

Training Working Memory in Childhood Enhances Coupling between Frontoparietal Control Network and Task-Related Regions

Jessica J. Barnes,¹ Anna Christina Nobre,² Mark W. Woolrich,² Kate Baker,³ and Duncan E. Astle¹

¹Medical Research Council, Cognition and Brain Sciences Unit, Cambridge CB2 7EF, United Kingdom, ²Oxford Centre for Human Brain Activity, University of Oxford, Oxford OX3 7JX, United Kingdom, and ³Department of Medical Genetics, University of Cambridge, Cambridge CB2 0QQ, United Kingdom

Working memory is a capacity upon which many everyday tasks depend and which constrains a child's educational progress. We show that a child's working memory can be significantly enhanced by intensive computer-based training, relative to a placebo control intervention, in terms of both standardized assessments of working memory and performance on a working memory task performed in a magnetoencephalography scanner. Neurophysiologically, we identified significantly increased cross-frequency phase amplitude coupling in children who completed training. Following training, the coupling between the upper alpha rhythm (at 16 Hz), recorded in superior frontal and parietal cortex, became significantly coupled with high gamma activity (at ~90 Hz) in inferior temporal cortex. This altered neural network activity associated with cognitive skill enhancement is consistent with a framework in which slower cortical rhythms enable the dynamic regulation of higher-frequency oscillatory activity related to task-related cognitive processes.

Key words: cognitive training; development; developmental disorders; MEG; oscillations; working memory

Significance Statement

Whether we can enhance cognitive abilities through intensive training is one of the most controversial topics of cognitive psychology in recent years. This is particularly controversial in childhood, where aspects of cognition, such as working memory, are closely related to school success and are implicated in numerous developmental disorders. We provide the first neurophysiological account of how working memory training may enhance ability in childhood, using a brain recording technique called magnetoencephalography. We borrowed an analysis approach previously used with intracranial recordings in adults, or more typically in other animal models, called "phase amplitude coupling."

Introduction

Many everyday activities depend upon our ability to hold in mind and manipulate small amounts of information for brief periods.

This ability, "working memory," is severely tested as children progress through years of intensive education (Gathercole et al., 2004), and poor working memory skills are highly predictive of educational underachievement and neurodevelopmental disorder (Alloway, 2009). Recent studies demonstrate that, in childhood (8–11 years of age), working memory can be trained, with benefits transferring to similarly structured but untrained working memory tests (e.g., Holmes et al., 2009).

In contrast to the wealth of behavioral evidence, studies examining the impact of training on neurophysiology are rare, especially so in childhood. Little is known about the brain mechanisms by which training might enhance working memory. The purpose of this study is to redress this. We focus on children in middle childhood because working memory differences here are closely associated with academic performance (Gathercole et al., 2004), and there is strong evidence of training gains at this age (Holmes et al., 2009).

Previously, we explored spontaneous fluctuations in resting-state MEG after working memory training (Astle et al., 2015). Following training, there was increased coordination of oscillations

Received Jan. 11, 2016; revised July 1, 2016; accepted July 8, 2016.

Author contributions: D.E.A. designed research; J.J.B. and D.E.A. performed research; A.C.N., M.W.W., and D.E.A. analyzed data; J.J.B., A.C.N., M.W.W., K.B., and D.E.A. wrote the paper.

D.E.A. was supported by British Academy postdoctoral fellowship and United Kingdom Medical Research Council intramural funding MC-A060-5PQ41. A.C.N. was supported by Wellcome Trust Senior Investigator Award 104571/Z/14/Z. M.W.W. was supported by the Wellcome Trust, the National Institute for Health Research, and the Oxford Biomedical Research Centre based at Oxford University Hospitals Trust, Oxford University. A.C.N. and M.W.W. were supported by Medical Research Council United Kingdom MEG Partnership Grant MR/K005464/1. K.B. was supported by National Institute for Health Research Academic Clinical Lectureship.

The authors declare no competing financial interests.

This article is freely available online through the *J. Neurosci.* Author Open Choice option.

Correspondence should be addressed to Dr. Duncan Astle, Medical Research Council Cognition and Brain Sciences Unit, 15 Chaucer Road, Cambridge, UK, CB2 7EF. E-mail: duncan.astle@mrc-cbu.cam.ac.uk.

DOI:10.1523/JNEUROSCI.0101-16.2016

Copyright © 2016 Barnes et al.

This is an Open Access article distributed under the terms of the Creative Commons Attribution License Creative Commons Attribution 4.0 International, which permits unrestricted use, distribution and reproduction in any medium provided that the original work is properly attributed.

tory brain activity between a bilateral dorsal frontoparietal network (incorporating frontal eye fields and superior parietal cortex) and an area in inferior temporal cortex. This frontoparietal network has previously been implicated in the control of attention (Gregoriou et al., 2009) and visual working memory (Kuo et al., 2011). One interpretation of this result is that repeatedly activating the control network over 20 training sessions alters its coupling with task-related areas, with these changes being robust enough to be detected at rest. A further implication is that training-induced changes are of functional benefit during the performance of relevant cognitive tasks. However, this resting-state analysis establishes candidate areas but does not address these key questions about the functional purpose of this altered connectivity. These questions are addressed in the current study.

One possible consequence of enhanced connectivity is more effective top-down regulation of the sensory activity to be encoded and maintained in working memory. Dynamic neural activity in areas of the frontal and parietal lobe is thought to be involved in coding task-critical information (Desimone and Duncan, 1995; Duncan and Owen, 2000). Long-range connections enable these brain areas to organize ongoing sensory processing in a goal-directed fashion. Previous research in human adults (e.g., Voytek et al., 2010), nonhuman primates (e.g., Haegens et al., 2011), and rodents (e.g., Axmacher et al., 2010) provides a neurophysiological framework by which this functional integration may occur. Slow oscillations, putatively expressed in long-range functional networks, modulate the excitability of local neural assemblies (Jensen and Colgin, 2007; Jensen et al., 2012), and the high-frequency activity reflected by these local processes can subsequently become coupled to the slower rhythm (Canolty et al., 2006). This mechanism, cross-frequency phase amplitude coupling, provides a basis by which attentional demands, including spatial location (Sauseng et al., 2008), timing (Lakatos et al., 2005; Cravo et al., 2011), and relevance (Cohen et al., 2009; Siegel et al., 2009), can influence ongoing sensory, cognitive, or motor processes.

The purpose of the current study was to understand better the neurophysiological changes that accompany improvements following working memory training, using MEG. We wanted to test this account of neuronal coupling previously proposed in the attentional control literature. The data were acquired as part of a double-blind randomized placebo-controlled study of working memory training. In addition to the acquisition of resting-state data, participants performed a working memory task in the MEG scanner before and after training. To assess training-associated changes in task-related activity, we used the network previously shown to change following training in these same children at rest (as reported by Astle et al., 2015) as an initial starting point for the coupling analyses. However, in view of the limited analysis of neurophysiological coupling in childhood to date, we subsequently applied several anatomical approaches (region of interest, network-based, and whole-brain) and alternative statistical methods to our data to ensure robust, convergent findings.

Materials and Methods

Participants. The study protocol was approved by the University of Cambridge Psychology Research Ethics Committee. Thirty-three children, 8–11 years of age (mean age = 119.2 months; SD = 11.3 months; 12 males) were recruited via local schools (see also Astle et al., 2015). Parents provided written informed consent. Children were randomly allocated to either the Adaptive or Placebo group. Parents were made aware of the

existence of an intergroup experimental manipulation but were not informed as to the nature of it. Children and parents were blind to group allocation. These numbers were recruited on the basis of a previous behavioral study (Holmes et al., 2009), which used the same training program and set of assessments. With the effect sizes reported in this previous study, significant training effects on verbal working memory should be apparent with just 7 children, and spatial working memory effects with 12 children. Six children failed to complete the training and did not return for the second session, so subsequent analysis was performed on the datasets from the remaining 27 children, 14 of whom had been allocated to the Placebo group and 13 to the Adaptive group. The Adaptive group included 3 boys, and the children in this group had a mean age of 119.5 months (SD = 9.8 months). The Placebo group included 7 boys, and the children in this group had a mean age of 118.8 months (SD = 10.8 months). A *post hoc* power calculation, using the effect sizes from Holmes et al. (2009) revealed that the study was well powered for detecting significant training effects in both verbal (0.99) and spatial working memory (0.95). There was no significant difference in age between the two groups ($t_{(25)} = 0.182$, $p = 0.857$), the gender balance did not differ significantly between the groups ($\chi^2 = 2.095$, $p = 0.148$), and the two groups had equivalent IQ scores (as assessed using vocabulary and matrix reasoning scales from Wechsler Abbreviated Scale of Intelligence; $t_{(25)} = 0.226$, $p = 0.822$).

Cognitive training. Both groups performed the training at home, and each child's progress was remotely monitored online by the research team. Children performed a minimum of 20 sessions, with an upper limit of 25 sessions, over 4–6 weeks. Each training session comprised 8 tasks (from a bank of 12 possible tasks). These tasks all have in common the requirement to maintain pieces of information in serial order and to provide a full report of that sequence. In some tasks, this was spatial information (e.g., locations flashing up on a grid) and in others verbal information (e.g., digits appearing in visual or auditory format). In some cases, the retention of serial order was in isolation, and in other cases in combination with some concurrent processing (e.g., remembering locations lighting up on a disc that keeps spinning) (for full details of these tasks, see www.cogmed.com/RM). In each session, children performed 15 trials of each task (a total duration of 30–45 min per session). The children in the Adaptive group received a commercially available version of this program, with task difficulty being adjusted on a trial-by-trial basis, such that children were always being worked at the limits of their current capacity. For all games, difficulty was increased in the Adaptive condition by increasing the span length. The children assigned to the Placebo group completed an identical training program, save for the fact that the tasks had a span level limited to 2 items across all trials (see also Holmes et al., 2009; Dunning et al., 2013; Astle et al., 2015).

For the Adaptive group, the training program provides a “training index.” This is the difference between the maximum span level reached, minus that achieved on session 3 of the training. In this study, the mean training index was 21, which is in line with other studies that have used this training program (e.g., Holmes et al., 2009). We also compared the training index across children according to the number of training sessions completed. There is little variability in the number of sessions completed: 7 children completed 20 sessions, 2 completed 21 sessions, 1 completed 22 sessions, 2 completed 23 sessions, and 1 completed 24 sessions. Using a linear regression analysis, we tested whether the number of sessions modulated the training index. There was no significant impact of session number on the training index (standardized beta = -0.254 , $t = -0.870$, $p = 0.403$).

Standardized cognitive assessments. Standardized assessments of short-term and working memory ability were taken before and after training using the Automated Working Memory Assessment (Alloway et al., 2008), as a way of testing whether the training activities had impacted upon untrained working memory abilities. Like the games in the training program, these assessments all require the retention of item information in serial order, and the full report of that information at the end of each trial. We used verbal and spatial measures of short-term and working memory from this battery (Forwards Digit Span, Backwards Digit Span, Dot Matrix, and Spatial Span subtests). For one child, we only had scores in the spatial domain. Assessments at both time points were conducted

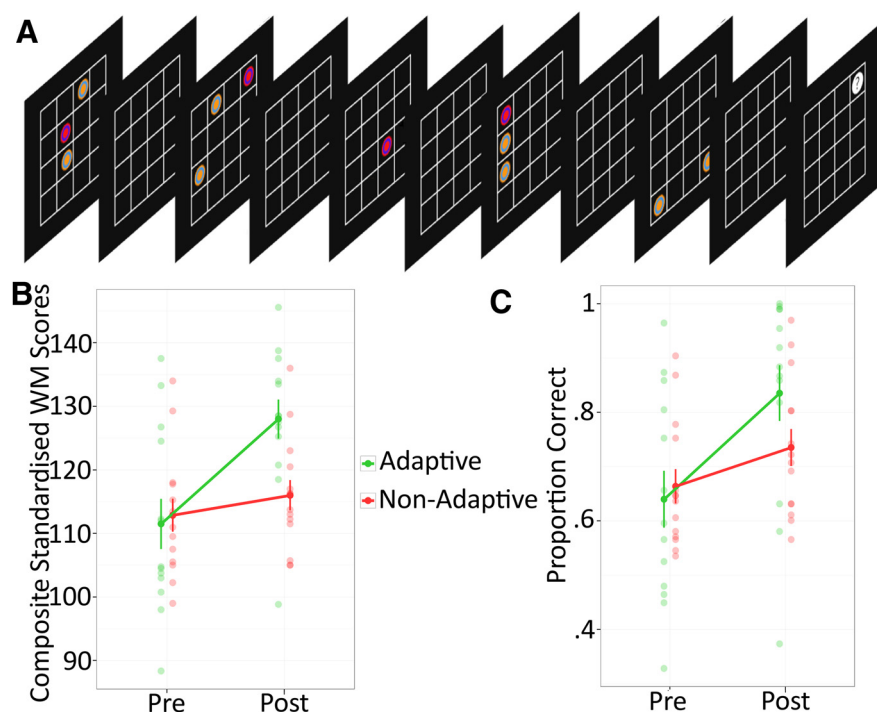


Figure 1. Trial schematic and behavioral data. **A**, Example of the in-scanner task. Children are required to remember the location and order of the four red target items. Any nonred distracter items are to be ignored. At the end of the trial, a single location in the matrix is probed, and children are required to report the sequential position of the target presented at that location. **B**, The composite age-standardized working memory scores, for both groups of children, before and after training. These scores are collapsed across the four subtests used from the Automated Working Memory Assessment. Error bars indicate SEM. **C**, Mean proportion correct performance on our in-scanner task, for both groups, before and after training. Error bars indicate SEM.

by a member of the research team who was blind to group allocation. To characterize our sample, we converted performance into age-independent standardized scores. There were no pretraining differences between the groups. Verbal short-term memory: Adaptive = 116 (SD = 19) versus Placebo = 118 (SD = 16) ($t_{(24)} = 0.397$, $p = 0.695$). Verbal working memory: Adaptive = 106 (SD = 21) versus Placebo = 111 (SD = 13) ($t_{(24)} = 0.645$, $p = 0.525$). Spatial short-term memory: Adaptive = 109 (SD = 15) versus Placebo = 107 (SD = 14) ($t_{(25)} = 0.335$, $p = 0.741$). Spatial working memory: Adaptive = 116 (SD = 14) versus Placebo = 115 (SD = 11) ($t_{(25)} = 0.139$, $p = 0.891$).

We separately analyzed the data from the 6 children who did not complete the training. These children tended to have lower scores on the pretraining assessments, relative to the children who completed the training: verbal short-term memory (105, $t_{(29)} = 1.659$, $p = 0.108$), spatial short-term memory (86, $t_{(30)} = 3.378$, $p = 0.002$), verbal working memory (93, $t_{(29)} = 2.123$, $p = 0.042$), and spatial working memory (99, $t_{(30)} = 2.880$, $p = 0.007$).

In-scanner task design. Working memory task-related MEG data were collected before and after the training intervention. We used a probed recall task in the scanner; like the standardized assessments, this was not included in the training regimen. The task required children to remember the spatial locations of targets presented in serial order. An example trial sequence can be seen in Figure 1A. The task was administered in two runs of 20 min each, separated by a break. A probed recall task was used rather than full report because of the constraints imposed by the scanner; it was not possible for children to use either a touch screen or computer mouse in the scanner. The probed recall design enabled us retain the requirement for children to remember the full sequence of locations in serial order, but without the need for full report. There were no performance differences, in terms of accuracy, on this task between the groups before training ($t_{(25)} = 0.288$, $p = 0.776$).

On each trial of this task, children were shown a sequence of 5 grids (each comprised of a 4×4 array of cells), with grids presented for 300 ms each and spaced 700 ms apart. Four of the grids contained a target dot,

and task required subjects to remember the position of these targets in the order that they appeared. In addition, each grid could also contain two distracter dots in other positions. After a delay of 1100–1480 ms following the fifth grid, a final grid appeared containing a “?” in one of the 4 target positions. Participants were asked to indicate which of the 4 grids had contained a target dot in that probed position, using a button box. The children were encouraged to take their time but to try and respond as accurately as possible. As a consequence, the accuracy data are likely to be the most informative.

Each grid spanned $3.08^\circ \times 3.08^\circ$, and the targets and distracters spanned 0.53° in diameter. The targets were made of a red background (R:255, G:0, B:0) with a blue ring (R:0, G:0, B:255). The distracters were also defined in RGB space, with a yellow background and a cyan ring. On half the trials, these were defined as follows: R:255, G:77, B:0 and R:0, G:77, B:255, respectively. On the other half of trials, these were defined as follows: R:255, G:179, B:0 and R:0, G:77, B:255, respectively. We manipulated the hue of the distracters to see whether this would interact with the training manipulation (see also Shimi and Astle, 2013; Shimi et al., 2015).

MEG data were recorded throughout this task. Below we describe the basic data acquisition and preprocessing steps, an initial analysis, before proceeding to phase-coupling analysis. Before embarking on the phase amplitude analysis, we first wanted to establish that, during task performance, there is an ongoing oscillatory process, at the same frequencies, and in the same network as were implicated in the original resting-state result (reported by Astle et al., 2015).

MEG data acquisition and basic preprocessing. MEG data were acquired with a high-density whole-head VectorView MEG system (Elekta-Neuromag), containing a magnetometer and 2 orthogonal planar gradiometers at 102 positions (306 sensors in total), housed in a magnetically shielded room. Data were sampled at 1 kHz, and signals slower than 0.01 Hz were not recorded. A 3D digitizer (Fastrack Polhemus) was used to record the positions of 5 head position indicator coils and 50–100 additional points evenly distributed over the scalp, all relative to the nasion and left and right preauricular points. We also attached an electrode to each wrist to measure the pulse, and bipolar electrodes to obtain horizontal and vertical electro-oculograms. Head position was monitored throughout the recording using the head position indicator coils. Particularly small children were seated on a booster seat to ensure that their head was optimally positioned within the scanner helmet.

External noise was removed from the MEG data using the signal-space separation method, and adjustments in head position within the recording were compensated for using the MaxMove software, both implemented in MaxFilter version 2.1 (Elekta Neuromag). The continuous data were visually inspected, and any short sections with large signal jumps were removed. A sensor-space temporal Independent Component Analysis (ICA) was then used to remove artifacts arising from blinks, saccades, and pulse-related cardiac artifacts using a combination of metrics and manual inspection. The temporal ICA was conducted separately for each subject using fastICA run on the sensor space data, and then the time course of each IC was correlated with the time course of the vertical electro-oculograms, horizontal electro-oculograms, and cardiac channels, respectively. Components with a Pearson correlation value >0.1 with any of the artifact channels were subsequently removed from the data. Components dominated by 50 Hz noise were also removed. The data were then epoched into overlapping epochs running from -400 to

1000 ms relative to the onset of each target item in the sequence. These long, overlapping epochs meant we could later focus on a shorter period (0–600 ms relative to onset of each target item) without incurring edge effects in the time-frequency decomposition. The data were then checked by hand for any remaining trials that were likely contaminated by artifacts. The end result was that, for each child, we had on average 647 pretraining and 752 post-training epochs for the adaptive group, and 689 pretraining and 727 post-training epochs for the placebo group (there were no significant differences between the groups in terms of trial numbers, p values >0.19).

The 13–20 Hz network analysis. A resting-state analysis (reported by Astle et al., 2015), using a set of networks defined using an independent components analysis on fMRI data (Smith et al., 2012), identified one network as being particularly susceptible to working memory training effects. This network incorporated the left and right superior parietal cortex (MNI coordinates, left: $-24, -60, 56$; and right: $24, -64, 56$, respectively) and the bilateral frontal eye fields (MNI coordinates, left: $-28, -6, 56$; and right: $28, -2, 56$, respectively). In the resting-state data (Astle et al., 2015), we demonstrated that there was a strong ongoing oscillatory process at 13–20 Hz in this network. Before using the same network in the phase amplitude coupling analysis, we wanted to test whether this was also true in the task-related data.

To do this, we conducted an analysis identical to that used on the resting-state data (reported by Astle et al., 2015). For 23 of the children, we coregistered the MEG data to the child's T1-weighted structural MRI image acquired using a 3T Siemens Tim Trio and an MPRAGE sequence. For the remaining children who did not undergo an MRI scan, we used a standard MNI template for coregistration (as in Barnes et al., 2016). For this coregistration, we used the digitized scalp locations and fiducials, via an iterative closest point algorithm using SPM8 (<http://www.fil.ion.ucl.ac.uk/spm/>). The forward model was estimated using a single homogeneous shell model of the head shape of each subject (Mosher et al., 1999). This was also calculated using SPM8. The data were bandpass-filtered to capture activity in the 13–20 Hz band. We then applied a linearly constrained minimum variance beamformer (Van Veen et al., 1997). The beamformer combined information from both the magnetometers and planar gradiometers while taking into account the reduced dimensionality of the data introduced by the signal-space separation algorithm (Woolrich et al., 2011). This approach was used to produce a whole-brain source reconstruction, using a 6 mm dipole grid.

The data were then temporally concatenated, and the oscillatory amplitude envelopes of the reconstructed time series were estimated via computation of the absolute value of the analytic signal, which was found using a Hilbert transform. This yielded an estimate of instantaneous signal amplitude at each voxel. The envelope time series for every voxel were low-pass filtered by dividing each envelope time course into 0.8 s windows and averaging within those windows to focus on the low-frequency power fluctuations that have previously been shown to demonstrate functional connectivity in MEG data (Brookes et al., 2011).

We submitted our source-projected temporally concatenated downsampled envelope data to a temporal ICA. This ICA reduces the 8417 location-specific time courses to a set of 25 underlying components that have temporally distinct power fluctuations in the 13–20 Hz range. Importantly, the ICA is blind to the spatial locations of those original recordings. The resulting component time courses could then be correlated with the original location-specific envelope time courses to produce spatial maps of the underlying components. This is the same approach used by Brookes et al. (2011) to study resting-state networks using oscillatory data, and identical to that used by Astle et al. (2015).

Despite not providing the temporal ICA with any spatial information, we wanted to test whether one of the components would have a distribution of activity highly overlapping with the fMRI network used in the previously reported resting-state analysis (Astle et al., 2015), and which we would subsequently use in the phase amplitude coupling analysis. We used spatial cross-correlations (as implemented in *fsfcc*) (Woolrich et al., 2009) to calculate the degree of spatial overlap between the maps produced by the ICA procedure and the fMRI network. We selected the component map with the strongest cross-correlation value. This ap-

proach also allowed us to test the spatial overlap between this component and an equivalent component derived from the resting-state data using the same procedure.

MEG source reconstruction for phase amplitude coupling. Before beamforming, the data were bandpass-filtered to capture frequencies between 3 and 100 Hz. As a starting point, for each subject, we extracted source space activity for the dipoles (using a 6 mm grid) within our network of interest and region of interest (see Astle et al., 2015). For the single inferior temporal region, we concatenated all trials for each given subject, submitted the individual dipole time courses to a temporal PCA, and then used the weights from the first PC to combine the individual time courses to make a single region time course for each subject. This was then divided back into the individual epochs for the subsequent analyses. For the network of interest, we initially averaged across the dipoles within the network to produce the network time course: a PCA would potentially have biased the time course toward one particular location. However, in a subsequent analysis, we used the individual dipoles within the network, to specify which specific areas within the frontoparietal network most strongly show any coupling effect. All source reconstructions were performed using a linearly constrained minimum variance beamformer (Van Veen et al., 1997), as described above.

Time-frequency decomposition. We used a continuous Morlet Wavelet Transform, with three full cycles of each frequency, to decompose the signal into distinct time-frequency bins. We extracted phase and amplitude values between 3 Hz and 100 Hz in 1 Hz steps. This was done for the network time course, and for that of the inferior temporal region, for each subject. Once we had the frequency specific phase and amplitude values for these two signals, we could explore the extent to which they were coupled.

Cross-region phase amplitude coupling. We estimated the degree of cross-frequency coupling using circular-linear correlations (Berens, 2009). These were performed separately for each child, collapsing across trials and time within each trial, before and after training, for 3–20 Hz for phase and 30–100 Hz for amplitude. These coupling values were converted to z scores, using a standardized Fisher r -to- z transform, before a linear regression analysis. The result was that, for each child, before and after training, we had estimates of phase amplitude coupling across our frequency range. These values were then submitted to a linear regression analysis.

Linear regression analysis. To ascertain whether the working memory training had any influence on the degree of cross-frequency coupling, we performed a linear regression. The linear regression included a regressor representing the interaction between group (Adaptive vs Placebo) and time (pretraining vs post-training). This produced a new dataset in which we could ascertain the effect of working memory training on the degree of coupling across the full spectrum of frequencies included in the analysis. For ease of interpretation, and to simplify the subsequent multiple-comparisons correction, the outputs of the linear regression step were expressed as standardized t values (as in Barnes et al., 2016). We tested for significance using a cluster-based permutation procedure that corrected for multiple comparisons.

Multiple-comparisons correction. We first thresholded the result of our linear regression analysis at $t = 2.3$, which approximates an uncorrected significance level of $p = 0.05$. We identified any clusters of consecutive comparisons (across consecutive phase or amplitude values) and calculated the size of these clusters: that is, the number of consecutive samples that surpass this arbitrary clustering threshold. We then adopted a permutation procedure to build a null distribution of the size of clusters that we would expect by chance. To do this, we performed an identical analysis to that described above, save for the final step. Rather than using the actual regressor of time (pre vs post) and group (Adaptive vs Placebo), we randomly shuffled the values in that regressor (i.e., we randomly assigned data to pre vs post and Adaptive vs Placebo conditions). We again thresholded the results and recorded the size of the largest cluster. This process was repeated 1000 times, with each permutation having a new shuffled allocation of the conditions. Once complete, this provided a null distribution of cluster size, against which we could compare any clusters from our original analysis, to produce a p value. This randomization procedure is an extremely powerful way of testing for significance in an analysis

such as this. Any nonlinearities in the data or biases occurring as a result of the various necessary transforms are corrected for because they are recreated perfectly in each permutation. Furthermore, it makes no a priori assumptions about when or where effects are likely to be apparent, correcting for multiple comparisons over both frequency for phase and frequency for amplitude (Kilner, 2013).

This sequence of steps forms our basic phase amplitude analysis pipeline. In Results, we also report the findings from various follow-up analyses, including control regions and a modified version of the pipeline that provides a whole-brain analysis.

Individual differences. In addition to the group average approach described above in the basic coupling pipeline, we also explored individual differences in training effects. A general linear model was tested with gain in phase amplitude coupling as the outcome variable and gain in in-scanner accuracy as the predictor. We also applied a general linear model, with change in phase amplitude coupling as the outcome measure, and change in resting-state connectivity as the predictor (taken from Astle et al., 2015). In both of these models, we included age, gender, and IQ as control variables. Both models were applied across the time course of the phase amplitude coupling, to build a time course of the effect of the respective predictor variable, expressed as t scores. In both cases, we used a clustering threshold of $t > 2.3$ and then tested the extent to which clusters of this duration would occur by chance, using the same permutation procedure described above. In both analyses, we removed one child because they had a Cook's Distance score > 1 , suggesting that their data were having undue influence on the size of the relationship.

Alternative analysis of phase amplitude coupling. In view of the lack of previous studies exploring neurophysiological coupling in childhood, we also used an alternative approach to corroborate the above analysis. Phase estimates from the frontoparietal network at 16 Hz and amplitude estimates from the region in inferior temporal cortex at 83–93 Hz were aligned for each subject, again collapsing across time and trials. The amplitude estimates were then sorted, in order, according to the corresponding phase estimates, and converted into z scores. Because there are nonidentical trial numbers across subjects, we then divided these sorted values into equally spaced bins of 0.0016 radians per bin and averaged the amplitude values within each bin. This produced 2000 values for each subject, before and after training. If there is no consistent relationship between phase and amplitude, then these phase-ordered amplitude bins will be approximately uniform. However, if there is some consistent phase amplitude relationship, then these binned values ought to change systematically across the cycle of the phase angles. A nonparametric Wolf-Wolfowitz calculation was used to test the resulting distributions. This is a test of randomness. A significant statistic indicates that the phase-ordered amplitude values have a systematic order.

The beamformer source reconstruction includes a dipole sign ambiguity, which necessarily needs resolving for this shuffled data analysis to be meaningful, effects running in opposing directions across subjects, because of the sign ambiguity of the phase angles, will cancel one another out. For this reason, for each subject, before and after training, we organized the signs such that all of the phase amplitude relationships ran in the same direction. In this case, we opted for a negative relationship, with high gamma amplitudes being associated with negative phase angles of 16 Hz activity, because this inverse relationship is well reported in the literature (e.g., Haegens et al., 2011). A downside of this approach is that it will match random noise, as well as any genuine coupling effect, across subjects. It is important to consider that this procedure was identical across the different conditions and groups, so if this introduces a bias, then it ought to be apparent and of an equivalent magnitude across conditions and groups.

Results

Behavioral results

The behavioral results section starts with an analysis of the bespoke in-scanner task. We used the pretraining data to explore task performance, and the extent to which it varied by serial order

(whether the first, second, third, or fourth item was probed) and distracter type. We subsequently tested for training effects in both the standardized assessments and the in-scanner task.

In-scanner task performance

Accuracy was significantly higher on the “dissimilar” (67%) distracter trials relative to “similar” (61%) distracter trials ($F_{(1,25)} = 12.823$, $p < 0.001$). We also observed a recency effect, with accuracy improving for the later items (serial position 1, 53%; position 2, 55%; position 3, 66%; position 4, 84%) in the sequence ($F_{(3,75)} = 50.457$, $p < 0.001$). These two factors interacted significantly ($F_{(3,75)} = 3.142$, $p = 0.030$), with the difference between distracter trial types being significant for the first and last serial positions ($t_{(26)} = 3.638$, $p = 0.001$, and $t_{(26)} = 3.525$, $p = 0.002$, respectively), whereas they were not significantly different for the second and third serial positions ($t_{(26)} = 0.512$, $p = 0.613$, and $t_{(26)} = 1.384$, $p = 0.178$, respectively).

To foreshadow the training findings, although the training had an overall effect on performance, it did not significantly interact with either of these factors (p values > 0.5). For simplicity, in subsequent analyses, we collapse across these factors.

Standardized assessments of short-term and working memory

Consistent with a number of recent studies of cognitive training, we explored the potential impact of the working memory training on the standardized assessments using a hierarchical multiple regression approach. This has the advantage of taking into account initial pretraining performance differences across subjects. Post-training accuracy was our outcome measure. In the first block, we entered pretraining accuracy. In the second block, we entered group (adaptive vs placebo). This first level controls for variability in performance that can be accounted for by baseline differences in ability, and the second level then tests whether the remaining differences across the children can be accounted for by group membership. In short, it tests whether improvements in task performance across children can be explained by the type of training the children experienced. When tested in this way, we observed significant effects of group membership on all subtests, apart from forwards digit recall: Digit Recall (R^2 change = 0.06, $F_{(1,23)} = 3.309$, $p = 0.082$, standardized beta coefficient = 0.245); Dot Matrix (R^2 change = 0.34, $F_{(1,24)} = 19.414$, $p < 0.001$, standardized beta coefficient = 0.584); Backwards Digit Recall (R^2 change = 0.14, $F_{(1,23)} = 4.254$, $p = 0.051$, standardized beta coefficient = 0.375); and Spatial Span (R^2 change = 0.278, $F_{(1,24)} = 11.605$, $p = 0.002$, standardized beta coefficient = 0.528). A composite memory score, before and after training, can be seen Figure 1B.

In-scanner behavioral task performance

The impact of the working memory training on in-scanner performance accuracy was assessed using the same linear regression approach as described above. Group membership was a significant predictor of post-training accuracy levels, even after accounting for pretraining ability differences ($R^2 = 0.13$; $F_{(1,24)} = 7.090$, $p = 0.014$, standardized beta coefficient = 0.359) (Fig. 1C).

Our five untrained working memory tasks included both simple and complex span measures, and verbal and visuo-spatial paradigms. They also included elements that are structurally different from the training activities, such as interleaved encoding and maintenance episodes and probed-recall (the in-scanner task). Apart from Forwards Digit Span, these all showed signifi-

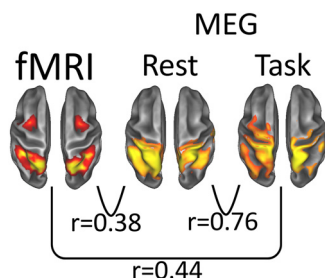


Figure 2. Results of the 13–20 Hz temporal ICA analysis, in both the resting state and task-related data, alongside the network from the independent fMRI dataset. The correlation values refer to spatial cross-correlations.

cant improvements in the Adaptive group, relative to the Placebo group, replicating the well-established near transfer to untrained working memory exercises (Dunning et al., 2013). We then explored whether neurophysiological coupling during task performance had been altered in the Adaptive training group relative to the Placebo control group.

MEG results

The 13–20 Hz network analysis

We submitted the source projected activity time courses in the 13–20 Hz frequency band to a temporal ICA. Despite not having any prior spatial information about the location of these recordings, the ICA produced a component highly overlapping with the network taken from the fMRI dataset (Smith et al., 2012), which was implicated in our previous resting-state findings (Aster et al., 2015), and which would subsequently use in the phase amplitude analysis. The component shown below is that with the strongest spatial overlap ($r = 0.38$, $z = 0.40$, $p < 0.001$). The fact that this component is produced by the spatially blind ICA demonstrates that there is a robust network with oscillatory power within the 13–20 Hz frequency band and incorporating the particular areas contained within the network, during the task. An identical analysis was conducted on the resting-state analysis (Aster et al., 2015), and we can also demonstrate that there is a strong spatial overlap with the equivalent independent component from that dataset ($r = 0.76$, $z = 0.99$, $p < 0.001$). These networks can be seen in Figure 2.

Cross-regional phase amplitude coupling

Our first phase amplitude coupling analysis of the MEG data used the above described network, alongside a region in inferior temporal cortex, also identified in the resting state result (Aster et al., 2015).

Following training, children in the adaptive training versus placebo group showed increased cross-frequency phase amplitude coupling between lower beta/upper alpha-band oscillations in the frontoparietal network and the upper-gamma band oscillations in the region of interest within inferior temporal cortex. This significant cluster was at 16 Hz for phase, and from 83 to 93 Hz for amplitude ($p_{\text{corrected}} < 0.001$) (Fig. 3A). We also estimated the time course of this coupling by taking these frequencies and recalculating the coupling for a 400 ms sliding window, which incremented in 4 ms steps, enabling us to build a time course of the cross-frequency relationship. This was done separately for both groups, before and after training (Fig. 3B).

As a control, we also performed the coupling analysis, but using both the phase and amplitude estimates from the same frontoparietal network. If our result reflects a general effect of training on coupling at these frequency bands, rather than any

cross-regional effect, then the same effect should be apparent (or to an even stronger degree) when we take the phase angles and amplitude values from the same regions. However, there was no significant training effect on coupling in this case (Fig. 3C).

In addition to this group-average approach, we also explored individual differences. The change in frontoparietal to inferior temporal phase amplitude coupling between 124 and 174 ms ($p_{\text{corrected}} < 0.001$) and between 556 and 600 ms ($p_{\text{corrected}} < 0.001$) was significantly predictive of children's improved performance on the in-scanner task (Fig. 4A). This relationship can be seen in the bar graph in Figure 4B and the scatter plot in Figure 4C. Between 184 and 208 ms, the enhancement in task-related phase amplitude coupling was also predicted by the individual gains in resting-state connectivity ($p_{\text{corrected}} < 0.001$), as taken from Aster et al. (2015) (Fig. 4D).

Alternative analysis of phase amplitude coupling

An alternative means of demonstrating this cross-frequency coupling, between frontoparietal network phase angles (at 16 Hz) and the inferior temporal amplitude values (from 83 to 93 Hz), is to plot amplitude values sorted by the phase angles. For each subject, before and after training, we aligned the phase and amplitude values, and then sorted the two vectors by the phase angles in ascending order. The resulting phase-ordered amplitude values were then divided into equally spaced bins. Figure 5 shows these bins, for both groups, before and after training. Following training, in the Adaptive group, the sorted amplitude values were significantly nonuniform ($z = -2.59$, $p = 0.0095$). By contrast, the sorted amplitude values did not differ significantly from a random distribution before training ($z = -1.09$, $p = 0.27$), or for either pre or post training in the Placebo group ($z = -0.44$, $p = 0.65$; and $z = -0.69$, $p = 0.49$, respectively).

Induced analysis

A systematic pattern of ongoing or evoked activity within these two areas, at these particular frequency bands (upper alpha/lower beta and high gamma), that is sensitive to training, might give the false appearance of a training-induced enhancement to coupling. To test this, we performed an induced analysis, using the -300 to -100 time period relative to each stimulus as a baseline period. The resulting mean induced responses for both our network of interest (left-hand panel) and our region of interest (right-hand panel) can be seen in Figure 6A. We next tested for any significant changes as a result of training at the frequencies from our original coupling analysis (16 Hz and 83–93 Hz). To do this, we repeated the same linear regression analysis that we had applied to the coupling data. The results of this can be seen in Figure 6B. None of these effects approached our clustering threshold. The increased coupling following training does not stem from changes to the evoked or background rhythms in these frequency bands.

Further specifying areas within frontoparietal cortex that underpin coupling

In our initial analyses, we used the phase angles from numerous locations within the frontoparietal network. However, any more specific contribution is masked by us treating this collection of dipoles as a single network. To test this, we repeated our initial analysis for each dipole separately. To do this, we restricted this analysis to just the frequencies implicated in the initial analysis: 90 Hz for power in the inferior temporal cortex

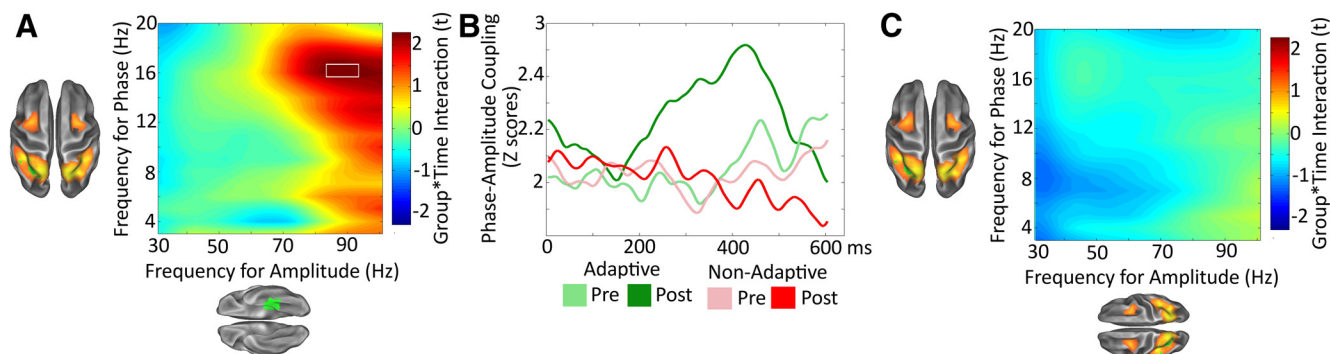


Figure 3. Cross-frequency phase amplitude coupling. **A**, Impact of training (training group \times time interaction) for all combinations of phase from the dorsal frontoparietal network (3–20 Hz) and amplitude from the inferior temporal region (30–100 Hz). White box represents values where $t > 2.3$. **B**, Time course of the phase amplitude (at 16 and 83–93 Hz, respectively) coupling shown in **A**, shown separately for the two groups, before and after training. This is calculated using a 400 ms window, which was incremented along the full dataset in 4 ms increments. For plotting purposes, 40 ms temporal smoothing was also applied. **C**, Identical analysis to that in **A**, but with both phase and amplitude estimates taken from the dorsal frontoparietal network.

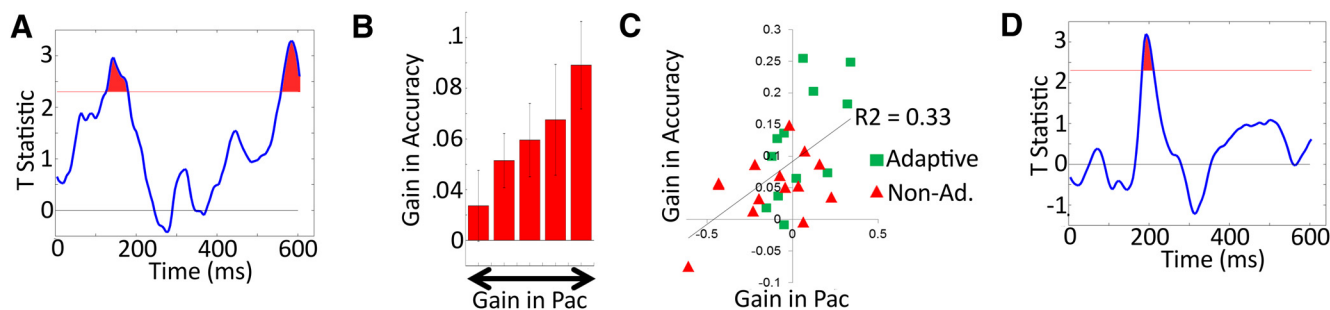


Figure 4. Individual differences plots. **A**, Time course of the relationship between changes in phase amplitude coupling and changes in performance, controlling for age, IQ, and gender. **B**, Behavioral gains on the in-scanner tasks of five groups of children, sorted on the basis of the size of change in phase amplitude coupling (collapsing across the two time windows highlighted in **A**). Error bars indicate SEM. **C**, Same data as in **B**, but depicted as a scatter plot. The partial correlation, controlling for age, IQ, and gender, is $r = 0.64$ ($p = 0.001$). **D**, Time course of the relationship between changes in phase amplitude coupling and changes in resting state connectivity taken from Astle et al. (2015), controlling for age, IQ, and gender.

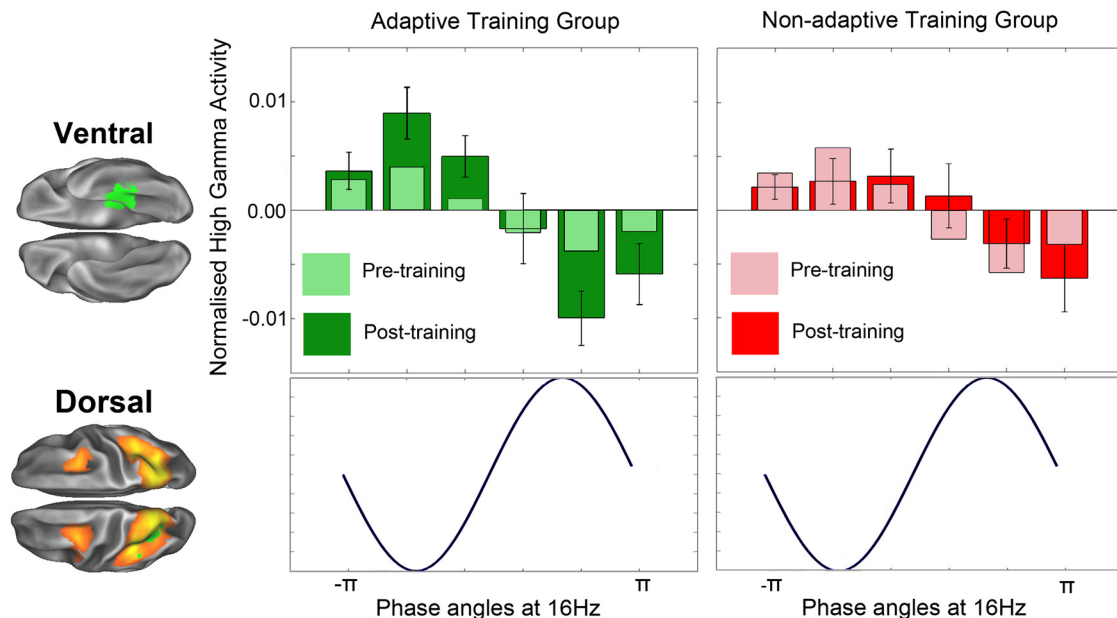


Figure 5. Phase-ordered amplitude plots. High gamma (83–93 Hz) inferior temporal cortex, replotted according to their position in the cycle of the 16 Hz rhythm from the superior frontoparietal network. For plotting purposes, these are divided into 6 equally spaced bins. However, for the statistical analysis, 2000 bins are used to provide a more detailed distribution. Error bars indicate SEM.

and 16 Hz for phase at each dipole within the frontoparietal network. We applied the same phase amplitude coupling and linear regression analysis steps as before. Figure 7 highlights the areas within the frontoparietal network that contribute

most strongly to this coupling effect. The areas shaded in blue highlight regions within the frontoparietal network where the training effect on the coupling exceeds $t = 2.3$ (our clustering threshold in the initial analysis).

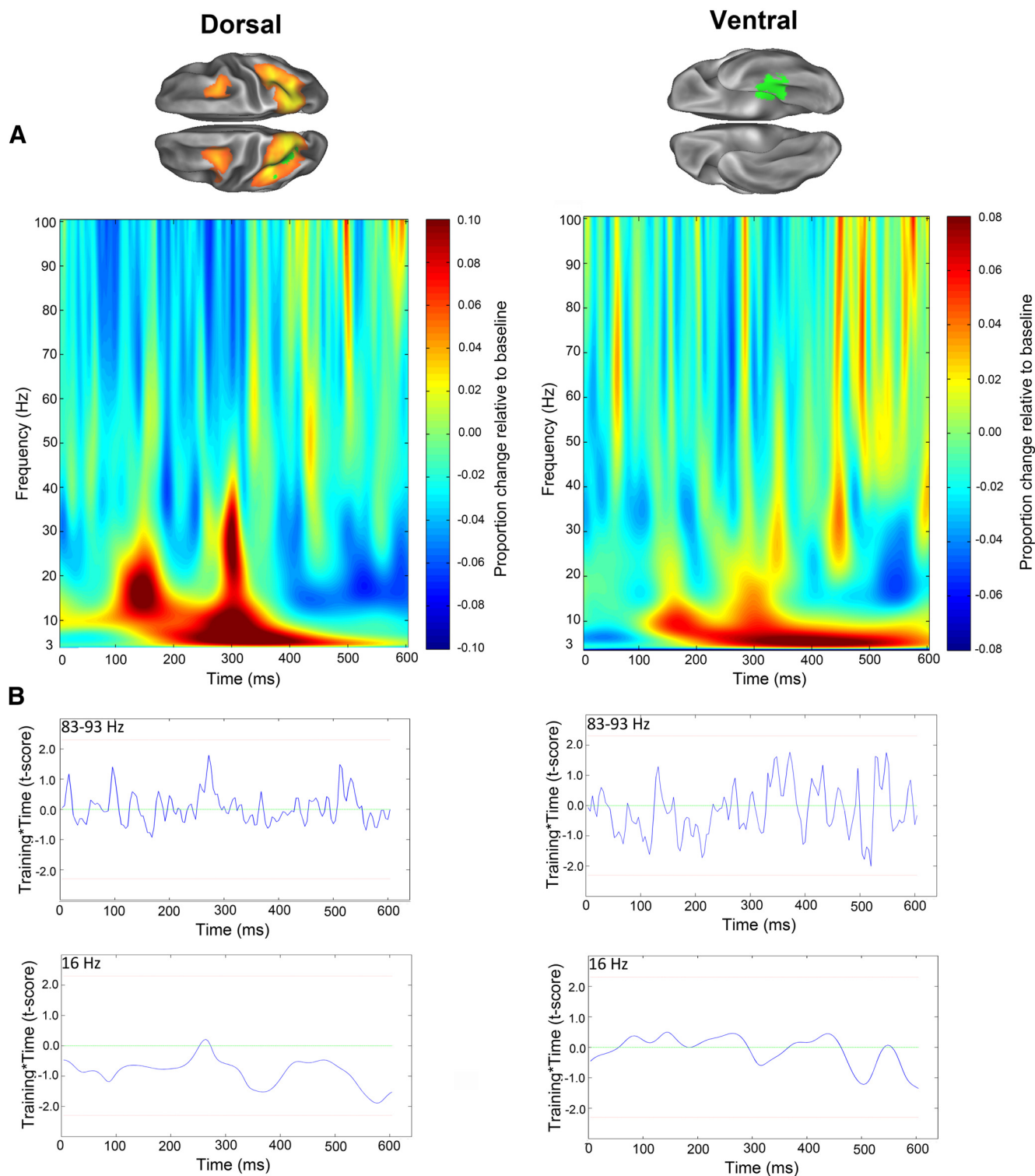


Figure 6. Induced analysis. **A**, Induced analysis, from 3–100 Hz, of activity in the dorsal frontoparietal network (left) and the region of inferior temporal cortex (right). In both cases, the –300 to –100 ms window, relative to target stimulus onset at 0 ms, was used as a baseline period. **B**, Linear regression analysis applied to 83–93 Hz and 16 Hz activity from the dorsal frontoparietal network (left panels) and region in inferior temporal cortex (right panels). Results are expressed as t statistics for the training group \times time interaction. In all plots, red line indicates our uncorrected threshold for the cluster-based permutation correction ($t > 2.3$).

Whole-brain phase amplitude coupling

We also explored whether high-frequency gamma-band activity in areas outside our inferior temporal region of interest showed training-induced coupling changes. To do this, we used the phase angles in the frontoparietal network in the alpha-lower-beta range (16 ± 1 Hz) and used the beamformer

to create a whole-brain reconstruction (using a 6 mm dipole grid) of gamma amplitudes (at 90 Hz). We then performed the coupling and linear regression steps exactly as before, but across all locations in our source reconstruction. We then passed the resulting values to a cluster-based multiple-comparisons permutations correction. This was similar to that

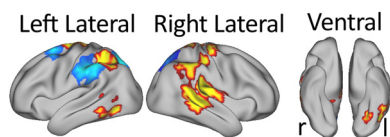


Figure 7. Whole-brain analysis. A whole-brain phase amplitude coupling analysis between 16 Hz phase within the frontoparietal network and 90 Hz activity estimated throughout the brain. This is whole-brain corrected for multiple comparisons using a cluster-based permutation correction procedure. All results are $p_{\text{corrected}} < 0.001$.

previously used, but with the clusters defined in spatial rather than frequency terms.

Following this procedure, we identified a set of significant clusters (all at $p_{\text{corrected}} < 0.001$). These can be seen in Figure 7, with the significant clusters shown in red. These clusters include the left inferior temporal cortex, as well as regions within the right hemisphere, bilateral inferior parietal activity, and some premotor areas. Following intensive working memory training, relative to placebo, high gamma activity in these regions became better coupled to the phase of the 16 Hz activity within the frontoparietal network during task performance.

Discussion

A child's working memory capacity is an excellent predictor of their educational progress (Gathercole et al., 2004), and impairments in working memory are increasingly seen as core to numerous neurodevelopmental disorders (Westerberg et al., 2004; Martinussen et al., 2005). We show that over 20 sessions of intensive adaptive training, children's working memory significantly improves, relative to that of an active placebo control group.

Having replicated the well-established near transfer working memory training effect (Holmes et al., 2009, 2010; Dunning et al., 2013), our MEG data identified significant changes in underlying neurophysiology that mirror these behavioral improvements. We tested one particular account of neuronal integration, called phase amplitude coupling. Within this account, slow oscillations organize cortical excitability within a distributed network, and this provides a mechanism by which one set of brain regions can regulate the activity of other cortical areas. In our data, the upper alpha rhythm in superior frontal and parietal areas provided a temporal structure within which high-frequency activity in inferior parietal, temporal and high-order visual processing areas became increasingly nested following training. That our whole-brain analysis highlighted a number of areas in inferior temporal and inferior parietal, which show a similar pattern of coupling, suggests that the 16 Hz rhythm generated by frontoparietal cortex may play a role in coordinating the ongoing processing across a number of different cortical areas. Enhancement following training was not just apparent at a group level; individual changes in coupling were predictive of individual gains in performance on standardized assessments. Moreover, the magnitude of task-related coupling was significantly associated with previously observed changes in resting-state activity (Astle et al., 2015).

A limitation of our approach is that it is not an exhaustive search of all potential coupling changes following training. Using the resting-state results to guide our task-related coupling analysis could mean that other changes that are not spatially overlapping with these original effects are missed. An alternative approach would be to conduct a fully exploratory analysis, across all locations and all frequencies. Our reason for not doing this

here is that doing so would introduce a massive multiple-comparison problem.

In previous work, the regulatory role of alpha-band activity has been related to the attentional modulation of the sensitivity of sensory processes (e.g., Jensen et al., 2012). Working memory training may enhance these hierarchical coupling relationships because it heavily taxes these attention control systems (Gazzaley and Nobre, 2012). At a cellular level, the alpha rhythm provides this temporal structure by modulating the excitability of local neuronal assemblies (Fries, 2005; Canolty et al., 2006, 2010, 2012; Canolty and Knight, 2010). This is possible because these rhythmic changes alter a neuron's output spike timing and its sensitivity to synaptic input (Canolty et al., 2010), and these local neuronal computations are apparent as ongoing gamma activity (Fries, 2009).

Changes in synchrony with development have been shown across childhood and adolescence and are tied to important changes in cognitive ability (Uhlhaas et al., 2010; Uhlhaas and Singer, 2011). A strong possibility is that these changes in neuronal synchrony mirror underlying changes in white matter integrity, as nodes within neural networks become better integrated with age (Barnea-Goraly et al., 2005; Hagmann et al., 2010). For example, studies combining MEG with structural brain imaging have shown that white matter pathways play a strong role in coordinating spontaneous alpha oscillations at rest (Hindriks et al., 2015). Specifically, in the case of attentional modulation, the integrity of particular tracks (e.g., the super longitudinal fasciculus, a white matter tract connecting frontal and parietal areas) significantly predict subjects' degree of alpha and gamma synchrony (Marshall et al., 2015). A strong possibility is that the changes we observe in neuronal synchrony are matched by corresponding changes in structural connectivity (Takeuchi et al., 2010; Caeyenberghs et al., 2016). It remains unclear whether these changes are akin to those associated with typical development or whether they reflect highly specific reorganization in response to the particular working memory activities that we trained.

Some have made strong claims that working memory training results in fundamental changes in the general attentional control of working memory (Klingberg, 2010). Other theorists (e.g., Sprenger et al., 2013) have suggested that transfer is best understood in terms of overlap in specific processes between trained and untrained tasks, implying multiple potential sources of training-induced change. These may in part reflect the development and refinement of metacognitive strategies through intensive training (e.g., Gibson et al., 2013; Dunning and Holmes, 2014). We do not view these accounts as necessarily mutually exclusive; the neurophysiological changes that we observe could be in conjunction with, or distinct from, the development or refinement of particular strategies. Future studies capable of distinguishing different strategies, alongside neurophysiological measures, will be needed to explore whether and how these two levels of analysis are related.

In conclusion, through our cross-frequency phase amplitude coupling analyses of task-related MEG data collected before and after training, we provide a detailed neurophysiological account of gains in working memory following training. This has broad implications for our understanding of these neurophysiological processes, their plasticity, their role in higher-order cognition, and the applicability of these interventions.

References

- Alloway BJ (2009) Soil factors associated with zinc deficiency in crops and humans. *Environ Geochem Health* 31:537–548. [CrossRef Medline](#)
- Alloway TP, Gathercole SE, Kirkwood H, Elliott J (2008) Evaluating the

- validity of the automated working memory assessment. *Educ Psychol* 28:725–734. [CrossRef](#)
- Astle DE, Barnes JJ, Baker K, Colclough GL, Woolrich MW (2015) Cognitive training enhances intrinsic brain connectivity in childhood. *J Neurosci* 35:6277–6283. [CrossRef Medline](#)
- Axmacher N, Henseler MM, Jensen O, Weinreich I, Elger CE, Fell J (2010) Cross-frequency coupling supports multi-item working memory in the human hippocampus. *Proc Natl Acad Sci U S A* 107:3228–3233. [CrossRef Medline](#)
- Barnea-Goraly N, Menon V, Eckert M, Tamm L, Bammer R, Karchemskiy A, Dant CC, Reiss AL (2005) White matter development during childhood and adolescence: a cross-sectional diffusion tensor imaging study. *Cereb Cortex* 15:1848–1854. [CrossRef Medline](#)
- Barnes JJ, Woolrich MW, Baker K, Colclough GL, Astle DE (2016) Electrophysiological measures of resting state functional connectivity and their relationship with working memory capacity in childhood. *Dev Sci* 19:19–31. [CrossRef Medline](#)
- Berens P (2009) CircStat: a MATLAB toolbox for circular statistics. *J Stat Software* 31:1–21.
- Brookes MJ, Woolrich M, Luckhoo H, Price D, Hale JR, Stephenson MC, Barnes GR, Smith SM, Morris PG (2011) Investigating the electrophysiological basis of resting state networks using magnetoencephalography. *Proc Natl Acad Sci USA* 108:16783–16788. [CrossRef Medline](#)
- Caeyenberghs K, Metzler-Baddeley C, Foley S, Jones DK (2016) Dynamics of the human structural connectome underlying working memory training. *J Neurosci* 36:4056–4066. [CrossRef Medline](#)
- Canolty RT, Knight RT (2010) The functional role of cross-frequency coupling. *Trends Cogn Sci* 14:506–515. [CrossRef Medline](#)
- Canolty RT, Edwards E, Dalal SS, Soltani M, Nagarajan SS, Kirsch HE, Berger MS, Barbaro NM, Knight RT (2006) High gamma power is phase-locked to theta oscillations in human neocortex. *Science* 313:1626–1628. [CrossRef Medline](#)
- Canolty RT, Ganguly K, Kennerley SW, Cadieu CF, Koepsell K, Wallis JD, Carmena JM (2010) Oscillatory phase coupling coordinates anatomically dispersed functional cell assemblies. *Proc Natl Acad Sci U S A* 107:17356–17361. [CrossRef Medline](#)
- Canolty RT, Ganguly K, Carmena JM (2012) Task-dependent changes in cross-level coupling between single neurons and oscillatory activity in multiscale networks. *Plos Comput Biol* 8:e1002809. [CrossRef Medline](#)
- Cohen MX, Elger CE, Fell J (2009) Oscillatory activity and phase amplitude coupling in the human medial frontal cortex during decision making. *J Cogn Neurosci* 21:390–402. [CrossRef Medline](#)
- Cravo AM, Rohenkohl G, Wyart V, Nobre AC (2011) Endogenous modulation of low frequency oscillations by temporal expectations. *J Neurophysiol* 106:2964–2972. [CrossRef Medline](#)
- Desimone R, Duncan J (1995) Neural mechanisms of selective visual attention. *Annu Rev Neurosci* 18:193–222. [CrossRef Medline](#)
- Duncan J, Owen AM (2000) Common regions of the human frontal lobe recruited by diverse cognitive demands. *Trends Neurosci* 23:475–483. [CrossRef Medline](#)
- Dunning DL, Holmes J (2014) Does working memory training promote the use of strategies on untrained working memory tasks? *Mem Cognit* 42:854–862. [CrossRef Medline](#)
- Dunning DL, Holmes J, Gathercole SE (2013) Does working memory training lead to generalized improvements in children with low working memory? A randomized controlled trial. *Dev Sci* 16:915–925. [CrossRef Medline](#)
- Fries P (2005) A mechanism for cognitive dynamics: neuronal communication through neuronal coherence. *Trends Cogn Sci* 9:474–480. [CrossRef Medline](#)
- Fries P (2009) Neuronal gamma-band synchronization as a fundamental process in cortical computation. *Annu Rev Neurosci* 32:209–224. [CrossRef Medline](#)
- Gathercole SE, Pickering SJ, Knight C, Stegmann Z (2004) Working memory skills and educational attainment: evidence from national curriculum assessments at 7 and 14 years of age. *Appl Cogn Psychol* 18.
- Gazzaley A, Nobre AC (2012) Top-down modulation: bridging selective attention and working memory. *Trends Cogn Sci* 16:129–135. [CrossRef Medline](#)
- Gibson BS, Gondoli DM, Kronenberger WG, Johnson AC, Steeger CM, Morrissey RA (2013) Exploration of an adaptive training regimen that can target the secondary memory component of working memory capacity. *Mem Cognit* 41:726–737. [CrossRef Medline](#)
- Gregoriou GG, Gotts SJ, Zhou H, Desimone R (2009) High-frequency, long-range coupling between prefrontal and visual cortex during attention. *Science* 324:1207–1210. [CrossRef Medline](#)
- Haegens S, Nacher V, Luna R, Romo R, Jensen O (2011) alpha-Oscillations in the monkey sensorimotor network influence discrimination performance by rhythmic inhibition of neuronal spiking. *Proc Natl Acad Sci U S A* 108:19377–19382. [CrossRef Medline](#)
- Hagmann P, Sporns O, Madan N, Cammoun L, Pienaar R, Wedeen VJ, Meuli R, Thiran JP, Grant PE (2010) White matter maturation reshapes structural connectivity in the late developing human brain. *Proc Natl Acad Sci U S A* 107:19067–19072. [CrossRef Medline](#)
- Hindriks R, Woolrich M, Luckhoo H, Joensuu M, Mohseni H, Kringelbach ML, Deco G (2015) Role of white-matter pathways in coordinating alpha oscillations in resting visual cortex. *Neuroimage* 106:328–339. [CrossRef Medline](#)
- Holmes J, Gathercole SE, Dunning DL (2009) Adaptive training leads to sustained enhancement of poor working memory in children. *Dev Sci* 12:F9–F15. [CrossRef Medline](#)
- Holmes J, Gathercole SE, Place M, Dunning DL, Hilton KA, Elliott JG (2010) Working memory deficits can be overcome: impacts of training and medication on working memory in children with ADHD. *Appl Cogn Psychol* 24:827–836. [CrossRef](#)
- Jensen O, Colgin LL (2007) Cross-frequency coupling between neuronal oscillations. *Trends Cogn Sci* 11:267–269. [CrossRef Medline](#)
- Jensen O, Bonnefond M, VanRullen R (2012) An oscillatory mechanism for prioritizing salient unattended stimuli. *Trends Cogn Sci* 16:200–206. [CrossRef Medline](#)
- Kilner JM (2013) Bias in a common EEG and MEG statistical analysis and how to avoid it. *Clin Neurophysiol* 124:2062–2063. [CrossRef Medline](#)
- Klingberg T (2010) Training and plasticity of working memory. *Trends Cogn Sci* 14:317–324. [CrossRef Medline](#)
- Kuo BC, Yeh YY, Chen AJ, D'Esposito M (2011) Functional connectivity during top-down modulation of visual short-term memory representations. *Neuropsychologia* 49:1589–1596. [CrossRef Medline](#)
- Lakatos P, Shah AS, Knuth KH, Ulbert I, Karmos G, Schroeder CE (2005) An oscillatory hierarchy controlling neuronal excitability and stimulus processing in the auditory cortex. *J Neurophysiol* 94:1904–1911. [CrossRef Medline](#)
- Marshall TR, Bergmann TO, Jensen O (2015) Frontoparietal structural connectivity mediates the top-down control of neuronal synchronization associated with selective attention. *PLoS Biol* 13:e1002272. [CrossRef Medline](#)
- Martinussen R, Hayden J, Hogg-Johnson S, Tannock R (2005) A meta-analysis of working memory impairments in children with attention-deficit/hyperactivity disorder. *J Am Acad Child Adolesc Psychiatry* 44:377–384. [CrossRef Medline](#)
- Mosher JC, Leahy RM, Lewis PS (1999) EEG and MEG: forward solutions for inverse methods. *IEEE Trans Biomed Eng* 46:245–259. [CrossRef Medline](#)
- Sauseng P, Klimesch W, Gruber WR, Birbaumer N (2008) Cross-frequency phase synchronization: a brain mechanism of memory matching and attention. *Neuroimage* 40:308–317. [CrossRef Medline](#)
- Shimi A, Astle DE (2013) The strength of attentional biases reduces as visual short-term memory load increases. *J Neurophysiol* 110:12–18. [CrossRef Medline](#)
- Shimi A, Woolrich MW, Mantini D, Astle DE (2015) Memory load modulates graded changes in distracter filtering. *Front Hum Neurosci* 8:1025. [CrossRef Medline](#)
- Siegel M, Warden MR, Miller EK (2009) Phase-dependent neuronal coding of objects in short-term memory. *Proc Natl Acad Sci U S A* 106:21341–21346. [CrossRef Medline](#)
- Smith SM, Miller KL, Moeller S, Xu J, Auerbach EJ, Woolrich MW, Beckmann CF, Jenkinson M, Andersson J, Glasser MF, Van Essen DC, Feineberg DA, Yacoub ES, Ugurbil K (2012) Temporally-independent functional modes of spontaneous brain activity. *Proc Natl Acad Sci USA* 109:3131–3136. [CrossRef](#)
- Sprenger AM, Atkins SM, Bolger DJ, Harbison JL, Novick JM, Chrabaszcz JS, Weems SA, Smith V, Bobb S, Bunting MF, Dougherty MR (2013) Training working memory: limits of transfer. *Intelligence* 41:638–663. [CrossRef](#)
- Takeuchi H, Sekiguchi A, Taki Y, Yokoyama S, Yomogida Y, Komuro N, Yamanouchi T, Suzuki S, Kawashima R (2010) Training of working memory impacts structural connectivity. *J Neurosci* 30:3297–3303. [CrossRef Medline](#)
- Uhlhaas PJ, Singer W (2011) The development of neural synchrony and

- large-scale cortical networks during adolescence: relevance for the pathophysiology of schizophrenia and neurodevelopmental hypothesis. *Schizophr Bull* 37:514–523. [CrossRef Medline](#)
- Uhlhaas PJ, Roux F, Rodriguez E, Rotarska-Jagiela A, Singer W (2010) Neural synchrony and the development of cortical networks. *Trends Cogn Sci* 14:72–80. [CrossRef Medline](#)
- Van Veen BD, van Drongelen W, Yuchtman M, Suzuki A (1997) Localization of brain electrical activity via linearly constrained minimum variance spatial filtering. *IEEE Tran Biomed Eng* 44:867–880. [CrossRef Medline](#)
- Voytek B, Canolty RT, Shestyuk A, Crone NE, Parvizi J, Knight RT (2010) Shifts in gamma phase amplitude coupling frequency from theta to alpha over posterior cortex during visual tasks. *Front Hum Neurosci* 4:191. [CrossRef Medline](#)
- Westerberg H, Hirvikoski T, Forssberg H, Klingberg T (2004) Visuo-spatial working memory span: a sensitive measure of cognitive deficits in children with ADHD. *Child Neuropsychol* 10:155–161. [CrossRef Medline](#)
- Woolrich MW, Jbabdi S, Patenaude B, Chappell M, Makni S, Behrens T, Beckmann C, Jenkinson M, Smith SM (2009) Bayesian analysis of neuroimaging data in FSL. *Neuroimage* 45:S173–S186. [CrossRef Medline](#)
- Woolrich M, Hunt L, Groves A, Barnes G (2011) MEG beamforming using Bayesian PCA for adaptive data covariance matrix regularization. *Neuroimage* 57:1466–1479. [CrossRef Medline](#)