LETTER

doi:10.1038/nature12486

Video game training enhances cognitive control in older adults

J. A. Anguera[1,2,3], J. Boccanfuso[1,3], J. L. Rintoul[1,3], O. Al-Hashimi[1,2,3], F. Faraji[1,3], J. Janowich[1,3], E. Kong[1,3], Y. Larraburo[1,3], C. Rolle[1,3], E. Johnston[1] & A. Gazzaley[1,2,3,4]

Cognitive control is defined by a set of neural processes that allow us to interact with our complex environment in a goal-directed manner[1]. Humans regularly challenge these control processes when attempting to simultaneously accomplish multiple goals (multitasking), generating interference as the result of fundamental information processing limitations[2]. It is clear that multitasking behaviour has become ubiquitous in today's technologically dense world[3], and substantial evidence has accrued regarding multitasking difficulties and cognitive control deficits in our ageing population[4]. Here we show that multitasking performance, as assessed with a custom-designed three-dimensional video game (NeuroRacer), exhibits a linear age-related decline from 20 to 79 years of age. By playing an adaptive version of NeuroRacer in multitasking training mode, older adults (60 to 85 years old) reduced multitasking costs compared to both an active control group and a no-contact control group, attaining levels beyond those achieved by untrained 20-year-old participants, with gains persisting for 6 months. Furthermore, age-related deficits in neural signatures of cognitive control, as measured with electroencephalography, were remediated by multitasking training (enhanced midline frontal theta power and frontal–posterior theta coherence). Critically, this training resulted in performance benefits that extended to untrained cognitive control abilities (enhanced sustained attention and working memory), with an increase in midline frontal theta power predicting the training-induced boost in sustained attention and preservation of multitasking improvement 6 months later. These findings highlight the robust plasticity of the prefrontal cognitive control system in the ageing brain, and provide the first evidence, to our knowledge, of how a custom-designed video game can be used to assess cognitive abilities across the lifespan, evaluate underlying neural mechanisms, and serve as a powerful tool for cognitive enhancement.

In a first experiment, we evaluated multitasking performance across the adult lifespan. A total of 174 participants spanning six decades of life (ages 20–79; ∼30 individuals per decade) played a diagnostic version of NeuroRacer to measure their perceptual discrimination ability ('sign task') with and without a concurrent visuomotor tracking task ('driving task'; see Supplementary Information for details of NeuroRacer). Performance was evaluated using two distinct game conditions: 'sign only' (respond as rapidly as possible to the appearance of a sign only when a green circle was present); and 'sign and drive' (simultaneously perform the sign task while maintaining a car in the centre of a winding road using a joystick (that is, 'drive'; see Fig. 1a)). Perceptual discrimination performance was evaluated using the signal detection metric of discriminability (d'). A 'cost' index was used to assess multitasking performance by calculating the percentage change in d' from 'sign only' to 'sign and drive', such that greater cost (that is, a more negative percentage cost) indicates increased interference when simultaneously engaging in the two tasks (see Methods Summary).

Prior to the assessment of multitasking costs, an adaptive staircase algorithm was used to determine the difficulty levels of the game at which each participant performed the perceptual discrimination and visuomotor tracking tasks in isolation at ∼80% accuracy. These levels were then used to set the parameters of the component tasks in the multitasking condition, so that each individual played the game at a customized challenge level. This ensured that comparisons would inform differences in the ability to multitask, and not merely reflect disparities in component skills (see Methods, Supplementary Figs 1 and 2, and Supplementary Information for more details).

Multitasking performance diminished significantly across the adult lifespan in a linear fashion (that is, increasing cost, see Fig. 2a and Supplementary Table 1), with the only significant difference in cost between adjacent decades being the increase from the twenties (−26.7% cost) to the thirties (−38.6% cost). This deterioration in multitasking performance is consistent with the pattern of performance decline across the lifespan observed for fluid cognitive abilities, such as reasoning[5] and working memory[6]. Thus, using NeuroRacer as a performance assessment tool, we replicated previously evidenced age-related multitasking deficits[7,8], and revealed that multitasking performance declines linearly as we advance in age beyond our twenties.

In a second experiment, we explored whether older adults who trained by playing NeuroRacer in multitasking mode would exhibit improvements in their multitasking performance on the game[9,10] (that is, diminished NeuroRacer costs). Critically, we also assessed whether this training

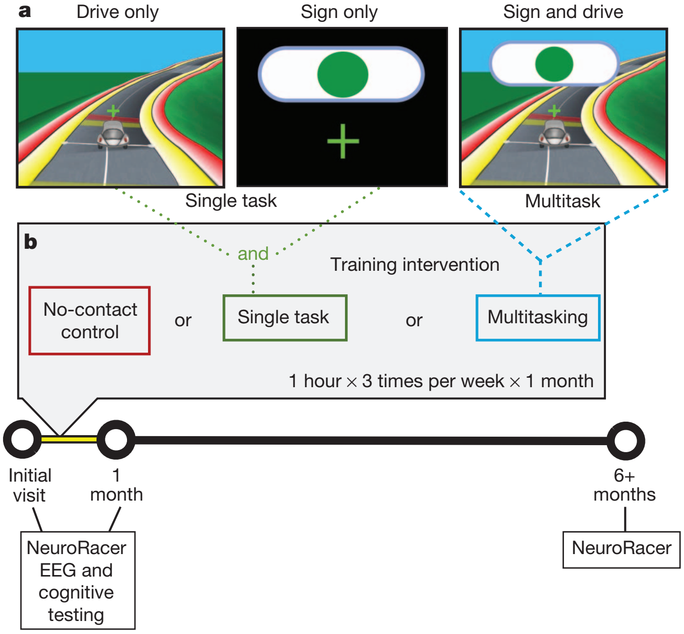

Figure 1 | NeuroRacer experimental conditions and training design. **a**, Screen shot captured during each experimental condition. **b**, Visualization of training design and measures collected at each time point.

---

[1] Department of Neurology, University of California, San Francisco, California 94158, USA. [2] Department of Physiology, University of California, San Francisco, California 94158, USA. [3] Center for Integrative Neuroscience, University of California, San Francisco, California 94158, USA. [4] Department of Psychiatry, University of California, San Francisco, California 94158, USA.

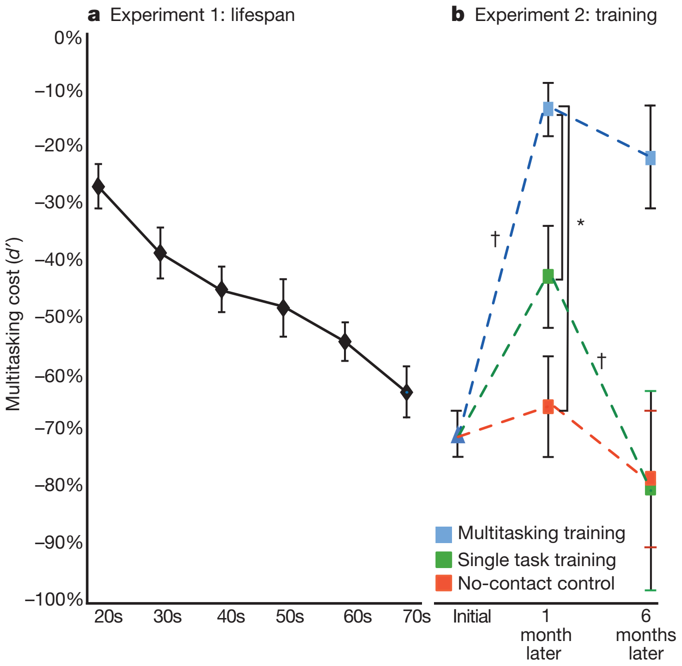

Figure 2 | NeuroRacer multitasking costs. **a**, Costs across the lifespan ($n = 174$) increased (that is, a more negative percentage) in a linear fashion when participants were grouped by decade ($F(1,5) = 135.7$, $P < 0.00001$) or analysed individually ($F(1,173) = 42.8$, $r = 0.45$, $P < 0.00001$; see Supplementary Fig. 3), with significant increases in cost observed for all age groups versus the 20-year-old group ($P < 0.05$ for each decade comparison). **b**, Costs before training, 1 month post-training, and 6 months post-training showed a session X group interaction ($F(4,72) = 7.17$, $P < 0.0001$, Cohen's $d = 1.10$), with follow-up analyses supporting a differential benefit for the MTT group (Cohen's d for MTT vs STT = 1.02; MTT vs NCC = 1.20). †$P < 0.05$ within group improvement from pre to post, *$P < 0.05$ between groups ($n = 46$). Error bars represent s.e.m.

transferred to enhancements in their cognitive control abilities[11] beyond those attained by participants who trained on the component tasks in isolation. In designing the multitasking training version of NeuroRacer, steps were taken to maintain both equivalent difficulty and engagement in the component tasks to assure a prolonged multitasking challenge throughout the training period: difficulty was maintained using an adaptive staircase algorithm to independently adjust the difficulty of the 'sign' and 'driving' tasks following each 3-min run based on task performance, and balanced task engagement was motivated by rewards given only when both component tasks improved beyond 80% on a given run.

We assessed the impact of training with NeuroRacer in a longitudinal experiment that involved randomly assigning 46 naive older adults (60–85 years old; 67.1 ± 4.2 (mean age ± s.d.)) to one of three groups: multitasking training (MTT; $n = 16$), single task training (STT; $n = 15$) as an active control, or no-contact control (NCC; $n = 15$). Training involved playing NeuroRacer on a laptop at home for 1 h a day, 3 times a week for 4 weeks (12 h of training in total), with all groups returning for a post-training assessment after 1 month, and a follow-up assessment after 6 months (Fig. 1b). The MTT group played the 'sign and drive' condition exclusively during the training period, whereas the STT group divided their time between a 'sign only' and a 'drive only' condition, and so were matched for all factors except the presence of interference. In addition to a battery of cognitive control tests used to assess the breadth of training benefits (see Supplementary Table 2), the neural basis of training effects was evaluated using electroencephalography (EEG) recorded at pre- and post-training visits while participants performed a neural assessment version of NeuroRacer.

Analysis showed that only the MTT group's multitasking performance index significantly improved from pre- (−64.2% cost) to post-training (−16.2% cost; Fig. 2b), thus supporting the role of interference during game play as a key mechanistic feature of the training approach. In addition, although cost reduction was observed only in the MTT group, equivalent improvement in component task skills was exhibited by both STT and MTT (see Supplementary Figs 4 and 5). This indicates that enhanced multitasking ability was not solely the result of enhanced component skills, but a function of learning to resolve interference generated by the two tasks when performed concurrently. Moreover, the d' cost improvement following training was not the result of a task trade-off, as driving performance costs also diminished for the MTT group from pre- to post-training (see Supplementary Information). Notably in the MTT group, the multitasking performance gains remained stable 6 months after training without booster sessions (at 6 months, −21.9% cost). Interestingly, the MTT group's post-training cost improved significantly beyond the cost level attained by a group of 20 year olds who played a single session of NeuroRacer (−36.7% cost; experiment 3; $P < 0.001$).

Next, we assessed if training with NeuroRacer led to generalized enhancements of cognitive control abilities that are known to be impaired in ageing (for example, sustained attention, divided attention, working memory; see Supplementary Table 2)[12]. We hypothesized that being immersed in a challenging, adaptive, high-interference environment for a prolonged period of time (that is, MTT) would drive enhanced cognitive performance on untrained tasks that also demanded cognitive control. Consistent with our hypothesis, significant group X session interactions and subsequent follow-up analyses evidenced pre- to post-training improvements in both working memory (delayed-recognition task with and without distraction[7]; Fig. 3a, b) and sustained attention

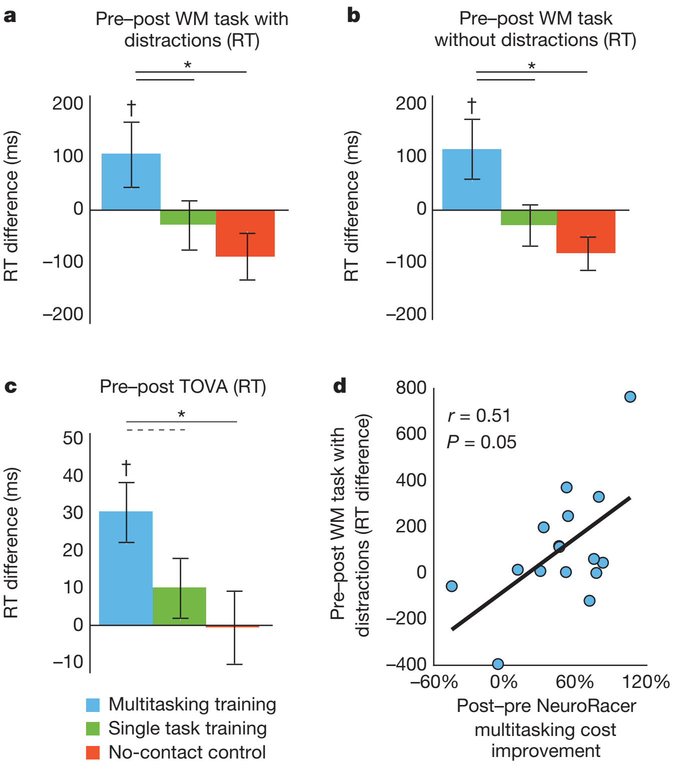

Figure 3 | Change in performance across sessions on independent tests of cognition for each experimental group. For each test, a group X session ANOVA revealed a significant interaction ($F(2,43) > 3.39$, $P < 0.04$, Cohen's $d > 0.73$), with follow-up analyses demonstrating improvement only for MTT ($n = 15$). **a**, Response time (RT) change for a delayed-recognition working memory (WM) task with the presence of distraction ($n = 46$). **b**, Response time change for a delayed-recognition WM task without distraction. **c**, Response time change for the test of variables of attention (TOVA). **d**, Correlation between data from (**a**) and NeuroRacer multitasking cost improvement 1 month after training for the MTT group ($n = 16$). †$P < 0.05$ within group improvement from pre to post, *$P < 0.05$ between groups, - - - $P = 0.08$. Error bars represent s.e.m.

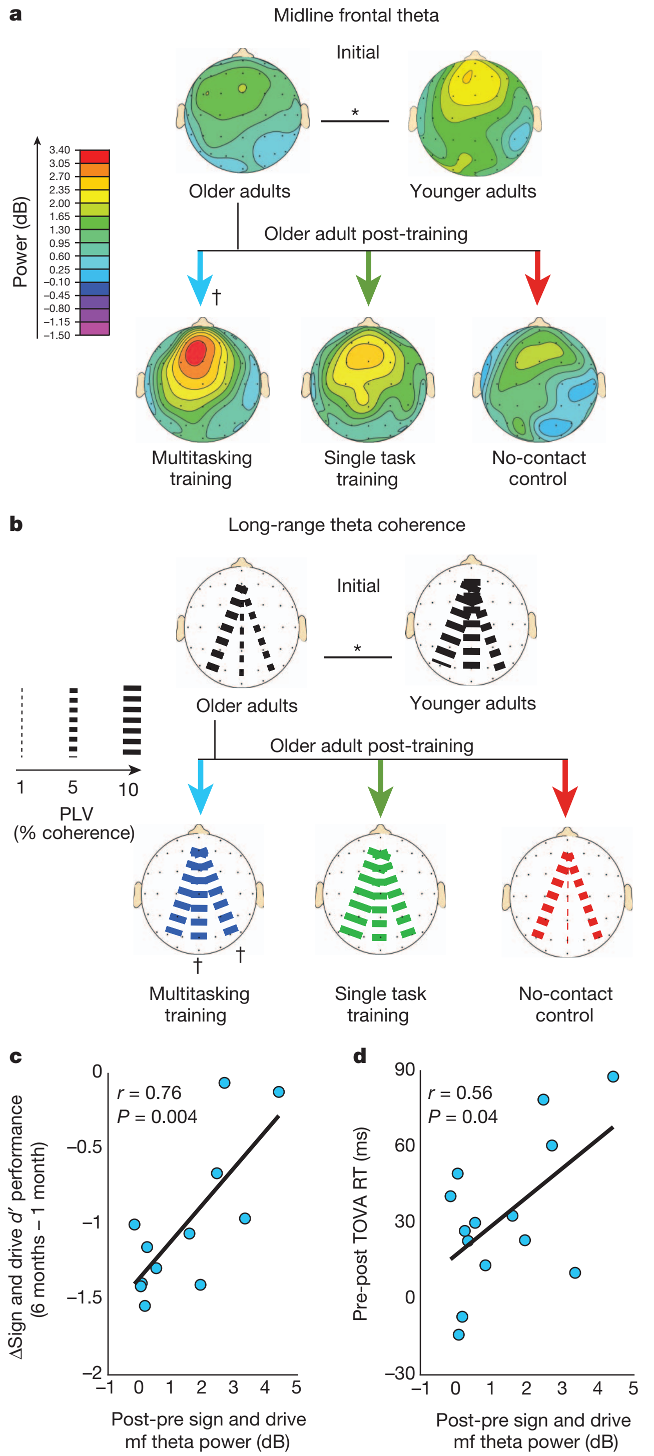

Figure 4 | 'Sign and drive' midline frontal theta activity and long-range theta coherence in younger adults and older adults pre- and post-training. **a**, **b**, For older adult training assessments, a group X session X condition ANOVA for each neural measure revealed significant interactions (in each case, $F(2,41) > 4.98$, $P < 0.01$, Cohen's $d > 0.93$; see Supplementary Fig. 6a, b), with follow-up analyses demonstrating improvement only for MTT during 'sign and drive' ($n = 15$). For younger ($n = 18$) vs older adult ($n = 44$) assessments, both neural measures revealed significant reductions in older adults (see Supplementary Fig. 8a, b). **c**, Correlation in the MTT group between the change in midline frontal theta power and multitasking behavioural gain preservation 6 months later ($n = 12$). **d**, Correlation in the MTT group between the change in midline frontal (mf) theta power and behavioural improvement on the TOVA ($n = 14$). †$P < 0.05$ within group improvement from pre- to post-training, *$P < 0.05$ between groups.

(vigilance; test of variables of attention (TOVA)[13]) only for the MTT group (Fig. 3c; see Supplementary Table 2). In addition, there were several statistical trends suggestive of improved post-training performance on other cognitive control tasks (dual-tasking, useful field of view, and change detection task; see analysis of covariance (ANCOVA) results in Supplementary Table 2). Note that although the working memory and sustained attention improvements were documented as more rapid responses to test probes, neither impulsivity (assessed with the alternative version of the TOVA) nor accuracy results showed significant group differences, revealing that training effects were not the result of a speed/accuracy trade-off. Importantly, these cognitive improvements were specific to working memory and sustained attention processes, and not the result of generalized increases in speed of processing, as no group X session interactions were found on two processing speed tasks (a stimulus detection task and the digit symbol substitution task; see Supplementary Table 2). Finally, only the MTT group exhibited a significant correlation between multitasking cost reduction (assessed with NeuroRacer) and improvements on an untrained cognitive control task (delayed-recognition with distraction) from pre- to post-training (Fig. 3d).

These important 'transfer of benefits' suggest that a common, underlying mechanism of cognitive control was challenged and enhanced by MTT with NeuroRacer. To assess this further, we examined the neural basis of training effects by quantifying event-related spectral perturbations (ERSP) and long-range phase coherence time-locked to the onset of each sign presented during NeuroRacer both pre- and post-training. We specifically assessed midline frontal theta (4–7 Hz), a well-described EEG measure of cognitive control (for example, working memory[14], sustained attention[15] and interference resolution[16]) localized to the medial prefrontal cortex. In addition, we analysed long-range theta coherence between frontal and posterior brain regions, a functional connectivity measure also associated with cognitive control (for example, working memory[14] and sustained attention[15]). Separate ANOVAs for theta power and coherence each revealed significant three-way interactions of condition ('sign and drive', 'sign only') X session (pre, post) X group (MTT, STT, NCC; see Supplementary Fig. 6). Further analysis revealed that for the 'sign and drive' condition, only the MTT group demonstrated a significant increase from pre- to post-training in both neural measures (see Fig. 4a, b). These findings are consistent with other reports of training-driven modulations in prefrontal cortical activity of older adults[9,17]. Furthermore, the coherence results demonstrate for the first time modulation of a neural network in response to cognitive training in older adults. These findings evidence a shift in the rapid engagement of prefrontal cognitive control processes less than 400 ms after a sign appears and before the motor response (see Supplementary Fig. 7 and Supplementary Table 1b), supporting training-induced neuroplasticity as the mechanistic basis of these training effects.

As described above, both MTT and STT resulted in equivalent improvements on the NeuroRacer component tasks (see Supplementary Table 2 and Supplementary Fig. 4), whereas only MTT led to broad enhancements both behaviourally (diminished multitasking costs, improved sustained attention and working memory) and neurally (enhanced midline frontal theta power and long-range coherence). This indicates that the training factor driving these effects was the interference generated when participants were motivated to engage in the two tasks simultaneously. Given that there were no clear differences in sustained attention or working memory demands between MTT and STT, transfer of benefits to these untrained tasks must have resulted from challenges to overlapping cognitive control processes. Of note, the use of a three-dimensional immersive and fun video game for training (see Methods) diverges from the sparse environments typically used in dual-task training studies[9,10], which have not documented a similar degree of far transfer[10].

Coupled with previous findings of increased midline frontal theta on a variety of cognitive control tasks[18], the current results support a common neural basis of cognitive control processes, which can be enhanced by immersion in an adaptive, high-interference environment. This interpretation is bolstered by evidence here indicating that MTT-induced increases in midline frontal theta power during 'sign and drive' were positively correlated with both sustained multitasking performance improvements (6 months post performance, Fig. 4c) and improvements in TOVA response times (Fig. 4d). Thus, MTT-induced enhancement of midline frontal theta power was associated with the preservation of multitasking performance over time and with generalized benefits on an untrained cognitive control task, reflecting its utility as a neural signature of plastic cognitive control processes.

Finally, we questioned whether these neural measures that exhibited training effects in older adults were actually altered at baseline compared to younger adults, or if training boosted non-deficient neural processes. In a third experiment, we compared midline frontal theta power and long-range coherence from older adults before training to a naive group of younger adults who were not trained ($n = 18$; 20–29 years old (24.1 ± 2.9)). The multitasking costs for each group replicated findings of age-matched cohorts from experiment 1. Both neural measures showed a main effect of group (see Supplementary Fig. 8), indicating less theta power and coherence in older adults when processing signs in either condition ('sign and drive' depicted in Fig. 4a, b). The absence of a significant condition X age group interaction for either neural measure (see Supplementary Fig. 8) revealed that ageing was associated with a general reduction in theta power and coherence when older adults discriminate visual stimuli, regardless of whether they are multitasking or single tasking. Notably, MTT led to changes in the neural processing of signs during 'sign and drive' that reached a level comparable to neural activity patterns observed in younger adults.

The mechanism underlying these neural findings are informed by a growing literature that shows deactivation of medial prefrontal cortical activity (suppression of a node of the 'default network'[19]) during cognitively demanding tasks is associated with reduced susceptibility to internal distraction and better task performance[20]. Given that medial prefrontal activity is inversely correlated with midline frontal theta power[21], increased levels of midline frontal theta exhibited by older adults following MTT may reflect more deactivation of medial prefrontal activity. NeuroRacer training may benefit cognitive control abilities by improving the ability of older adults to suppress the default network during task engagement, a process known to be compromised in ageing[22]. Future studies using neurochemical and physiological manipulations are warranted to inform the causal nature of the relationship between medial prefrontal activity and training-induced performance effects observed here.

This study offers neural and behavioural evidence of generalized positive effects from video game training on cognitive control abilities of older adults, with enhancements comparable to those observed in younger adults who are habitual action video-game players: interference resolution[23], working memory[24] and sustained attention[25]. Although reports of transfer of benefits following cognitive training in the older population are relatively rare[11,26], the observed generalization supports the results of larger-scale training studies that demonstrate some degree of transfer to untrained cognitive tasks[27,28] and subjective measures of daily living[29]. In contrast to these studies, and most other cognitive training experiments on older adults that report small to medium effect sizes for untrained tasks, the current findings document medium to large effect sizes (all > 0.50–1.0 (using Cohen's d, see Methods)) for both cognitive control performance and neural measures versus either control group. The sustained multitasking cost reduction over time and evidence of generalizability to untrained cognitive control abilities provide optimism for the use of an adaptive, interference-rich, video game approach as a therapeutic tool for the diverse populations that suffer from cognitive control deficits (for example, attention deficit hyperactivity disorder (ADHD), depression, dementia). These findings stress the importance of a targeted training approach, as reinforced by a recent study that observed a distinct lack of transfer following non-specific online cognitive exercises[30]. In conclusion, we provide evidence of how a custom-designed video game targeting impaired neural processes in a population can be used to diagnosis deficits, assess underlying neural mechanisms and enhance cognitive abilities.

METHODS SUMMARY

All participants had normal or corrected vision, no history of neurological, psychiatric or vascular disease, and were not taking any psychotropic or hypertension medications. In addition, they were considered 'non-gamers' given that they played less than 2 h of any type of video game per month. For NeuroRacer, each participant used their left thumb for tracking and their right index finger for responding to signs on a gamepad controller (Logitech). Participants engaged in three 3-min runs of each condition in a randomized fashion. Signs were randomly presented in the same position over the fixation cross for 400 ms every 2, 2.5 or 3 s, with the speed of driving dissociated from sign presentation parameters. The multitasking cost index was calculated as follows: (('sign and drive' performance − 'sign only' performance)/'sign only' performance) × 100. EEG data for 1 MTT post-training participant and 1 STT pre-training participant were corrupted during acquisition. Two MTT participants, two STT participants and four NCC participants were unable to return to complete their six-month follow-up assessments. Critically, no between-group differences were observed for neuropsychological assessments ($P = 0.52$) or pre-training data involving: (1) NeuroRacer thresholding for both road ($P = 0.57$) and sign ($P = 0.43$); (2) NeuroRacer component task performance ($P > 0.10$ for each task); (3) NeuroRacer multitasking costs ($P = 0.63$); (4) any of the cognitive tests (all ANOVAs at pre-training: $P \geq 0.26$); (5) ERSP power for either condition ($P \geq 0.12$); and (6) coherence for either condition ($P \geq 0.54$).

Full Methods and any associated references are available in the online version of the paper.

Received 16 January; accepted 18 July 2013.

1. Botvinick, M. M., Braver, T. S., Barch, D. M., Carter, C. S. & Cohen, J. D. Conflict monitoring and cognitive control. *Psychol. Rev.* **108,** 624–652 (2001).
2. Dux, P. E. *et al.* Training improves multitasking performance by increasing the speed of information processing in human prefrontal cortex. *Neuron* **63,** 127–138 (2009).
3. Foehr, U. G. Media multitasking among American youth: prevalence, predictors, and pairings. (Kaiser Family Foundation, 2006).
4. Gazzaley, A. in *Principles of Frontal Lobe Function* 2nd edn (eds D. T. Stuss & R. T. Knight) Top-down modulation and cognitive aging. (Oxford Univ. Press, 2013).
5. Tucker-Drob, E. M. & Salthouse, T. A. Adult age trends in the relations among cognitive abilities. *Psychol. Aging* **23,** 453–460 (2008).
6. Park, D. C. *et al.* Models of visuospatial and verbal memory across the adult life span. *Psychol. Aging* **17,** 299–320 (2002).
7. Clapp, W. C., Rubens, M. T., Sabharwal, J. & Gazzaley, A. Deficit in switching between functional brain networks underlies the impact of multitasking on working memory in older adults. *Proc. Natl Acad. Sci. USA* **108,** 7212–7217 (2011).
8. Verhaeghen, P., Steitz, D. W., Sliwinski, M. J. & Cerella, J. Aging and dual-task performance: a meta-analysis. *Psychol. Aging* **18,** 443–460 (2003).
9. Erickson, K. I. *et al.* Training-induced plasticity in older adults: effects of training on hemispheric asymmetry. *Neurobiol. Aging* **28,** 272–283 (2007).
10. Lussier, M., Gagnon, C. & Bherer, L. An investigation of response and stimulus modality transfer effects after dual-task training in younger and older. *Front. Hum. Neurosci.* **6,** 129 (2012).
11. Zelinski, E. M. Far transfer in cognitive training of older adults. *Restor. Neurol. Neurosci.* **27,** 455–471 (2009).
12. Gazzaley, A., Cooney, J. W., Rissman, J. & D'Esposito, M. Top-down suppression deficit underlies working memory impairment in normal aging. *Nature Neurosci.* **8,** 1298–1300 (2005).
13. Greenberg, L. M. *T.O.V.A. Continuous Performance Test Manual.* (The TOVA company, 1996).
14. Onton, J., Delorme, A. & Makeig, S. Frontal midline EEG dynamics during working memory. *Neuroimage* **27,** 341–356 (2005).
15. Sauseng, P., Hoppe, J., Klimesch, W., Gerloff, C. & Hummel, F. C. Dissociation of sustained attention from central executive functions: local activity and interregional connectivity in the theta range. *Eur. J. Neurosci.* **25,** 587–593 (2007).
16. Nigbur, R., Ivanova, G. & Sturmer, B. Theta power as a marker for cognitive interference. *Clin. Neurophysiol.* **122,** 2185–2194 (2011).
17. Dahlin, E., Nyberg, L., Backman, L. & Neely, A. S. Plasticity of executive functioning in young and older adults: immediate training gains, transfer, and long-term maintenance. *Psychol. Aging* **23,** 720–730 (2008).
18. Mitchell, D. J., McNaughton, N., Flanagan, D. & Kirk, I. J. Frontal-midline theta from the perspective of hippocampal 'theta'. *Prog. Neurobiol.* **86,** 156–185 (2008).
19. Buckner, R. L., Andrews-Hanna, J. R. & Schacter, D. L. The brain's default network: anatomy, function, and relevance to disease. *Ann. NY Acad. Sci.* **1124,** 1–38 (2008).
20. Grady, C. L., Springer, M. V., Hongwanishkul, D., McIntosh, A. R. & Winocur, G. Age-related changes in brain activity across the adult lifespan. *J. Cogn. Neurosci.* **18,** 227–241 (2006).
21. Scheeringa, R. *et al.* Frontal theta EEG activity correlates negatively with the default mode network in resting state. *Int. J. Psychophysiol.* **67,** 242–251 (2008).

22. Damoiseaux, J. S. *et al.* Reduced resting-state brain activity in the 'default network' in normal aging. *Cereb. Cortex* **18,** 1856–1864 (2008).
23. Strobach, T., Frensch, P. A. & Schubert, T. Video game practice optimizes executive control skills in dual-task and task switching situations. *Acta Psychol. (Amst.)* **140,** 13–24 (2012).
24. Boot, W. R., Kramer, A. F., Simons, D. J., Fabiani, M. & Gratton, G. The effects of video game playing on attention, memory, and executive control. *Acta Psychol. (Amst.)* **129,** 387–398 (2008).
25. Dye, M. W., Green, C. S. & Bavelier, D. Increasing speed of processing with action video games. *Curr. Dir. Psychol. Sci.* **18,** 321–326 (2009).
26. Berry, A. S. *et al.* The influence of perceptual training on working memory in older adults. *PLoS ONE* **5,** e11537 (2010).
27. Smith, G. E. *et al.* A cognitive training program based on principles of brain plasticity: results from the improvement in memory with plasticity-based adaptive cognitive training (IMPACT) study. *J. Am. Geriatr. Soc.* **57,** 594–603 (2009).
28. Wolinsky, F. D., Vander Weg, M. W., Howren, M. B., Jones, M. & Dotson, M. M. A randomized controlled trial of cognitive training using a visual speed of processing intervention in middle aged and older adults. *PLoS ONE* **8,** e61624 (2013).
29. Ball, K. *et al.* Effects of cognitive training interventions with older adults: a randomized controlled trial. *J. Am. Med. Assoc.* **288,** 2271–2281 (2002).
30. Owen, A. M. *et al.* Putting brain training to the test. *Nature* **465,** 775–778 (2010).

Supplementary Information is available in the online version of the paper.

Acknowledgements We thank J. Avila, N. Barbahiya, M. Gugel, B. Jensen, R. Moustafa, Y. Rezaeihaghighi, P. Sztybel, C. Vong, A. Wang, B. Yang and D. Yerukhimov for their help with data collection and analyses, and B. Benson for assistance with the NeuroRacer behavioral analysis stream. Thanks to D. Ellingson, N. Falstein, and M. Omernick for insights and support of NeuroRacer development. Thanks to J. Bollinger, J. Kalkstein, J. Mishra, B. Voytek and T. Zanto for support on ERSP and coherence analyses, and Z. Chadick, W. Clapp, J. Fung, M. Hough, E. Morsella, J. Pa, M. Rubens, P. Wais, C. Walsh, and D. Ziegler for helpful discussions. Thanks to all of our participants whose time and efforts made this work possible, and Apple who generously loaned the Gazzaley laboratory all of the MacBook Pro laptops used in this study. Support for this research was provided by the Robert Wood Johnson Foundation's Pioneer Portfolio through a grant from its national program, 'Health Games Research: Advancing Effectiveness of Interactive Games for Health' (A.G.) and the National Institute of Aging (A.G.). J.A.A. was supported by a UCSF Institutional Research and Career Development Award (IRACDA).

Author Contributions J.A.A., J.B., J.L.R., O.A., E.J. and A.G. designed the experiments; J.A.A., J.L.R., O.A., E.J. and A.G. developed the NeuroRacer software; J.A.A., J.B., O.A., F.F., E.K., Y.L. and C.R. collected the data; J.A.A., J.B., O.A., J.J. and C.R. analysed the data; and J.A.A. and A.G. wrote the paper. All authors discussed the results.

Author Information Reprints and permissions information is available at www.nature.com/reprints. The authors declare competing financial interests: details are available in the online version of the paper. Readers are welcome to comment on the online version of the paper. Correspondence and requests for materials should be addressed to J.A.A. (joaquin.anguera@ucsf.edu) or A.G. (adam.gazzaley@ucsf.edu).

METHODS

Participants. All participants were recruited through online and newspaper advertisements. For experiment 1, 185 (90 male) healthy, right-handed individuals consented to participate according to procedures approved by the University of California at San Francisco. For experiments 2 and 3, 60 (33 males) older adult individuals and 18 (9 male) young adult individuals participated without having been a part of experiment 1 (see Supplementary Table 3 for demographic descriptions and Supplementary Fig. 9 for experiment 2 participant enrollment). Participants who were unable to perform the tasks, as indicated by tracking performance below 15% (6 individuals from experiment 1, 8 individuals from experiment 2), or a false positive rate greater than 70% (5 individuals from experiment 1, 6 individuals from experiment 2) during any one visit or across more than 4 individual training sessions, were excluded.

Thresholding. Prior to engaging in NeuroRacer, participants underwent an adaptive thresholding procedure for discrimination (nine 120-s runs) and tracking ability (twelve 60-s runs) to determine a 'sign' and 'drive' level that each participant would perform at ~80% accuracy (see Supplementary Figs 1 and 2). Having individuals engage each condition in their own 'space' following thresholding procedures facilitated a fairer comparison across ages and abilities. This is a frequently omitted procedure in other studies, and leads to difficulty interpreting performance differences (especially multitasking) as being the result of differences in interference processing or due to differences in component task skills.

For the perceptual discrimination thresholding, each participant's performance for a given run was determined by calculating a proportion correct score involving: (1) correctly responding to targets; (2) correctly avoiding non-targets; (3) late responses to targets; and (4) responding to non-targets. At the end of each run, if this score was greater than 82.5%, the subsequent run would be played at a higher level which had a corresponding shorter time window for responses to targets. More specifically, the adaptive algorithm would make proportional level changes depending upon participants' performance from this ~80% median, such that each 1.75% increment away from this median corresponded with a change in level (see Supplementary Fig. 1a). Thus, a 90% performance would lead to a 40-ms reduction in the time window, while a 55% (or less) performance would lead to a 100-ms lengthening of said window. Thresholding parameters for road levels followed a similar pattern with each 0.58% increment away from the same median corresponded with a change in level (see Supplementary Fig. 1b). These parameters were chosen following extensive pilot testing to: (1) minimize the number of trial runs until convergence was reached and (2) minimize convergence instability, while (3) maximizing sampling resolution of user performance.

The first three driving thresholding blocks were considered practice to familiarize participants with the driving portion of the task and were not analysed. A regression over the nine thresholding runs in each a case was computed to select the ideal time window and road speed to promote a level of ~80% accuracy on each distraction free task throughout the experiment (see Supplementary Fig. 2). All participants began the thresholding procedures at the same road (level 20) and sign levels (level 29).

Conditions. Following the driving and sign thresholding procedures, participants performed five different three minute 'missions', with each mission performed three times in a pseudo-randomized fashion. In addition to the 'sign only', 'drive only', and 'sign and drive' conditions, participants also performed a 'sign with road' condition in which the car was placed on 'auto pilot' for the duration of the run and participants responded to the signs, and a 'drive with signs' condition where participants were told to ignore the presence of signs appearing that and continue to drive as accurately as possible. Data from these two conditions are not presented here. Feedback was given at the end of each run as the proportion correct to all signs presented for the perceptual discrimination task (although we used the signal detection metric of discriminability (d')[31] to calculate our 'cost' index throughout the study), and percentage of time spent on the road (see Supplementary Fig. 10). Prior to the start of the subsequent run, participants were informed as to which condition would be engaged in next, and made aware of how many experimental runs were remaining. Including thresholding, the testing session encompassed 75 min of game play.

NeuroRacer training and testing protocol. For experiment 1, participants were seated in a quiet room in front of an Apple MacBook Pro 5.3 laptop computer at an approximate distance of 65 cm from the 15 inch screen. For experiment 2 and 3, participants were seated in a dark room with the screen ~100 cm from the participants. All training participants trained at their homes using an Apple MacBook Pro 5.3 laptop computer while sitting ~60 cm from the screen (see Supplementary Fig. 11a). For experiment 1, each perceptual discrimination-based experimental run (180 s) contained 36 relevant targets (green circles) and 36 lures (green, blue and red pentagons and squares). For experiments 2 and 3, the sign ratio was 24/48.

Prior to training, each participant was given a tutorial demonstrating how to turn on the laptop, properly set-up the joystick, navigate to the experiment, shown what the first day of training would be like in terms of the task, how to interpret what the feedback provided meant, and were encouraged to find a quiet environment in their home for their training sessions. If requested by the participant, a laboratory member would visit the participant at their home to help set up the computer and instruct training. In addition, to encourage/assess compliance and hold participants to a reasonable schedule, participants were asked to plan their training days and times with the experimenter for the entire training period and enter this information into a shared calendar. Each participant (regardless of group) was informed that their training protocol was designed to train cognitive control faculties, using the same dialogue to avoid expectancy differences between groups. There was no contact between participants of different groups, and they were encouraged to avoid discussing their training protocol with others to avoid potentially biasing participants in the other groups.

Each day of training, the participants were shown a visualization of a map that represented their 'training journey' to provide a sense of accomplishment following each training session (Supplementary Fig. 11b). They were also shown a brief video that reminded them how to hold the controller, which buttons to use, their previous level(s) reached, and what the target would be that day for the perceptual discrimination condition. In addition, the laptop's built-in video camera was also activated (indicated by a green light) for the duration of said run, providing (1) visual assessment of task engagement; (2) motivation for participants to be compliant with the training task instructions; and (3) information about any run where performance was dramatically poorer than others.

Participants were discouraged from playing 2 days in a row and they were encouraged to play at the same time of day. MTT participants were reminded that an optimal training experience was dependent upon doing well on both their sign and drive performance without sacrificing performance on one task for the other. Although the STT group were provided a 'driving' or 'sign' score following each training run, the MTT group were also provided an 'overall' score following each run as a composite of performance on both tasks (see Supplementary Figs 5 and 11). Following the completion of every fourth run, participants were rewarded with a 'fun fact' screen regarding basic human physiology (http://faculty.washington.edu/chudler/ffacts.html) before beginning their subsequent training run. To assess if training was a 'fun' experience, participants in each training group rated the training experience on their final visit to the laboratory on a scale of 1 (minimally) to 10 (maximally) (MTT: 6.5 ± 2.2; STT 6.9 ± 2.4; $t = 0.65$, $P = 0.52$). Critically, training groups did not differ on their initial thresholding values for both road ($F(2,45) = 0.58$, $P = 0.57$) and sign ($F(2,45) = 0.87$, $P = 0.43$).

Each laptop was configured to transmit NeuroRacer performance data to our secure laboratory server wirelessly using Dropbox as each run was completed. This facilitated monitoring for compliance and data integrity in a relatively real-time fashion, as participants would be contacted if (1) there was a failure to complete all 20 training runs on a scheduled training day; (2) 'sign only' and 'drive only' performance was suggestive that a problem had occurred within a given training session; or (3) a designated training day was missed. Individuals without wireless internet in their home were instructed to visit an open wireless internet location (for example, a coffee shop or public library) at least once a week to transfer data, and if this was not an option, researchers arranged for weekly home visits to acquire said data. All participants were contacted through email and/or phone calls on a weekly basis to encourage and discuss their training; similarly, in the event of any questions regarding the training procedures, participants were able to contact the research staff through phone and email.

Pre- and post-training evaluations involving cognitive testing and NeuroRacer EEG took place across 3 different days (appointment and individual test order were counterbalanced), with all sessions completed approximately within the span of a week (total number of days to complete all pre-training testing: 6.5 ± 2.2 days; post-training testing: 6.1 ± 1.5 days). Participants returned for their first post-training cognitive assessments 2.0 ± 2.2 days following their final training session. Although scheduled for 6 months after their final testing session, the follow-up visits at 6 months actually occurred on average 7.6 ± 1.1 months afterwards owing to difficulties in maintaining (and rescheduling) these distant appointments. Critically, no group differences were present regarding any of these time-of-testing measures ($F < 1.81$, $P > 0.18$ for each comparison).

Cognitive battery. The cognitive battery (see Supplementary Table 2) consisted of tasks spanning different cognitive control domains: sustained attention (TOVA; see Supplementary Fig. 12a), working memory (delayed-recognition- see Supplementary Fig. 12b), visual working memory capacity (see Supplementary Fig. 13), dual-tasking (see Supplementary Fig. 14), useful field of view (UFOV; see Supplementary Fig. 15), and two control tasks of basic motor and speed of processing (stimulus detection task, digit symbol substitution task; see Supplementary Fig. 16). Using the analysis metrics regularly reported for each measure, we performed a mixed model ANOVA of group (3: MTT, STT, NCC) X session (2: pre, post) X cognitive test (11; see Supplementary Table 2), and observed a significant three-way interaction ($F(20, 400) = 2.12$, $P = 0.004$) indicative that training had selective benefits across group and test. To interrogate this interaction, each cognitive test was analysed separately with session X group ANOVAs to isolate those measures that changed significantly following training. We also present the *P* value associated with the ANCOVAs for each measure in Supplementary Table 2 (dependent measure = post-training performance, covariate = pre-training performance), which showed a similar pattern of effects as most of the two-way ANOVAs. The ANCOVA approach is considered to be a more suitable approach when post-test performance that is not conditional/predictable based on pre-test performance is the primary outcome of interest following treatment, as opposed to characterizing gains achieved from pre-training performance (for example, group X session interaction(s))[32]; however, both are appropriate statistical tools that have been used to assess cognitive training outcomes[27,33] (see Supplementary Fig. 17 as an example).

EEG recordings and eye movements. Neural data were recorded using an active two head cap (Cortech Solutions) with a BioSemiActiveTwo 64-channel EEG acquisition system in conjunction with BioSemiActiView software (Cortech Solutions). Signals were amplified and digitized at 1024 Hz with a 16-bit resolution. Anti-aliasing filters were used and data were band-pass filtered between 0.01–100 Hz during data acquisition.

For each EEG recording session, the NeuroRacer code was modified to flash a 1 × 1 inch white box for 10 ms at one of the corners on the stimulus presentation monitor upon the appearance of a sign. A photodiode (http://www.gtec.at/Products/Hardware-and-Accessories/g.TRIGbox-Specs-Features) captured this change in luminance to facilitate precise time-locking of the neural activity associated with each sign event. During the experiment, these corners were covered with tape to prevent participants from being distracted by the flashing light.

To ensure that any training effects were not due to changes in eye movement, electro-oculographic data were analysed. Using a previously validated approach[34], vertical (VEOG = FP_2-IEOG electrodes) and horizontal (HEOG = REOG-LEOG electrodes) difference waves were calculated from the raw data and baseline corrected to the mean pre-stimulus activity. The magnitude of eye movement was computed as follows: $(\mathrm{VEOG}^2 + \mathrm{HEOG}^2)^{0.5}$. The variance in the magnitude of eye movement was computed across trials and the mean variance was specifically examined from −200 to 1000 ms post-stimulus onset. The variance was compared: (1) between sessions for each group's performance on the 'sign and drive' and 'sign only' conditions; (2) between groups at each session for each condition; and (3) between young and older adults on each condition. We used two-tailed *t*-test that were uncorrected for multiple comparisons at every ms time point to be as conservative as possible. There was no session difference for any group on the 'sign only' condition ($P > 0.05$ for each group comparison); similarly, there were no differences for the MTT or NCC groups on the 'sign and drive' condition ($P > 0.30$ for each comparison), with the STT group showing more variance following training ($P = 0.01$). With respect to experiment 3, there were also no age differences on either condition ($P > 0.45$ for each comparison). This indicates that the training effects observed were not due to learned eye movements, and that the age effects observed were also not a function of age-related differences in eye movements as well.

EEG analysis. Preprocessing was conducted using Analyzer software (Brain Vision) then exported to EEGLAB[35] for event-related spectral perturbations (ERSP) analyses. ERSP is a powerful approach to identifying stable features in a spontaneous EEG spectrum that are induced by experimental events, and have been used to successfully isolate markers of cognitive control[36,37]. We selected this approach because we felt that a measure in the frequency domain would be more stable than other metrics given the dynamic environment of NeuroRacer. Blinks and eye-movement artefacts were removed through an independent components analysis, as were epochs with excessive peak-to-peak deflections (± 100 μV). Given the use of d', which takes into account performance on every trial, we collapsed across all trial types for all subsequent analyses. Epochs of −1000 to +1000 ms were created for ERSP total power analysis (evoked power + induced power), with theta band activity analysed by resolving 4–100 Hz activity using a complex Morlet wavelet in EEGLAB and referenced to a −900 to −700 pre-stimulus baseline (thus relative power (dB)).

Assessment of the 'sign and drive' ERSP data in 40 ms time bins collapsing across all older adult participants and experimental sessions revealed the onset of peak midline frontal activity to be between 360–400 ms post-stimulus, and so all neural findings were evaluated within this time window for the older adults (see Supplementary Fig. 7 for these topographies). For younger adults, peak theta activity occurred between 280–320 ms, and so for across-group comparisons, data from this time window was used for younger adults.

The cognitive ageing literature has demonstrated delayed neural processing in older adults using EEG[38,39]. For example, previous work has demonstrated that older adults show similar patterns of activity as younger adults in terms of selective processing, but there is a time shift to delayed processing with ageing[38]. For the data generated in this study, presented topographically in Supplementary Fig. 7, it was clear that the peak of the midline frontal theta was delayed in older versus younger adults. To fairly assess if there was a difference in power, it was necessary to select different comparison windows in an unbiased, data-driven manner for each group.

Coherence data for each channel was first filtered in multiple pass bands using a two-way, zero phase-lag, finite impulse response filter (eegfilt.m function in EEGLAB toolbox) to prevent phase distortion. We then applied a Hilbert transform to each of these time series (hilbert.m function), yielding results equivalent to sliding window FFT and wavelet approaches[40], giving a complex time series:

$$h_x[n] = a_x[n] \exp(i\phi_x[n]) \tag{1}$$

in which $a_x[n]$ and $\phi_x[n]$ are the instantaneous amplitudes and phases, respectively. The phase time series ϕ_x assumes values within $(-\pi, \pi$ radians with a cosine phase such that π radians corresponds to the trough and 0 radians to the peak. In order to compute PLV for theta phase, for example, we extract instantaneous theta phases $\phi_\theta[n]$ by taking the angle of $h_\theta[n]$. Event-related phase time-series are then extracted and, for each time point, the mean vector length $R_\theta[n]$ is calculated across trials (circ_r.m function in CircStats toolbox)[41]. This mean vector length represents the degree of PLV in which an R of 1 reflects perfect phase-locking across trials and a value of 0 reflects perfectly randomly distributed phases. These PLVs were controlled for individual state differences at each session by baseline correcting each individual's PLVs using their −200 to 0 period (thus, a relative PLV score was calculated for each subject).

Statistical analyses. Mixed model ANOVAs with: (1) decade of life (experiment 1); (2) training group (experiment 2); or (3) age (experiment 3) as the between-group factor were used for all behavioural and neural comparisons, with planned follow-up *t*-tests and the Greenhouse-Geisser correction used where appropriate. One-tailed *t*-tests were used to interrogate group differences for all transfer measures given our a priori hypothesis of the direction of results following multitask training. All effect size values were calculated using Cohen's d[42] and corrected for small sample bias using the Hedges and Olkin[43] approach. The neural-behavioural correlations presented included only those MTT participants who demonstrated increased midline frontal theta power following training (14/15 participants). For statistical analyses, we created 1 frontal and 3 posterior composite electrodes of interest (EOI) from the average of the following electrodes: AFz, Fz, FPz, AF3, and AF4 (medial frontal), PO8, P8, and P10 (right-posterior), PO7, P7, and P9 (left-posterior); POz, Oz, O1, O2 and Iz (central-posterior), with PLVs calculated for each frontal-posterior EOI combination separately. For the coherence data, the factor of posterior EOI location (3) was modelled in the ANOVA, but did not show either a main effect or interaction with the other factors.

31. Macmillan, N. A. & Creelman, C. D. *Detection Theory: A User's Guide*. 2nd edn (Lawrence Erlbaum Associates, 2005).
32. Knapp, T. R. & Schafer, W. D. From Gain Score t to ANCOVA F (and vice versa). *Pract. Assess. Res. Eval.* **14,** 6 (2009).
33. Jaeggi, S. M., Buschkuehl, M., Jonides, J. & Perrig, W. J. Improving fluid intelligence with training on working memory. *Proc. Natl Acad. Sci. USA* **105,** 6829–6833 (2008).
34. Berry, A. S., Zanto, T. P., Rutman, A. M., Clapp, W. C. & Gazzaley, A. Practice-related improvement in working memory is modulated by changes in processing external interference. *J. Neurophysiol.* **102,** 1779–1789 (2009).
35. Delorme, A. & Makeig, S. EEGLAB: an open source toolbox for analysis of single-trial EEG dynamics including independent component analysis. *J. Neurosci. Methods* **134,** 9–21 (2004).
36. Makeig, S. Auditory event-related dynamics of the EEG spectrum and effects of exposure to tones. *Electroencephalogr. Clin. Neurophysiol.* **86,** 283–293 (1993).
37. Neuper, C. & Klimesch, W. *Event-Related Dynamics of Brain Oscillations*. Vol. 159 (Elsevier Science, 2006).
38. Zanto, T. P., Toy, B. & Gazzaley, A. Delays in neural processing during working memory encoding in normal aging. *Neuropsychologia* **48,** 13–25 (2010).
39. Gazzaley, A. *et al.* Age-related top-down suppression deficit in the early stages of cortical visual memory processing. *Proc. Natl Acad. Sci. USA* **105,** 13122–13126 (2008).
40. Bruns, A. Fourier-, Hilbert- and wavelet-based signal analysis: are they really different approaches? *J. Neurosci. Methods* **137,** 321–332 (2004).
41. Berens, P. CircStat: A MATLAB Toolbox for Circular Statistics. *J. Stat. Softw.* **31,** 1–21 (2009).
42. Cohen, J. *Statistical Power Analysis for the Behavioral Sciences*. 2nd edn (Lawrence Erlbaum Associates, 1988).
43. Hedges, L. V. & Olkin, I. *Statistical Methods For Meta-analysis*. (Academic Press, 1985).