to observe above-chance classification accuracies when classifying vertical EOG ($t_{(23)} = 0.88$, $p = 0.39$), horizontal EOG ($t_{(23)} = 0.28$ $p = 0.78$), and both EOGs ($t_{(23)} = 0.59$, $p = 0.56$). Therefore, eye movements cannot explain the link between EEG activity and the stored orientation.

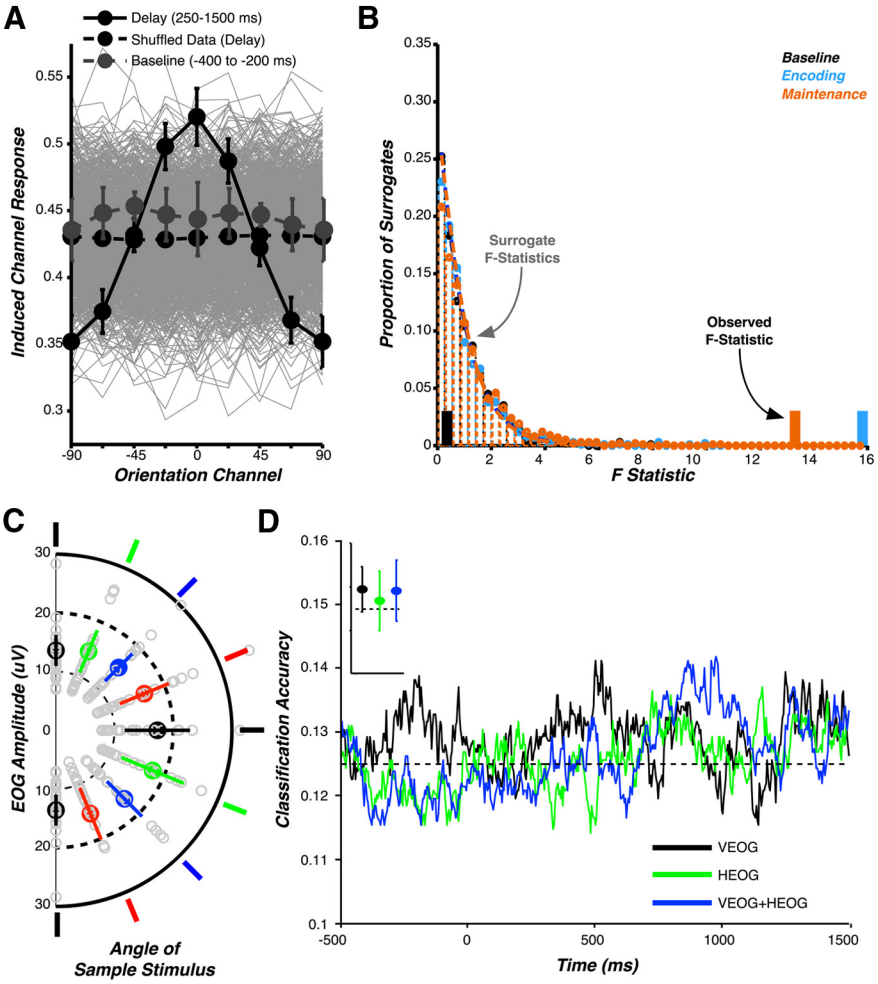

Figure 4. Ruling out alternative explanations. ***A***, To demonstrate that significant CTFs would not be observed under random conditions, we compared the observed empirical delay-specific CTF (solid black line) against a chance distribution (thin gray lines) derived from the analysis of 1000 surrogate time series that share statistical properties with the original data, which generated the expected flat average CTF (dotted black line). A similar flat CTF was observed during the baseline period (dark gray dotted line), during which time no stimulus was present. ***B***, Empirically observed *F* statistics were compared against null distributions of surrogate *F* statistics derived from the randomization routine described in Materials and Methods, Randomization analyses. For baseline (black), encoding (blue), and maintenance (orange) epochs, 1000 surrogate CTFs were generated for each subject, and *F* statistics were calculated for each surrogate series. Each distribution of surrogate *F* statistics (dotted histograms) served to determine the deviation of the empirically observed pattern (thick vertical lines) from the null distribution. These data were used to estimate the probability of observing similar *F* statistics between null and empirical data. ***C***, To demonstrate that stimulus-specific eye movement patterns did not contribute to the observed CTFs, we examined horizontal EOG amplitudes across each orientation channel (colored filled circles with error bars; thin gray circles represent each subject), and observed no significant modulation of EOG amplitude as a function of orientation channel. ***D***, A linear classifier was trained on vertical EOG (black line), horizontal EOG (green line), and both horizontal and vertical (blue line) EOG amplitudes to determine if the pattern of EOGs could allow for above-chance (black dotted line) classification accuracy. None of the EOG measures could reliably predict the angle of the remembered stimulus during the delay period. Inset figure shows the average classification accuracy for each EOG measure.

Experiment 2: examining the role of volitional storage demands on evoked and induced population tuning functions

So far, we have demonstrated that induced oscillatory activity specific to the alpha frequency band tracks the content and quality of representations in WM. Experiment 2 examines whether the sustained CTFs that track the stored feature value are contingent on the volitional storage goals of the observer. We used the same WM task as in Experiment 1, except that immediately after the offset of the sample display, a retrocue (i.e., a change in fixation color to either blue or green) instructed observers to store or drop the memorandum (color assignment was counterbalanced across subjects). If the sustained CTFs observed in Experiment 1 are contingent on the observer's intent to store the memorandum, then CTFs should be maintained in the "store" condition and abolished in the "drop" condition. In addition, we examined whether the store and drop cues would have distinct consequences for CTFs generated from evoked and induced activity. If evoked CTFs are determined by the initial volley of stimulus-driven activity, then the postsample cue should not influence the content of the evoked CTF.

We examined the pattern of evoked and induced CTFs across time for both store and drop conditions (Fig. 5). While there was no apparent difference in the evoked CTFs (from the 0 the 250 ms) between the store (Fig. 5*A*) and drop (Fig. 5*B*) conditions, there was a clear difference between delay-specific (i.e., 250–1500 ms) induced CTFs observed in store (Fig. 5*D*) and drop (Fig. 5*E*) conditions. Confirming this qualitative observation, randomization analyses of the evoked CTF during the encoding period (0–250 ms; Fig. 5*C*) revealed a significant effect of channel ($F = 9.16$, $p = 0.047$), in line with the orderly CTF evoked by the stimulus, but there was no significant effect of condition (store vs drop; $F = 0.23$, $p = 0.89$) and no channel by condition interaction ($F = 1.21$, $p = 0.65$). Randomization analyses of the induced CTF during maintenance (250–1500 ms; Fig. 5*F*) revealed a main

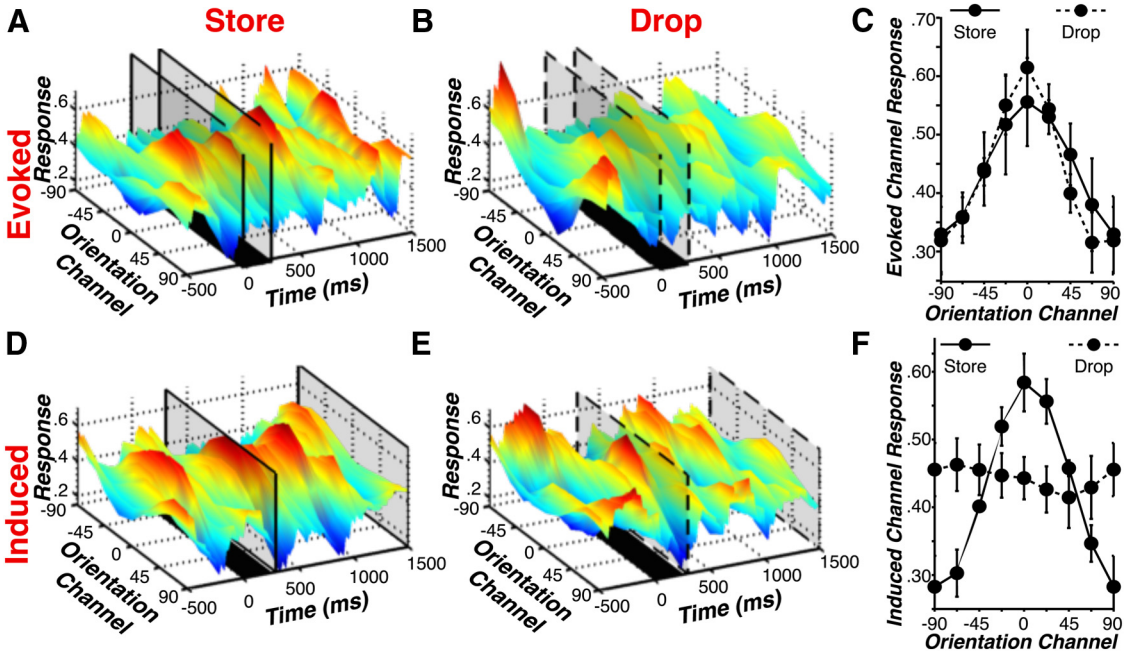

Figure 5. CTF time course across different storage demands. In Experiment 2, subjects were cued to store or drop the sample stimulus during the maintenance period. Evoked CTFs during store (***A***) and drop (***B***) trials revealed similar profiles during the initial encoding period (boundaries of gray windows). ***C***, Average encoding-specific evoked CTFs revealed no difference in CTF selectivity during store (solid black line) and drop (dotted black line) trials. In contrast, induced CTFs revealed a sustained delay-specific CTF during store (***D***), but not drop (***E***), trials (boundaries of gray windows). ***F***, Average delay-specific-induced CTFs revealed clear channel tuning in store trials, whereas a flat CTF was observed in drop trials.

effect of both channel ($F = 7.77$, $p = 0.038$) and condition ($F = 21.35$, $p = 0.014$), as well as a trending condition by channel interaction ($F = 10.30$, $p = 0.08$; Fig. 7*H*), driven by significant channel modulation in the store condition ($F = 25.45$, $p = 0.023$) and a lack of significant channel modulation in the drop condition ($F = 0.98$, $p = 0.68$).

Thus, the CTFs generated from induced alpha band activity reflect the volitionally stored contents of working memory rather than the neural activity evoked by the presentation and initial encoding of the sample.

Experiment 3: examining the role of feature-based attention on evoked and induced population tuning functions

To further examine the observers' ability to exert voluntary control over the feature values that drive EEG-based CTFs, Experiment 3 required observers to use feature-based attention to encode a target of one color into WM while ignoring a distractor of a different color. Because observers had advance knowledge of which color was relevant, this design allowed us to examine whether feature-based selection would influence both the initial encoding and subsequent maintenance of the targets and distractors.

Evoked alpha tuning functions track the deployment of feature-based attention

We observed a transient evoked feature-selective tuning function during the sample display (0–250 ms post stimulus; Fig. 6*A*,*B*). Examination of the evoked tuning functions from 0 to 250 ms (Fig. 6*E*) revealed a significant effect of channel ($F = 18.04$, $p = 0$) and a significant interaction between channel and target-distractor status ($F = 4.33$, $p = 0.024$). Compared with the distractor stimulus, attending the target stimulus led to an apparent enhancement of on-channel responses (channels at or near the encoded feature value, e.g., ±0, 22.5°) and an attenuation of off-channel responses (channels further from the encoded feature value, e.g., ±67.5, 90°), consistent with previous studies of feature-based attention (Treue and Martínez-Trujillo, 1999; Martinez-Trujillo and Treue, 2004; Serences et al., 2009b; Garcia et al., 2013). We then fitted evoked CTFs with a Gaussian function to evaluate the influence of feature-based attention on CTF amplitude and dispersion parameters, which revealed an effect of target/distractor status on amplitude (Fig. 6*F*; $t_{(24)} = 2.61$, $p < 0.01$), but not dispersion (Fig. 6*G*; $t_{(24)} = 1.59$, $p = 0.13$). Thus, the amplitude of evoked CTFs was reliably larger for the item of the attended color. In line with the findings from Experiment 1, the precision of the evoked CTF to the target did not predict between-subject differences in WM precision; a median split between high and low precision subjects revealed no difference (Fig. 7*A*; $F_{(1,23)} = 0.10$, $p = 0.76$), and there was no correlation between WM precision and CTF dispersion (Fig. 7*B*; $R^2 = 0.10$, $p = 0.13$) or amplitude (Fig. 7*C*; $R^2 = 0.02$; $p = 0.54$).

Induced alpha tuning functions track the content of selectively stored representations

The feature-based enhancement of target-evoked channel responses suggests that observers were selectively encoding the target stimulus, while actively suppressing the distractor stimulus. In line with this interpretation, we observed a sustained induced feature-selective tuning function during the delay period (250–1500 ms post stimulus) in response to the target stimulus (Fig. 6*C*), but not the distractor stimulus (Fig. 6*D*). Confirming this observation, the induced CTFs during the delay period (250–1500 ms; Fig. 6*H*) showed a significant effect of stimulus ($F = 55.85$, $p = 0.051$) and channel ($F = 9.27$, $p = 0.004$), and a significant stimulus by channel interaction ($F = 12.93$, $p = 0.001$). The significant interaction was driven by significant channel modulation in response to the target stimulus ($F = 51.20$, $p = 0$) and a lack of significant channel modulation in response to the distractor stimulus ($F = 0.53$, $p = 0.508$). Thus, Experiment 3 provides further evidence that EEG-based CTFs selectively represent the feature values that are relevant to the observers' current storage goals.

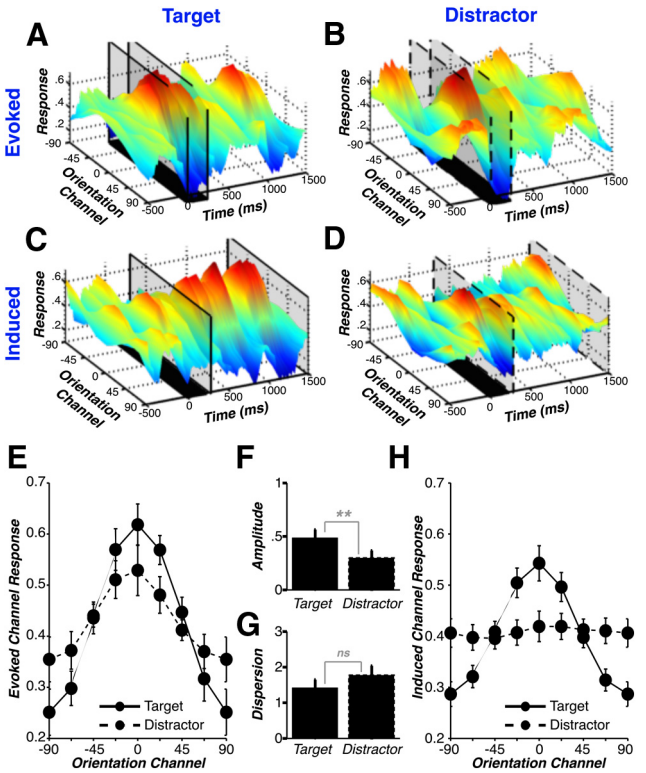

Figure 6. CTF time course during deployment of feature-based attention. In Experiment 3, subjects were presented with two differently colored sample stimuli during encoding, and cued to remember one (target) and ignore the other (distractor). Evoked CTFs revealed similar encoding-specific CTF profiles for target (***A***) and distractor (***B***) sample stimuli. In contrast, induced CTFs revealed a sustained delay-specific CTF profile for target stimuli (***C***), whereas we observed no delay-specific CTF for distractor stimuli (***D***). ***E***, Average evoked CTFs during encoding (0–250 ms) revealed a significant loss of selectivity in distractor CTFs relative to target CTFs. The loss of selectivity observed in distractor-evoked CTFs was driven by a decline in CTF amplitude (***F***), but not CTF dispersion (***G***). ***H***, Average induced CTFs during maintenance (250–1500 ms) revealed a significant target-induced CTF, whereas distractor-induced CTFs were flat.

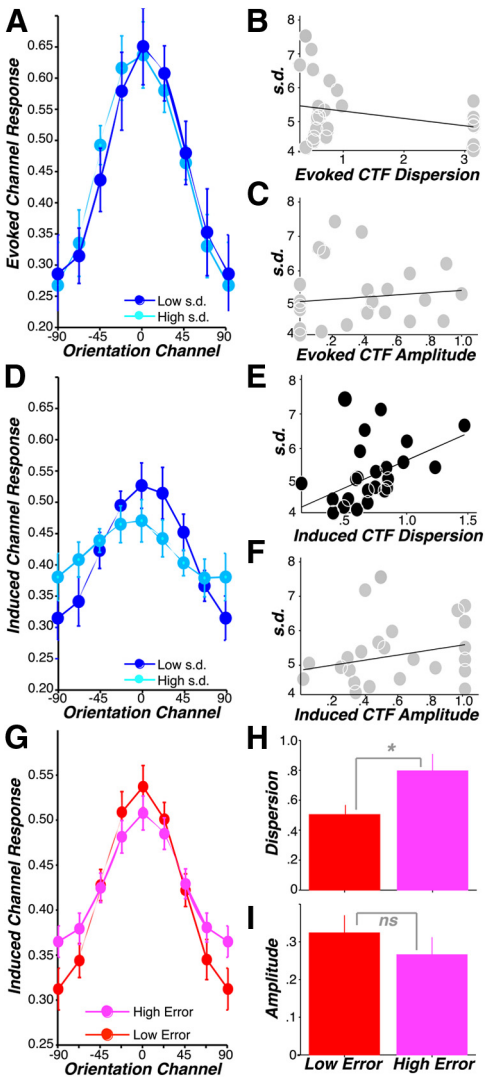

Figure 7. Replicating between- and within- subject links with mnemonic precision. ***A***, ***D***, A median-split analysis based on between-subject differences in mnemonic precision (s.d.) was performed on evoked and induced CTFs. No group difference was observed in evoked CTFs (***A***), and no apparent link was observed between mnemonic precision and CTF dispersion (***B***; $R^2 = 0.10$, $p = 0.13$) or amplitude (***C***; $R^2 = 0.02$; $p = 0.54$). ***D***, A clear loss of selectivity was observed for low precision (dark blue) relative to high precision (light blue) subjects in induced CTFs ($p < 0.05$). A strong link was observed between mnemonic precision and induced CTF dispersion (***E***; $R^2 = 0.30$, $p < 0.01$), but not amplitude (***F***; $R^2 = 0.09$, $p = 0.16$). ***G***, A median-split analysis based on within-subject differences in response error. CTFs were generated from trials grouped into high error (low precision; pink line) or low error (high precision; red line) bins. A clear loss of selectivity was observed in high error trials. The observed loss of selectivity was attributable to a significant difference in CTF dispersion (***H***; $p < 0.05$), but not amplitude (***I***; $p = 0.25$).

Replication of the link between CTF dispersion and memory quality

Replicating the findings from Experiment 1, we observed a significant relationship between mnemonic precision (s.d.) and channel responses observed in delay-specific induced CTFs (Fig. 7*D*; $F_{(1.77,44.92)} = 4.25$, $p < 0.05$), such that there was an apparent reduction in CTF selectivity in low precision subjects compared with high precision subjects. To evaluate how selectivity interacts with mnemonic precision, we fitted channel response functions with a Gaussian function to estimate CTF dispersion and amplitude. As in Experiment 1, we observed a significant link between individual differences in mnemonic precision and dispersion (Fig. 7*E*; $R^2 = 0.30$, $p < 0.01$), but not amplitude (Fig. 7*F*; $R^2 = 0.09$, $p = 0.16$). In addition, Experiment 3 replicated the interaction between CTF selectivity and within-subject variations in response error (Fig. 7*G*); CTF selectivity was reduced for low relative to high error trials, leading to a significant interaction between error bin (high vs low) and orientation channel ($F_{(1.87,44.83)} = 3.90$, $p < 0.05$). Next, we fitted low and high error CTFs with a Gaussian function to examine whether the observed error-related difference in selectivity was due to changes in response amplitude or dispersion. Consistent with the pattern observed in the between-subject analysis, we observed larger dispersion estimates in low– relative to high–error trials (Fig. 7*G*; $t_{(24)} = 2.46$, $p < 0.05$), whereas we observed no difference in amplitude (Fig. 7*G*; $t_{(23)} = 1.17$, $p = 0.25$). Thus, we replicated the basic empirical link between mnemonic precision and the selectivity of EEG-based CTFs.

Discussion

Mounting evidence has implicated oscillatory activity in the alpha frequency band in the on-line storage of information in WM, with multiple studies linking activity in this frequency band with the number of representations stored in WM (Jensen et al., 2002; Sauseng et al., 2009; Palva et al., 2010). The present work provides a critical extension of these findings by demonstrating that the spatial distribution of activity in the alpha frequency band tracks

both the content and quality of representations stored in WM. Using a forward modeling approach to reconstruct orientation-specific population response profiles, Experiment 1 revealed two temporally and functionally distinct generators of stimulus-specific CTFs: (1) evoked (i.e., stimulus-locked) alpha power enabled the construction of orderly CTFs during the initial encoding period (0–250 ms after sample onset) and (2) induced (not stimulus-locked) alpha power revealed sustained stimulus-specific CTFs from shortly after stimulus onset to the end of the delay period. Critically, whereas the dispersion of delay-specific (250–1500 ms) induced tuning functions strongly predicted both between- and within-subject variability in mnemonic precision, we observed no link between mnemonic precision and the dispersion of evoked tuning functions.

Experiments 2 and 3 showed that these EEG-based CTFs are specific to the intentionally stored aspects of the stimulus display. Experiment 2 showed that when a postsample cue instructed observers to drop the sample stimulus from memory, the CTF was abolished within about 300 ms; thus, the sustained CTFs observed during the delay period are not a passive consequence of the initial encoding of the sample. Similarly, Experiment 3 showed that when observers were instructed to store a target of a relevant color while ignoring a distractor of a different color, both evoked and induced CTFs were modulated by feature-based attention. The target-evoked CTF was more selective than the one for the distractor, reminiscent of past demonstrations of feature-similarity gain (Treue and Martínez-Trujillo, 1999; Martinez-Trujillo and Treue, 2004; Serences et al., 2009a; Scolari et al., 2012; Garcia et al., 2013). In addition, only the target stimulus elicited a sustained induced CTF during the delay period, while the distractor stimulus representation faded quickly after sample offset. Finally, Experiment 3 replicated the prior finding that the dispersion of the induced CTF was a robust predictor of both between- and within-subject variations in WM precision. Thus, EEG-based CTFs appear to provide an expedient method for measuring the integrity of the sensory population codes that are thought to support on-line memory representations (Awh and Jonides, 2001; Jonides et al., 2005; Postle, 2006; D'Esposito, 2007).

Our findings suggest that evoked and induced activity track qualitatively different aspects of stimulus-specific activity during the encoding and storage of information in visual working memory. While evoked activity yielded CTFs that were early and brief, induced activity yielded CTFs that were sustained throughout the delay period and predictive of memory quality. Also, recall that early and late trial epochs had distinct profiles in our analysis of MI. MI analyses of induced oscillatory alpha power during stimulus encoding revealed higher MI in posterior relative to anterior electrodes, while the same analysis during the subsequent delay period revealed uniform MI levels across anterior and posterior electrodes. One possibility is that these dynamic changes in feature-selective activity across time reflect the translation of the memoranda from early stimulus-driven format to an endogenously sustained representation that extends across a broader range of feature-selective neural populations (Engel et al., 2001; Varela et al., 2001; Jensen and Colgin, 2007; Siegel et al., 2012). Indeed, recent neuroimaging studies have revealed spatially global feature selectivity in visual cortex (Ester et al., 2009), parietal cortex (Christophel et al., 2012), and frontal eye fields (Serences and Yantis, 2007; Heinen et al., 2013). The present work raises the possibility that feature-selective assemblies may be coordinated across relatively large cortical distances by synchronized activity in the alpha band (Varela et al., 2001; Palva and Palva, 2007; Palva et al., 2010).

A broad array of evidence converges on the conclusion that representations in WM are maintained via sustained activity in the same brain regions enabling the sensory encoding of the memoranda, the so-called sensory recruitment hypothesis (Awh and Jonides, 2001; Jonides et al., 2005; Postle, 2006; Ester et al., 2009; Harrison and Tong, 2009; Serences et al., 2009a; Anderson et al., 2013; Emrich et al., 2013). One hypothesis suggests synchronization in a specific frequency band supports the integration of the sensory population codes that represent information in WM (Singer and Gray, 1995; Singer, 1999) while amplifying their downstream influence by facilitating the temporal summation of postsynaptic potentials (Salinas and Sejnowski, 2001; Fries, 2005). If alpha is the binding frequency in visual WM, this broad framework may explain why variations in the spatial distribution of alpha power tracked the stored orientation value. Because different units participate in the population code for different orientations, different units–with reliably different spatial distributions–may have been synchronized in the alpha band for different orientations. Thus, the clear links between alpha band activity and the content and quality of visual WM fall in line with a biologically plausible model of storage in working memory, with the following properties: (1) the number of on-line representations is determined by the number of simultaneously activated alpha oscillators. (2) the precision of the representations is determined by the quality of population codes integrated into each alpha oscillator, and (3) offsets in the phase of each alpha oscillator enable the segregation of each item's population code from that of other items (cf., Lisman and Idiart, 1995; Raffone and Wolters, 2001). This model makes specific predictions about the consequences of increasing the number of items stored in working memory: (1) the number of phase-offset alpha generators will increase, leading to a reduction in alpha power and (2) the selectivity of orientation-selective population codes–examined via activity in the alpha frequency band–should decline. Further work with the approach outlined here will provide a clear test of these predictions, and may provide an approach for evaluating neural mechanistic models of the inverse relationship between precision and storage load in working memory (Wilken and Ma, 2004; Zhang and Luck, 2008; Anderson et al., 2011, 2013b; Anderson and Awh, 2012).

Although we have emphasized the role of alpha band activity in the coordination of feature-specific population codes, there is active debate regarding the functional role of oscillatory alpha activity (Jensen and Mazaheri, 2010). The so-called inhibition-timing hypothesis proposes alpha as a gating mechanism for the transmission of information (Jensen et al., 2002; Klimesch et al., 2007; Sauseng et al., 2009). According to this hypothesis, alpha synchronization reflects functional inhibition of task-irrelevant cortical populations, whereas alpha desynchronization reflects release from inhibition and the emergence of functionally coordinated neural activity among differentiated neural assemblies. The present data does not offer a direct test of whether fluctuations in alpha power reflect changes in functional inhibition, but we note that the present data do not necessarily contradict the putative link between alpha power and inhibition. First, it is plausible that multiple neural processes cause changes in alpha power, and that our paradigm holds most of those processes constant while manipulating the stored orientation. In this case, further work that combines storage demands with inhibitory demands may help to elucidate whether fluctuations in alpha power may be linked with more than one aspect of cognitive processing. Sec-

ond, if increased alpha power is associated with increased inhibition, this inhibitory activity could theoretically help to determine which neural units participate in the sensory population codes that are thought to underlie storage in working memory.

Although we have highlighted past work (Sauseng et al., 2009; Palva et al., 2010) that has linked variations in alpha power with variations in the number of items stored in working memory, another prominent neural signature of storage in working memory has been documented in studies of contralateral delay activity (CDA; Vogel and Machizawa, 2004; McCollough et al., 2007). Similar to load-dependent alpha activity (Sauseng et al., 2009; Palva et al., 2010), CDA amplitude increases monotonically with set size, reaching an upper bound at set sizes predicted by behavioral measures of working memory capacity (Vogel and Machizawa, 2004; Anderson et al., 2011). Jensen et al. (2010) have proposed a possible link between alpha activity and CDA (Mazaheri and Jenson, 2008; van Dijk et al., 2010). Supported by observed asymmetric amplitude modulations of alpha activity (Stam et al., 1999), Jensen et al. (2010) have proposed that differences in CDA amplitude across contralateral and ipsilateral hemispheres result from asymmetric changes in the peak–but not the trough–of alpha oscillations across hemispheres. For example, event-related averaging of multiple trials containing a reduction in the peak of contralateral desynchronized alpha activity and an increase in the peak of ipsilateral synchronized alpha activity (Sauseng et al., 2009) would lead to a larger sustained negativity in contralateral, relative to ipsilateral, electrodes. Thus, the link between CDA activity and the number of items stored in working memory may reflect asymmetric modulations of alpha band activity, while the spatial distribution of those modulations may track the specific neural populations that encode the stored feature values.

References

Anderson DE, Awh E (2012) The plateau in mnemonic resolution across large set sizes indicates discrete resource limits in visual working memory. Atten Percept Psychophys 74:891–910. CrossRef Medline

Anderson DE, Vogel EK, Awh E (2011) Precision in visual working memory reaches a stable plateau when individual item limits are exceeded. J Neurosci 31:1128–1138. CrossRef Medline

Anderson DE, Ester EF, Serences JT, Awh E (2013a) Attending multiple items decreases the selectivity of population responses in human primary visual cortex. J Neurosci 33:9273–9282. CrossRef Medline

Anderson DE, Vogel EK, Awh E (2013b) The selection and storage of perceptual groups is constrained a discrete resource in working memory. J Exp Psychol Hum Percept Perform 39:824–835. CrossRef Medline

Awh E, Jonides J (2001) Overlapping mechanisms of attention and spatial working memory. Trends Cogn Sci 5:119–126. CrossRef Medline

Brainard DH (1997) The psychophysics toolbox. Spat Vis 10:433–436. CrossRef Medline

Brouwer GJ, Heeger DJ (2009) Decoding and reconstructing color from responses in human visual cortex. J Neurosci 29:13992–14003. CrossRef Medline

Brouwer GJ, Heeger DJ (2011) Cross-orientation suppression in human visual cortex. J Neurophysiol 106:2108–2119. CrossRef Medline

Christophel TB, Hebart MN, Haynes JD (2012) Decoding the contents of visual short term memory from human visual and parietal cortex. J Neurosci 32:12983–12989. CrossRef Medline

Delorme A, Makeig S (2004) EEGLAB: an open source toolbox for analysis of single-trial EEG dynamics including independent component analysis. J Neurosci Methods 134:9–21. CrossRef Medline

D'Esposito M (2007) From cognitive to neural models of working memory. Philos Trans R Soc Lond B Biol Sci 362:761–772. CrossRef Medline

Emrich SM, Riggall AC, Larocque JL, Postle BR (2013) Distributed patterns of activity in sensory cortex reflect the precision of multiple items maintained in visual short-term memory. J Neurosci 33:6516–6523. CrossRef Medline

Engel AK, Fries P, Singer W (2001) Dynamic predictions: oscillations and synchrony in top-down processing. Nat Rev Neurosci 2:704–716. CrossRef Medline

Ester EF, Serences JT, Awh E (2009) Spatially global representations in human primary visual cortex during working memory maintenance. J Neurosci 29:15258–15265. CrossRef Medline

Ester EF, Anderson DE, Serences JT, Awh E (2013) A neural measure of precision in visual working memory. J Cogn Neurosci 25:754–761. CrossRef Medline

Fries P (2005) A mechanism for cognitive dynamics: neuronal communication through neuronal coherence. Trends Cogn Sci 9:474–480. CrossRef Medline

Garcia JO, Srinivasan R, Serences JT (2013) Near-real-time feature-selective modulations in human cortex. Curr Biol 23:515–522. CrossRef Medline

Harrison SA, Tong F (2009) Decoding reveals the contents of visual working memory in early visual areas. Nature 458:632–635. CrossRef Medline

Heinen K, Feredoes E, Weiskopf N, Ruff CC, Driver J (2013) Direct evidence for attention-dependent influences on the frontal eye-fields on feature-responsive visual cortex. Cereb Cortex. Advance online publication. Retrieved March 5, 2014. doi:10.1093/cercor/bht157. CrossRef

Jensen O, Colgin LL (2007) Cross-frequency coupling between neuronal oscillations. Trends Cogn Sci 11:267–269. CrossRef Medline

Jensen O, Mazaheri A (2010) Shaping functional architecture by oscillatory alpha activity: gating by inhibition. Front Hum Neurosci 4:186. CrossRef Medline

Jensen O, Gelfand J, Kounios J, Lisman JE (2002) Oscillations in the alpha band (9–12 Hz) increase with memory load during retention in a short-term memory task. Cereb Cortex 12:877–882. CrossRef Medline

Jonides J, Lacey SC, Nee DE (2005) Processes of working memory in mind and brain. Curr Dir Psychol Sci 14:2–5. CrossRef

Klimesch W, Sauseng P, Hanslmayr S (2007) EEG alpha oscillations: the inhibition-timing hypothesis. Brain Res Rev 53:63–88. CrossRef Medline

Lisman JE, Idiart MA (1995) Storage of 7 ± 2 short-term memories in oscillatory subcycles. Science 267:1512–1515. CrossRef Medline

Martinez-Trujillo JC, Treue S (2004) Feature-based attention increases the selectivity of population responses in primate visual cortex. Curr Biol 14:744–751. CrossRef Medline

Mazaheri A, Jensen O (2008) Asymmetric amplitude modulations of brain oscillations generate slow evoked responses. J Neurosci 28:7781–7787. CrossRef Medline

McCollough AW, Machizawa MG, Vogel EK (2007) Electrophysiological measures of maintaining representations in visual working memory. Cortex 43:77–94. CrossRef Medline

Palva JM, Monto S, Kulashekhar S, Palva S (2010) Neuronal synchrony reveals working memory networks and predicts individual memory capacity. Proc Natl Acad Sci U S A 107:7580–7585. CrossRef Medline

Palva S, Palva JM (2007) New vistas for alpha-frequency band oscillations. Trends Neurosci 30:150–158. CrossRef Medline

Pelli DG (1997) The VideoToolbox software for visual psychophysics: transforming numbers into movies. Spat Vis 10:437–442. CrossRef Medline

Postle BR (2006) Working memory as an emergent property of the mind and brain. Neuroscience 139:23–38. CrossRef Medline

Raffone A, Wolters G (2001) A cortical mechanism for binding in visual working memory. J Cogn Neurosci 13:766–785. CrossRef Medline

Salinas E, Sejnowski TJ (2001) Correlated neural activity and the flow of neural information. Nat Rev Neurosci 2:539–550. CrossRef Medline

Sauseng P, Klimesch W, Heise KF, Gruber WR, Holz E, Karim AA, Glennon M, Gerloff C, Birbaumer N, Hummel FC (2009) Brain oscillatory substrates of visual short term memory capacity. Curr Biol 19:1846–1852. CrossRef Medline

Scolari M, Byers A, Serences JT (2012) Optimal deployment of attention gain during fine discriminations. J Neurosci 32:7723–7733. CrossRef Medline

Serences JT, Saproo S (2012) Computational advances towards linking BOLD and behavior. Neuropsychologia 50:435–446. CrossRef Medline

Serences JT, Yantis S (2007) Spatially-selective representations of voluntary and stimulus-driven attentional priority in human occipital, parietal, and frontal cortex. Cereb Cortex 17:284–293. Medline

Serences JT, Ester EF, Vogel EK, Awh E (2009a) Stimulus-specific delay activity in human primary visual cortex. Psychol Sci 20:207–214. CrossRef Medline

Serences JT, Saproo S, Scolari M, Ho T, Muftuler LT (2009b) Estimating the

influence of attention on population codes in human visual cortex using voxel based tuning functions. Neuroimage 44:223–231. CrossRef Medline
Shannon CE (1948) A mathematical theory of communication. Bell System Technical Journal 27:379–423.
Siegel M, Donner TH, Engel AK (2012) Spectral fingerprints of large-scale neuronal interactions. Nat Rev Neurosci 13:121–134. CrossRef Medline
Singer W (1999) Timing as coding space? Curr Opin Neurobiol 9:189–194. CrossRef Medline
Singer W, Gray CM (1995) Visual feature integration and the temporal correlation hypothesis. Annu Rev Neurosci 18:555–586. CrossRef Medline
Stam CJ, Pijn JP, Suffczynski P, Lopes da Silva FH (1999) Dynamics of the human alpha rhythm: evidence for non-linearity? Clin Neurophysiol 110: 1801–1813. CrossRef Medline
Todd JJ, Marois R (2004) Capacity limit of visual short-term memory in human posterior parietal cortex. Nature 428:751–754. CrossRef Medline
Treue S, Martínez-Trujillo JC (1999) Feature-based attention influences motion processing gain in macaque visual cortex. Nature 399:575–579. CrossRef Medline
van Dijk H, van der Werf J, Mazaheri A, Medendorp WP, Jensen O (2010) Modulations in oscillatory activity with amplitude asymmetry can produce cognitively relevant event-related responses. Proc Natl Acad Sci U S A 107:900–905. CrossRef Medline
Varela F, Lachaux JP, Rodriguez E, Martinerie J (2001) The brainweb: phase synchronization and large-scale integration. Nat Rev Neurosci 2:229–239. CrossRef Medline
Vogel EK, Machizawa MG (2004) Neural activity predicts individual differences in visual working memory capacity. Nature 428:748–751. CrossRef Medline
Vogel EK, Luck SJ, Shapiro KL (1998) Electrophysiological evidence for a postperceptual locus of suppression during the attentional blink. J Exp Psychol Hum Percept Perform 24:1656–1674. CrossRef Medline
Wilken P, Ma WJ (2004) A detection theory account of change detection. J Vis 4(12):1120–1135. CrossRef Medline
Xu Y, Chun MM (2006) Dissociable neural mechanisms supporting visual short-term memory for objects. Nature 440:91–95. CrossRef Medline
Zhang W, Luck SJ (2008) Discrete fixed-resolution representations in visual working memory. Nature 453:233–235. CrossRef Medline