



Flegal, K., Ragland, J. D. and Ranganath, C. (2019) Adaptive task difficulty influences neural plasticity and transfer of training. *NeuroImage*, 188, pp. 111-121.

There may be differences between this version and the published version. You are advised to consult the publisher's version if you wish to cite from it.

<http://eprints.gla.ac.uk/176640/>

Deposited on: 21 January 2019

Enlighten – Research publications by members of the University of Glasgow_
<http://eprints.gla.ac.uk>

Adaptive task difficulty influences neural plasticity and transfer of training

Kristin E. Flegal^{1,2*}, J. Daniel Ragland³, and Charan Ranganath^{2,4}

¹ Institute of Neuroscience and Psychology, University of Glasgow,
58 Hillhead Street, Glasgow, G12 8QB, Scotland, UK.

² Center for Neuroscience, University of California at Davis,
1544 Newton Court, Davis, CA, 95618, USA.

³ Department of Psychiatry and Behavioral Sciences, University of California at Davis,
4701 X Street, Sacramento, CA, 95817, USA.

⁴ Department of Psychology, University of California at Davis,
One Shields Avenue, Davis, CA, 95616, USA.

*Correspondence to: kristin.flegal@glasgow.ac.uk

Abstract

The efficacy of cognitive training is controversial, and research progress in the field requires an understanding of factors that promote transfer of training gains and their relationship to changes in brain activity. One such factor may be adaptive task difficulty, as adaptivity is predicted to facilitate more efficient processing by creating a prolonged mismatch between the supply of, and the demand upon, neural resources. To test this hypothesis, we measured behavioral and neural plasticity in fMRI sessions before and after 10 sessions of working memory updating (WMU) training, in which the difficulty of practiced tasks either adaptively increased in response to performance or was fixed. Adaptive training resulted in transfer to an untrained episodic memory task and activation decreases in striatum and hippocampus on a trained WMU task, and the amount of training task improvement was associated with near transfer to other WMU tasks and with hippocampal activation changes on both near and far transfer tasks. These findings suggest that cognitive training programs should incorporate adaptive task difficulty to broaden transfer of training gains and maximize efficiency of task-related brain activity.

Keywords: COGNITIVE TRAINING, TRANSFER, fMRI, PLASTICITY, MEMORY, EXECUTIVE FUNCTION

Introduction

Training cognitive processes such as memory and executive function can improve behavioral performance and drive changes on neural measures (Klingberg, 2010; Morrison and Chein, 2011; Hsu et al., 2014). However, a common criticism is that effects are often limited to the trained tasks, whereas transfer to untrained tasks is inconsistent. Some studies show “near transfer” within the same cognitive domain as trained tasks, but evidence of “far transfer”, or generalization across cognitive domains, is reported less frequently—and regarded more skeptically (Moody, 2009; Shipstead et al., 2012; Melby-Lervåg et al., 2016). Although the literature on training-induced plasticity has stimulated a great deal of interest in developing interventions to improve cognition (Ranganath et al., 2011; Vinogradov et al., 2012; Mishra and Gazzaley, 2014), a lack of understanding of the factors that mediate transfer effects has hindered translation of laboratory research into demonstrably effective programs. Given the wide variability in methodology across training studies, breadth of transfer may depend on how training is conducted. Here, we investigated the possibility that effective transfer depends, at least in part, on adapting the difficulty of training tasks to an individual’s current level of proficiency (i.e., adaptive training). That is, do successful cognitive training outcomes require an intervention that dynamically increases task demands? Previous studies have speculated that adaptivity may be a key to effective transfer (Holmes et al., 2009; Jaeggi et al., 2010b; Brehmer et al., 2012; Anguera et al., 2013), but systematic investigations are lacking. We sought to address this controversy by directly testing whether adaptive, relative to individualized but non-adaptive, difficulty mediates behavioral and neural effects of cognitive training.

According to a recent theoretical framework (Lövdén et al., 2010), effective transfer depends on how cognitive processes are trained—whereas transient cognitive challenges are only

sufficient to promote task-specific learning, sustained cognitive challenges are required to elicit lasting neural changes that underlie enhancement of a general cognitive function. Specifically, if environmental demand (e.g., the processing load of a working memory task) briefly approaches the upper limit of functional supply (e.g., working memory processing efficiency), then all available resources will be flexibly brought to bear, but actually raising the level of maximum function (e.g., improved processing efficiency) requires a prolonged mismatch in which environmental demand exceeds functional supply. Based on this model, we predicted that adaptively increasing training task difficulty would provide the necessary prolonged mismatch, thereby inducing plasticity that is associated with broader transfer and greater changes in task-related brain activity than non-adaptive training. If adaptive training successfully improves processing efficiency, then training gains should generalize beyond superficially similar tasks to untrained tasks that rely on the same processing components (Jonides, 2004; Dahlin et al., 2008b; 2009), resulting in far transfer. Additionally, improved processing efficiency should be reflected in decreased neural recruitment in task-related brain areas (Kelly and Garavan, 2005). A few studies have used functional magnetic resonance imaging (fMRI) to assess the neural effects of training and transfer by scanning untrained tasks as well as trained (criterion) tasks at pre- and post-training sessions (Dahlin et al., 2008b; Schneiders et al., 2012; Schweizer et al., 2013; Heinzl et al., 2016), establishing that training-induced plasticity generalizes across tasks that engage overlapping brain areas, but adaptivity-related effects on fMRI outcome measures have never been studied.

In the present study, we manipulated adaptivity in a training regimen targeting the core cognitive process of working memory updating (WMU), an executive function that controls updating of information that is active in working memory (Morris and Jones, 1990; Miyake et

al., 2000). We assessed behavioral change and neural plasticity in fMRI sessions before and after 10 sessions of computerized training with visuospatial and verbal WMU tasks. Participants were randomly assigned to either an adaptive training (AT) group or a non-adaptive (NA) active control group. Training procedures were identical across the two groups, except for one critical difference. For AT participants, as task performance increased the number of updating operations (i.e., update level) was consequently increased, in order to adaptively and selectively increase the environmental demands on WMU processes. For NA participants, however, task difficulty was fixed at a relatively low level across all training days. Update level was individually set for each NA participant, in an effort to equate subjective difficulty across the active control group, as any single level of objective difficulty could produce higher or lower environmental demands based on participants' pre-existing ability differences.

Pre- and post-training MRI scanning sessions evaluated functional brain activity during a WMU criterion task modified from the visuospatial training task, an untrained spatial n-back task, and an untrained object-location association episodic memory task (Fig. 1). On the basis of previous studies reporting that WMU and episodic memory processing components of interest involve activation of striatum and hippocampus—subcortical structures long understood to contribute to learning and memory processes (Packard & Knowlton, 2002; Squire, 2004)—fMRI analyses in the present study focused on these structures as *a priori* regions of interest (ROIs). In particular, fMRI studies of WMU training have identified striatum as a specific site of training-induced activity changes (Dahlin et al., 2008b; Kühn et al., 2012). Additionally, we examined hippocampus because it is known to show increased activation during the Object-Location Association task (Gould, 2005; de Rover et al., 2011), and it has also been implicated in visuospatial working memory tasks (Piekema et al., 2006; Hannula & Ranganath, 2008; see Nee

& Jonides, 2013 for review) similar to the scanned WMU tasks in the present study. Thus, guided by our hypothesis that adaptively increasing training task difficulty will broaden transfer of training gains across tasks that engage overlapping processing components and brain areas, we selected striatal and hippocampal ROIs to examine adaptivity-related effects of WMU training and to test the prediction that improved WMU processes would facilitate episodic memory encoding to induce far transfer of training gains.

Materials and Methods

Participants

63 healthy young adults (18-29 years old; $M = 20.8$, $SD = 2.4$) were recruited from the University of California at Davis (UCD) community. Participants were right-handed, native English speakers, with normal or correct-to-normal vision, no reported history of neurological or psychiatric illness, no current use of psychoactive medication, and no known MRI contraindications. 48 females and 15 males participated. The research protocol was approved by the UCD Institutional Review Board, and all participants provided written informed consent and were paid for their participation. Compensation was \$10 for each of nine behavioral-only sessions, \$20 for each of three sessions with MRI scanning, plus a \$50 bonus for completing all 12 of the study sessions.

In the initial enrollment phase, 26 participants were assigned to the adaptive training (AT) group and 19 participants were assigned to the non-adaptive (NA) active control group. Assignment was random and single-blind, with the restriction that the groups did not run simultaneously (due to the delivery of at-home training sessions that differed by group assignment), so recruitment occurred in blocks alternating between the two groups. In a later

enrollment phase, 18 additional participants were recruited into a no-contact control (NCC) group.

Two participants (both from the AT group) withdrew prior to study completion; one due to claustrophobia at the first scanning session and one due to personal reasons after completing five study sessions. Five other participants assigned to the AT group failed to meet inclusion criteria due to a lack of improvement within the training protocol itself, defined by a measure of training gain (linear slope calculated from the maximum level of performance achieved at each training session) that was negative for one or both of the training tasks. Notably, negative training slopes indicate that these participants' training task performance declined to, and never recovered from, a floor level of difficulty even lower than the fixed levels performed by non-adaptive active control participants. Because the adaptivity manipulation in this study is operationalized by increasing task difficulty in response to performance improvements, and this defining feature was not experienced by AT participants who failed to improve on the trained tasks, their data were excluded from the present analysis (they are to be reported in a separate paper investigating predictors of responsiveness to training). Thus, the final sample included in the results reported below consists of 19 AT participants, 19 NA participants, and 18 NCC participants. Mean age and gender ratio were equated across conditions ($F_s < 1$).

Materials

Training Tasks

The training protocol consisted of two tasks designed to target working memory updating (WMU) processes, using different modalities in order to discourage task-specific strategies and to promote transfer. Example trials from both tasks are depicted in Fig. 2, below graphs of their

respective training trajectories for AT participants. The training tasks were administered, and responses were collected, using Presentation software (Version 14.9, www.neurobs.com).

Matrix Updating (MU) is a visuospatial working memory task that requires updating the location of multiple dots within a 4x4 matrix (Chen and Li, 2007). On MU trials, a matrix with colored dots (red, orange, green, and blue) in four of its cells was first displayed for 5000 ms, then in the center of the empty matrix, colored arrows (pointing up, down, left, or right) were presented sequentially for 1750 ms each with a 250 ms interstimulus interval. Participants were instructed to follow each arrow by mentally moving the dot of the same color one cell in that direction. After a variable number of arrows, a colored pointer appeared in the center of the empty matrix, prompting the participant to respond by using the mouse to move the pointer and click on the current location of the dot of the same color. The MU task was divided into blocks of five trials each, with feedback (number of correct and incorrect responses) presented at the end of each block. Within each task block, stimuli (location of dots; color and direction of arrows) were randomized on a trial-by-trial basis with the constraint that each arrow must point its corresponding dot in a valid direction: always within the matrix boundaries and never into a cell currently occupied by another dot. MU task duration was approximately 25-30 min.

Keep Track (KT) is a verbal working memory task that requires updating the identity of the most recently studied words in multiple semantic categories (Yntema, 1963). On KT trials, the names of four categories were displayed in boxes at the bottom of the screen, while in the center of the screen, exemplar words from the categories were presented sequentially for 2000 ms each with a 1000 ms interstimulus interval. Participants were instructed to mentally place each presented word into the box for its corresponding category. After a variable number of words, the box belonging to one of the four categories was highlighted, prompting the participant

to respond by using the keyboard to type the last word that was placed into that box. Four novel categories (and their respective word lists) were used at each of the 10 training sessions. In order to create a total of 40 categorized word lists of sufficient length, stimuli were collected from multiple published word pools (Murdock, 1976; Howard, 1979; van Overschelde et al., 2004). At each training session, the KT task began with a screen listing all of the words in the lists to be used in that session, in order to familiarize participants with the correct category assignments. Within each task block, stimuli were randomized with the constraint that all four categories were sampled (in any order) before any category was sampled again. In addition, trials contained occasional “distractor” words that did not belong to any of the given categories, which participants were instructed to ignore. KT task duration was approximately 20-25 min.

For both training tasks, level of difficulty can be modulated by increasing or decreasing the update level, i.e., the number of updates on each trial. At each update level, to minimize the predictability of when in a trial the response would be required, the exact number of updates was randomly selected from the update level ± 1 . For example, at the 7-update level of the MU task, the number of arrows on a given trial could be 6, 7, or 8. Importantly, in both training tasks, the working memory load was constant (always four colored dots or four categories) while the adaptivity manipulation was achieved solely by varying the update level, allowing the training protocol to specifically target WMU processes.

[FIG. 1 ABOUT HERE]

Scanned Tasks

Matrix Updating was modified from the training task version to an event-related fMRI design, serving as a *criterion task* performed at all study sessions (see also Dahlin et al., 2008b).

The structure and timing of the criterion task trials were the same as the training task version, except the response phase was changed to yes/no recognition: instead of freely moving a pointer to identify the updated location of a particular dot, one of the four colored dots reappeared in the matrix after the updating phase and the task was to respond by pressing one button if it was the correct, current location of that dot and pressing a second button if it was not (see Fig. 1A). Additionally, the Matrix Updating criterion task was not adaptive but instead consisted of three trial types: 7-Updates, with a high updating demand of seven colored arrows presented during the delay period; 4-Updates, with a lower updating demand of four colored arrows; and 0-Updates, a maintenance-only baseline condition in which four gray arrows are presented and thus the recognition probe after the delay period simply referred to the original location of the colored dots on that trial. The task was divided into four runs of 11 trials each, for a total of 16 trials in each of the two active updating conditions and 12 trials in the baseline condition. For each trial type, the dependent variable was the proportion of correct trials. Trial order was unique across runs and optimized using optseq2 (Dale, 1999), with the intertrial interval varying between 2 and 10 s ($M = 4$ s). Total duration of the four runs was approximately 20 min.

Spatial N-Back was selected as a scanned task representing near transfer, based on the prediction that it and the WMU training tasks engage overlapping processing components and brain areas. Based on an n-back paradigm used by Jaeggi and colleagues (2010a), stimuli were blue squares that appeared in one of eight locations (the perimeter of an unseen 3x3 matrix) for 500 ms each with a 2500 ms interstimulus interval, and the task was to respond by pressing one button when the current location matched the location presented n trials earlier and pressing a second button when there was not a match (see Fig. 1B). Each block consisted of 12 trials, of which three were targets. The N-Back (NB) task was divided into two runs of nine blocks each,

in a counterbalanced order alternating among three trial types determined by the value of n : 3-Back, which presents a high updating demand; 2-Back, which presents a lower updating demand; and 0-Back, a baseline condition in which the target location was always the upper left corner of the screen. For each trial type, the dependent variable was overall accuracy. Total duration of the two NB runs was approximately 13 min.

Object-Location Association is a measure of visual episodic memory, selected as a scanned task representing far transfer. Based on a paired associate learning paradigm adapted for fMRI testing (Gould, 2005; de Rover et al., 2011), the task consisted of blocks of trials arranged into an encoding phase followed by a retrieval phase (see Fig. 1C). Stimuli were unique kaleidoscope images (“objects”) from Voss and colleagues (2008) that were presented sequentially for 3 s each at random locations within a 4x4 matrix during the encoding phase, and participants were instructed to remember which object appeared in which cell, for the subsequent retrieval phase (separated from the last encoding trial by a 4 s delay). On each retrieval trial, one of the cells in which an object had appeared was highlighted for 5 s, and the task was to make a button press response to select the object associated with that location from among three options displayed at the bottom of the screen (one target and two foils that also appeared during the encoding phase). Every object-location pair presented during an encoding phase was probed during the subsequent retrieval phase. The Object-Location (OL) task consisted of two trial types: 8-Associates, with a high memory load of eight object-location pairs (i.e., eight encoding trials followed by eight retrieval trials); and 6-Associates, with a lower memory load of six pairs. There were also baseline task blocks, to control for perceptual and motor processing in the absence of memory load. In the control “encoding” phase of the baseline condition, six gray squares were presented sequentially within the matrix and participants were instructed to rest

with their eyes open but not try to remember anything about the squares. In the control “retrieval” phase, another six gray squares were presented sequentially and participants were instructed to make a button press response for each square to report the row of the matrix in which it appeared. The OL task was divided into two runs of six blocks each, in a counterbalanced order alternating among the two active trial types and the baseline condition. For each trial type, the dependent variable was the proportion of correct retrieval trials. Total duration of the two OL runs was approximately 14 min.

Order of the three scanned tasks was counterbalanced across participants, but task order was held constant across the scanning sessions for each participant.

Other Transfer Tasks

To more broadly assess transfer and test for nonspecific effects of the training protocols, a battery of untrained tasks was administered outside of the scanner after both pre- and post-training sessions. The executive functions of Updating, Inhibition, and Shifting (Miyake et al., 2000) were measured, respectively, with a Letter Running Memory task (Pollack et al., 1959; Morris and Jones, 1990), a Counting Stroop task (Bush et al., 1998), and a Global/Local task (Navon, 1977). Working memory capacity was measured for verbal stimuli with the Automated Operation Span task (Unsworth et al., 2005), and for visual stimuli with a change localization (Gold et al., 2006) version of the Change Detection task (Luck and Vogel, 1997). Verbal episodic memory was measured with the Hopkins Verbal Learning Test–Revised (Benedict et al., 1998), fluid intelligence with Raven’s Advanced Progressive Matrices (Raven et al., 1998), sustained attention with the Paced Auditory Serial Addition Test (Gronwall, 1977; Fischer et al., 1999), and processing speed with the WAIS-III Digit-Symbol Substitution test (Wechsler, 1997).

Additionally, to measure individual differences in implicit beliefs about the malleability of intelligence (see also Jaeggi et al., 2014), the 3-item Theories of Intelligence Scale (Dweck and Henderson, 1988) was administered before the task battery at the first study visit only.

Alternate versions of the standard neuropsychological measures were used for pre- and post-training assessments, with order of the two versions counterbalanced across participants. For the computerized tasks, validated alternate versions were not available, but stimuli were randomized at each assessment to minimize practice effects. Comparing AT, NA, and NCC groups, and controlling for pre-training performance, there was no significant effect of group on post-training scores for any tasks in the battery.

Design and Procedure

The study consisted of a total of 12 visits scheduled over the course of three weeks. The first and last study visits involved MRI scanning sessions and other transfer tasks administered outside of the scanner, and the remaining 10 visits were training sessions. In addition, an early-training MRI scanning session was included in the third study visit; data from the scanned tasks at that session are to be reported in a separate paper. Four study visits were scheduled per week, and, across participants, pre- and post-training scanning sessions (i.e., the first and last study visits) were separated by 16-18 calendar days. The scanning sessions were conducted at the UCD MRI Facility for Integrative Neurosciences, using a 3T Siemens Skyra scanner (imaging parameters are detailed below). Any participant who had no previous experience in an MRI environment completed a brief mock scanning session prior to the first study visit. Practice trials were provided before each task, and Presentation (www.neurobs.com) and E-Prime (Psychology

Software Tools, Pittsburgh, PA) software was used to collect behavioral data from the computerized tasks.

The first training session (Visit 2) was completed in the laboratory under experimenter supervision, as was the second training session because it coincided with the early-training scanning session which required a laboratory visit. Thereafter, the remaining eight training sessions (Visits 4-11) were completed on participants' home computers according to the study schedule. Task performance was monitored from encrypted data files transmitted to study staff via e-mail at the end of each at-home training session. To monitor compliance remotely, a secure website logged each time the training program was run, and participants were contacted promptly by an experimenter if a scheduled session was missed. Every participant who completed the study performed all 10 training sessions.

Participants assigned to the AT group started their first training session at the 4-update level for MU and the 5-update level for KT. For each subsequent session, each task was started at the level determined by the final block of the previous session. Adaptive difficulty was implemented in both tasks with an algorithm that applied an 80% accuracy criterion after every five trials. If at least four of the preceding trials were answered correctly, the update level was increased by one for the next five trials. Otherwise, the update level was decreased by one for the next five trials (down to a minimum 3-update level for MU and 4-update level for KT). Across training sessions, as the number of updates progressively increased with increasing levels of difficulty, the algorithm reduced the total number of task blocks in order to preserve a relatively constant duration for each training session (e.g., it takes approximately the same amount of time to complete eight blocks of MU trials at an average of the 10-update level as it does six blocks of MU trials at an average of the 16-update level).

Participants assigned to the NA group started their first training session—and remained for that and all subsequent sessions—at an individualized level of difficulty between the 5- and 9-update level for MU and between the 6- and 8-update level for KT. Participants were assigned to levels approximating the number of updates they would be predicted to achieve by the end of a first training session under adaptive conditions, on the basis of a pre-training measure of working memory capacity (Operation Span), calculated using a regression equation derived from pilot data. Consequently, although all NA participants completed the training tasks at a fixed and relatively low level of difficulty, the cognitive demand was deliberately set not so low as to induce boredom and disengagement (which has been a complaint rightly levied against less-active non-adaptive control conditions in previous training studies; e.g., as discussed in Morrison and Chein, 2011). Furthermore, for NA participants as well as AT participants, the exact number of updates was unpredictable on each trial because it was randomly selected from the update level ± 1 .

The NCC group was included to assess practice effects in the transfer task behavioral data. Participants in this group performed the same battery of scanned tasks and other transfer tasks, in three sessions scheduled at the same intervals (pre-, early-, and post-training) as participants in the active updating conditions, but were not scanned and completed no training sessions between their study visits. To control for environmental influences on performance, NCC participants completed the three “scanned” tasks in a mock scanner, using the same display and response collection equipment as at the UCD MRI Facility, while listening to an EPI pulse sequence recording through headphones during the task runs.

MRI Acquisition and Processing

At each scanning session, a multi-band gradient-echo EPI sequence (repetition time [TR] = 1220 ms; echo time [TE] = 24 ms; multi-band factor = 2; flip angle = 67°; field of view [FOV] = 192 mm; 64 × 64 matrix; 38 slices; 3.0 mm isotropic voxels) was used to obtain functional images sensitive to BOLD contrast. In each functional run, the first four volumes were discarded to allow for signal equilibration. The total number of volumes collected was 248 in each Matrix Updating criterion task run, 320 in each Spatial N-Back run, and 331 in each Object-Location Association run. An MP-RAGE sequence (TR = 1800 ms; TE = 2.96 ms; flip angle = 7°; FOV = 256 mm; 256 × 256 matrix) was used to obtain high-resolution T1-weighted anatomical images at the end of each scanning session.

For each participant, anatomical images acquired at each session were averaged using the `mri_robust_template` program (Reuter et al., 2012) and the average image was used as an unbiased template for spatial coregistration across sessions. Data were preprocessed using SPM8 (www.fil.ion.ucl.ac.uk/spm). Each participant's functional images were realigned using a six-parameter rigid body transformation, coregistered to their average anatomical image, normalized to MNI (Montreal Neurological Institute) space using affine and nonlinear transformations, and spatially smoothed using a 6-mm isotropic FWHM Gaussian kernel. Each participant's average anatomical image was segmented, and the conjunction of their normalized, spatially smoothed gray and white matter images was used to calculate a brain-only explicit mask used in functional analyses.

Analysis

For each task, at each scanning session, BOLD responses were analyzed using the general linear model implemented in SPM8. Covariates of interest were constructed by

convolving vectors of predicted neural activity with a canonical hemodynamic response function. To account for residual variance because of head movement, motion parameters estimated at the realignment stage of preprocessing and motion spikes identified using the ArtRepair toolbox (cibsr.stanford.edu/tools/human-brain-project/artrepair-software.html) were included in each model as covariates of no interest.

The Matrix Updating criterion task was analyzed in an event-related design, with separate regressors modeling matrix, updating, and probe period activation as a function of trial type (7-Updates/4-Updates/0-Updates) and response accuracy (correct/incorrect). First-level analysis was performed using the general linear model and applying a high-pass filter with a 200-sec cutoff period. The primary contrast of interest for high vs. low demand on WMU processes evaluated probe-period activation on correct 7-update trials vs. correct 4-update trials.

The other scanned tasks were analyzed in block designs, with first-level analysis performed using the general linear model and applying a high-pass filter with a 128-sec cutoff period. Spatial N-Back had separate regressors for trial type (3-Back/2-Back/0-Back), and the primary contrast of interest evaluated 2- and 3-Back blocks vs. 0-Back blocks. Object-Location Association had separate regressors modeling encoding and retrieval phase activation as a function of trial type (8-Associates/6-Associates/baseline), and the primary contrast of interest evaluated encoding-phase activation in 6- and 8-Associate blocks vs. baseline blocks.

To examine task- and adaptivity-related effects in brain areas associated with the putative processing components—WMU and episodic memory—involved in the scanned tasks, *a priori* ROIs were defined by computing the intersections between bilateral caudate, putamen, and hippocampus anatomical ROIs from the LONI Probabilistic Brain Atlas (Shattuck et al., 2008)

and each participant's normalized, spatially smoothed gray matter image from their segmented average anatomical image. For the primary contrast of interest from each task, for each participant at each scanning session, mean parameter estimates were extracted from the mask images of each ROI. Thereafter, for each ROI analysis, post-training activation was entered as the dependent variable in an ANCOVA with group (AT/NA) as a fixed factor and pre-training activation as a covariate (reported in Table 2). An exploratory whole-brain analysis subsequently investigated adaptivity-related activation changes not restricted to *a priori* ROIs, in a mixed design ANOVA with group (AT/NA) as the between-subjects factor and session (pre-training/post-training) as a repeated measure (reported in Table 3).

To analyze behavioral data from the scanned tasks, post-training performance for each measure was entered as the dependent variable in an ANCOVA with group (AT/NA/NCC) as a fixed factor and pre-training performance as a covariate (reported in Table 1). For AT participants, training gains were analyzed using repeated-measures ANOVA on the maximum update level achieved in each training session, for each task. Amount of improvement on the trained tasks (i.e., Training Slope) was indexed by averaging the linear slopes calculated from the maximum update level achieved in each training session, for each task.

[FIG. 2 ABOUT HERE]

Results

As expected, AT participants showed significant improvements over the course of training – as indicated by a significant effect of training day on the maximum update level achieved in each session – for both training tasks ($F_s > 36.98$; $ps < .001$). Training trajectories are shown in Fig. 2. On average, by the last training day, AT participants were performing

visuospatial trials at the 24-update level and verbal trials at the 22-update level. While such gains in WMU performance from the first training day are notable, transfer effects are of greater interest. That is, did training task improvements transfer to untrained tasks? Data in pre- and post-training fMRI sessions were obtained from a WMU criterion task (Matrix Updating), which was the visuospatial training task modified for scanning, an untrained WMU task (Spatial N-Back) to assess near transfer, and an untrained episodic memory task for which improved WMU processes may support more effective encoding (Object-Location Association) to assess far transfer. Each of the three scanned tasks included high-difficulty, low-difficulty, and baseline trial types. Because the plasticity induced through sustained neurocognitive challenge is proposed to raise the level of maximum function (Lövdén et al., 2010), we predicted that the largest performance increases and changes in brain activity related to adaptive training would be found on high-difficulty trials. Behavioral data from the AT and NA groups were compared to a no-contact control (NCC) group that completed the same criterion and transfer tasks (without fMRI) but with no intervening WMU training (see Table 1). For post-training performance, controlling for pre-training performance, there was a significant effect of group on high-difficulty Matrix Updating trials ($F(2,52) = 4.50, p < .05, \eta_p^2 = .15$), and high-difficulty Object-Location Association trials ($F(2,52) = 3.75, p < .05, \eta_p^2 = .13$). As shown in Fig. 3, performance increases were largest for AT participants, those predicted to benefit from a prolonged mismatch between functional supply and environmental demand. The Spatial N-Back task showed no significant effect of group on any trial type ($F_s < 1$), reflecting near-ceiling performance¹.

[TABLE 1 ABOUT HERE]

¹ The same null result is found if the proportion of hits minus false alarms (Pr) is used as the dependent variable for Spatial N-Back instead of overall accuracy.

[FIG. 3 ABOUT HERE]

Our next analyses investigated individual differences in responsiveness to training and transfer to untrained tasks. Because progressively higher levels of environmental demand are proposed to induce proportionally larger increases in functional supply (Lövdén et al., 2010), we predicted that greater amounts of adaptive training task improvement would be associated with larger transfer effects. To index relative training gains among AT participants, linear slopes were calculated for each training task from the maximum level of difficulty achieved in each session, and averaged to create a Training Slope variable. Controlling for pre-training performance, partial correlations showed that Training Slope was significantly predictive of post-training performance for high-difficulty Spatial N-Back trials ($r_p = .50, p < .05$), with a marginal effect for high-difficulty Matrix Updating trials ($r_p = .42, p < .10$). For both tasks, greater post-training performance was associated with greater improvement on the trained tasks.

Having established that adaptive WMU training increased transfer to untrained tasks, we next analyzed fMRI data in order to determine the neural mechanisms of these behavioral effects. Region of interest (ROI) analyses were performed for brain areas associated with the putative processing components—WMU and episodic memory—involved in the scanned tasks. For each task, at each scanning session, a primary contrast of interest was computed for high vs. low demand on WMU processes, and mean parameter estimates were extracted from anatomically-defined *a priori* ROIs in bilateral striatum (caudate and putamen) and bilateral hippocampus. For the Matrix Updating criterion task, controlling for pre-training activation, there was a significant effect of group on post-training activation in all ROIs (see Table 2). As shown in Fig. 4, activation decreases were greater for AT than NA participants. To verify the selectivity of this adaptivity-related change in brain activity, we also investigated activation

changes in the bilateral occipital pole, which was not predicted to be sensitive to the adaptive training manipulation. No significant effect of group was found in this control region ($F < 1$). Within the AT group, partial correlations showed that Training Slope significantly predicted post-training activation, controlling for pre-training activation, in bilateral hippocampus ROIs for both the Spatial N-Back task and the Object-Location Association task ($|r_p|s > .49$; $ps < .05$). For both tasks, in both hippocampus ROIs, greater post-training deactivation was associated with greater improvement on the trained tasks.

[TABLE 2 ABOUT HERE]

[FIG. 4 ABOUT HERE]

As working memory and episodic memory tasks can be expected to recruit brain regions in addition to striatum and hippocampus (Wager & Smith, 2003; Spaniol et al., 2009; Ranganath & Ritchey, 2012; Nee et al., 2013), we supplemented the ROI analyses with an exploratory whole-brain analysis to identify all significant regions of adaptivity-related activation changes in the primary contrast of interest from each scanned task, using a cluster-corrected FWE threshold of $p < 0.05$. In the Matrix Updating criterion task, as shown in Fig. 5, the group by session interaction revealed greater activation decreases for AT than NA participants in bilateral striatum, consistent with ROI analysis, and also bilateral prefrontal, bilateral temporal, and left parietal regions. This analysis did not identify significant hippocampal clusters. The equivalent whole-brain group by session interaction analyses performed for the Spatial N-Back task and the Object-Location Association task yielded no suprathreshold clusters in either case (see Table 3 for the results summary of exploratory whole-brain interaction effects).

[FIG. 5 ABOUT HERE]

[TABLE 3 ABOUT HERE]

Discussion

The goal of the present study was to determine whether cognitive training outcomes depend on how processes are trained. Our results demonstrate that adaptive task difficulty is one key factor that can influence breadth of transfer and efficiency of brain activity. Adaptive WMU training resulted in transfer to an episodic memory task supported by WMU processes, and activation decreases in striatum and hippocampus ROIs on the scanned WMU criterion task. Notably, the detected transfer effects were reliably larger in the adaptive training group than in a closely matched non-adaptive active control group that performed the same training tasks for the same number of sessions. Furthermore, the amount of adaptive training task improvement was associated with near transfer to other WMU tasks, and with hippocampal activation changes on untrained tasks measuring both near and far transfer. Additionally, adaptivity-related transfer effects appeared at high levels of task difficulty. These findings are consistent with the proposal that sustained neurocognitive challenge is a mediator of behavioral and neural plasticity (Lövdén et al., 2010).

By demonstrating an important role for adaptive task difficulty in inducing plasticity, our fMRI data reveal novel information about the neural effects of adaptive training, with adaptivity-related activation decreases implicating increased neural efficiency (Kelly and Garavan, 2005; Lövdén et al., 2010) during task performance. The post-training activation decreases observed in the present study contribute to a literature in which training-induced changes in fMRI outcome measures are inconsistent, with activation increases, functional reorganization, and more complex dynamics of brain activity changes also found over the course of cognitive training

(Klingberg, 2010; Morrison and Chein, 2011; Hsu et al., 2014). Specifically regarding WMU training, previous studies have reported striatal activation increases after 15 training sessions (Dahlin et al., 2008b) and striatal activation increases after about 5 training sessions followed by decreases after more than 50 training sessions (Kühn et al., 2012), indicating that the temporal dynamics of changes in brain activity induced by WMU training can be nonmonotonic. For evaluating these past findings in light of the present study, an important difference is that both previous studies used training tasks in which the level of difficulty increased according to a predetermined schedule, or was capped within a restricted range, and thus was not continuously adaptive. In the present study, training task difficulty was individually adapted within sessions in response to performance in the AT group, or individually assigned on the basis of pre-training working memory capacity in the NA group, so that the group comparison would isolate the effects of improved processing efficiency realized through continuously adaptive WMU training.

A recent study in older adults examined fMRI data before and after an n-back training program which was continuously adaptive, and found training-related activation decreases in lateral prefrontal cortex on an n-back task and also on an untrained WMU task (Heinzel et al., 2016). No striatal activation changes were reported, however a number of methodological differences between this study and ours limit comparability of fMRI results, including the age group studied, the tasks scanned, the use of a no-contact rather than active control group for comparison, and the use of a whole brain voxelwise analysis approach rather than a focus on *a priori* ROIs. Nevertheless, the findings of Heinzel and colleagues (2016) associating adaptive training with decreased recruitment of task-related brain areas, and with transfer to untrained tasks, are consistent with our primary results. Moreover, our own exploratory whole-brain group by session interaction analysis also detected activation decreases after adaptive training in

prefrontal regions which have been previously associated with WMU processes (Wager & Smith, 2003; Nee et al., 2013).

Although an exploratory whole-brain approach identified adaptivity-related activation changes on the scanned WMU criterion task in prefrontal as well as temporal, parietal, and striatal regions (as shown in Fig. 5), *a priori* ROIs for our analyses specifically focused on striatum—where previous fMRI studies of WMU training have reported activation changes—and hippocampus—which is associated with episodic memory tasks and also visuospatial working memory tasks—in order to test the prediction that adaptive difficulty would broaden transfer of training gains across tasks that engage overlapping processing components and brain areas. As shown in Fig. 4, greater criterion task activation decreases in these subcortical areas were observed after 10 sessions of adaptive, relative to individualized but non-adaptive, WMU training. Furthermore, greater amounts of adaptive training task improvement were associated with greater activation decreases in bilateral hippocampus on untrained visuospatial working memory and episodic memory tasks. Evidence that interactions between striatal and hippocampal regions support episodic memory (Sadeh et al., 2011; Nyberg et al., 2016) along with computational models of working memory incorporating striatal and hippocampal connectivity (Hazy et al., 2006) offer a potential neural mechanism for our findings that adaptive WMU training resulted in transfer to episodic memory task performance and activation changes in caudate, putamen, and hippocampus ROIs.

The present study stands out from much of the cognitive training literature by showing what is conventionally accepted as far transfer, with training-related improvements in working memory generalizing to an untrained episodic memory task. Although there are some previous reports of transfer to episodic memory from working memory training (Rudebeck et al., 2012)

and from multi-domain training (Schmiedek et al., 2010; Toril et al., 2016), many studies have failed to find far transfer effects across cognitive domains, including other training regimens specifically targeting WMU (Dahlin et al., 2008a). As with interpreting differences in brain activity changes between the present study and previous fMRI studies of WMU training, methodological variations may partly account for why our training protocol was associated with far transfer while others were not. Continuously adaptive training task difficulty with no upper limit is a feature which our study has in common with a previous study that showed transfer to episodic memory was predicted by amount of improvement on an adaptive spatial working memory task (Rudebeck et al., 2012), and which is different from a previous WMU training study where all participants achieved the highest available level of training task difficulty and minimal evidence was found for far transfer (Dahlin et al., 2008a). The theoretical framework of Lövdén and colleagues (2010) proposes that cognitive challenges must be sustained (e.g., continuously increasing environmental demands) rather than transient in order to increase functional supply, therefore an adaptive training protocol in which the level of difficulty is capped within a restricted range may be insufficient to induce plasticity that is associated with far transfer. Additionally, adaptivity-related transfer effects in our study were captured by high-difficulty trials, consistent with the prediction that raising the level of maximum function through the manifestation of plasticity would enable previously unattainable high levels of task difficulty to be met. This interpretation is consistent with results from a recent study showing that transfer to a delayed matching-to-sample task was predicted by amount of improvement in high-difficulty blocks of a non-adaptive n-back training program (Beatty et al., 2015). It is possible that near transfer was not found on the Spatial N-Back task in our study because the 3-Back trial type was not sufficiently difficult for healthy young adult participants, as their pre-training scores suggest.

Notably, the far transfer observed in the present study was restricted to the scanned visual episodic memory task and not found within a battery of untrained tasks administered outside of the scanner including measures of verbal episodic memory and also fluid intelligence, the cognitive domain which has been the focus of much controversy regarding far transfer effects from working memory training (Redick et al., 2013; Au et al., 2015; Melby-Lervåg et al., 2016; Greenwood & Parasuraman, 2016). A measure of fluid intelligence also may be considered to represent transfer “farther” from the training tasks in the present study than a measure of visual episodic memory such as the Object-Location Association task. In this respect, although generalization of training gains from working memory to episodic memory is a standard for far transfer in the cognitive training literature (Ranganath et al., 2011; Rudebeck et al., 2012), breadth of transfer can be classified along a continuum (Barnett & Ceci, 2002) and thus skeptics may question whether far transfer to an episodic memory task supported by WMU processes is “far enough” to substantiate the efficacy of cognitive training. The Object-Location Association task used as the scanned task representing far transfer in this study shares features with the visuospatial WMU training task such as the binding of items and spatial context, in addition to demands on executive function. Although conventional models of memory assign the two tasks to different cognitive domains, previous studies demonstrating that processing components involved in working memory and episodic memory are not cleanly dissociable (Ranganath & Blumenfeld, 2005; Nee & Jonides, 2013) suggest that these putatively separate memory domains likewise can be conceptualized along a continuum, with some components—such as the executive function of updating—contributing to processing under both subspan and supraspan conditions. Indeed, shared core cognitive processes appear to account for the transfer from

working memory training to improved episodic memory task performance that was predicted and subsequently observed in this study.

The present study was designed to selectively manipulate the factor of adaptive difficulty between two groups otherwise performing the same WMU training tasks. An alternative interpretation of our results is that the group difference was driven by variable task difficulty, rather than adaptive difficulty *per se*. A recent behavioral study (Bastian and Eschen, 2016) compared conditions in which the difficulty of working memory training tasks was adaptive, self-selected, or randomly varied, and found that all three procedures for varying the level of difficulty produced equivalent improvement on trained tasks, relative to an active control group. However, they also found that transfer effects on untrained working memory tasks and far transfer (reasoning) tasks did not significantly differ among the training groups and the active control group, and thus could not draw conclusions about whether transfer effects such as those observed in the present study are likely to be driven by adaptivity or variability of task difficulty. This is an important issue for future studies to further explore. Another study examining mechanisms of training-induced plasticity compared a group that received adaptive working memory training with an active control group in which task difficulty was yoked to the performance of participants in the adaptive group, and was thus variable but not individually adaptive (McKendrick et al., 2014). The yoked group appeared to reach a performance limit towards the end of the course of training as their performance improvements attenuated relative to the adaptive group, and differential effects were also found in near infrared spectroscopy (NIRS) hemodynamic response measurements across sessions, refuting the hypothesis that adaptive task difficulty and variable task difficulty are similarly effective.

Progress in cognitive training research requires systematic investigations of the factors that influence transfer of training gains, and the neural mechanisms involved. Comparisons across studies are difficult to draw when training protocols differ not only in adaptivity but also frequency and intensity of training, and outcome measurement. In response to recent critiques of the wide variability in training study methodology emphasizing the need for greater experimental rigor and protocol standardization (Shipstead et al., 2012; Green et al., 2014; Noack et al., 2014), our findings support the use of adaptive training as a best practice, at least for targeting WMU processes. Although mixed findings in the cognitive training literature suggest that adaptively increasing training task difficulty is neither necessary nor sufficient to promote transfer, our data show that an optimal design should use adaptive, rather than non-adaptive, training when possible. The present results have important implications for development of cognitive training programs, by demonstrating that adaptive task difficulty influences neural plasticity and transfer of training.

Acknowledgements

This work was supported by NIH grants R01MH068721, R01MH105411, and F32MH096469. We thank Dennis Thompson, Costin Tanase, Luke Jenkins, and Maureen Ritchey for assistance with programming and data processing, Susanne Jaeggi and Martin Buschkuehl for task scripts and helpful discussions, and Audrey Tseng, Chris Kino, Claire Hooker, and Borah Kim for assistance with data collection.

References

- Anguera JA, Boccanfuso J, Rintoul JL, Al-Hashimi O, Faraji F, Janowich J, Kong E, Larraburo Y, Rolle C, Johnston E, Gazzaley A (2013) Video game training enhances cognitive control in older adults. *Nature* 501:97-101.
- Au J, Sheehan E, Tsai N, Duncan GJ, Buschkuhl M, Jaeggi SM (2015) Improving fluid intelligence with training on working memory: A meta-analysis. *Psychon Bull Rev* 22:366-77.
- Barnett SM, Ceci SJ (2002) When and where do we apply what we learn? A taxonomy for far transfer. *Psych Bull* 128:612-37.
- Bastian von CC, Eschen A (2016) Does working memory training have to be adaptive? *Psychol Res* 80:181-94.
- Beatty EL, Jobidon ME, Bouak F, Nakashima A, Smith I, Lam Q, Blackler K, Cheung B, Vartanian O (2015). Transfer of training from one working memory task to another: Behavioural and neural evidence. *Front Syst Neurosci*, 9:86.
- Benedict R, Schretlen D, Groninger L (1998) Hopkins Verbal Learning Test–Revised: Normative data and analysis of inter-form and test-retest reliability. *Clin Neuropsychol* 12:43-55.
- Brehmer Y, Westerberg H, Bäckman L (2012) Working-memory training in younger and older adults: Training gains, transfer, and maintenance. *Front Hum Neurosci* 6:63.
- Bush G, Whalen PJ, Rosen BR, Jenike MA, McInerney SC, Rauch SL (1998) The counting Stroop: an interference task specialized for functional neuroimaging—validation study with functional MRI. *Hum Brain Mapp* 6:270-82.
- Chen T, Li D (2007) The roles of working memory updating and processing speed in mediating age-related differences in fluid intelligence. *Aging Neuropsychol Cogn* 14:631-46.
- Dahlin E, Bäckman L, Stigsdotter Neely A, Nyberg L (2009) Training of the executive component of working memory: Subcortical areas mediate transfer effects. *Restor Neurol Neurosci* 27:405-19
- Dahlin E, Nyberg L, Bäckman L, Stigsdotter Neely A (2008a) Plasticity of executive functioning in young and older adults: Immediate training gains, transfer, and long-term maintenance. *Psychol Aging* 23:720-30
- Dahlin E, Stigsdotter Neely A, Larsson A, Bäckman L, Nyberg L (2008b) Transfer of learning after updating training mediated by the striatum. *Science* 320:1510-12
- Dale AM (1999) Optimal experimental design for event-related fMRI. *Hum Brain Mapp* 8:109-14.

de Rover M, Pironti VA, McCabe JA, Acosta-Cabronero J, Arana FS, Morein-Zamir S, Hodges JR, Robbins TW, Fletcher PC, Nestor PJ, Sahakian BJ (2011) Hippocampal dysfunction in patients with mild cognitive impairment: A functional neuroimaging study of a visuospatial paired associates learning task. *Neuropsychologia* 49:2060-70.

Dweck CS, Henderson VL (1988) *Theories of intelligence: Background and measures*. Champaign-Urbana, IL.

Fischer JS, Rudick RA, Cutter GR, Reingold SC (1999) The Multiple Sclerosis Functional Composite Measure (MSFC): An integrated approach to MS clinical outcome assessment. *Mult Scler* 5:244-50.

Gold JM, Fuller RL, Robinson BM, McMahon RP, Braun EL, Luck SJ (2006) Intact attentional control of working memory encoding in schizophrenia. *J Abnorm Psychol* 115:658-73.

Gould RL (2005) Functional neuroanatomy of successful paired associate learning in Alzheimer's Disease. *Am J Psychiatry* 162:2049-60.

Green CS, Strobach T, Schubert T (2014) On methodological standards in training and transfer experiments. *Psychol Res* 78:756-72

Greenwood PM, Parasuraman R (2016) The mechanisms of far transfer from cognitive training: Review and hypothesis. *Neuropsychology* 30:742-55.

Gronwall DM (1977) Paced auditory serial-addition task: A measure of recovery from concussion. *Percept Mot Skills* 44:367-73.

Hannula DE, Ranganath C (2008) Medial temporal lobe activity predicts successful relational memory binding. *J Neurosci* 28:116-24.

Hazy TE, Frank MJ, O'Reilly RC (2006) Banishing the homunculus: Making working memory work. *Neuroscience* 139:105-18

Heinzel S, Lorenz RC, Pelz P, Heinz A, Walter H, Kathmann N, Rapp MA, Stelzel C (2016). Neural correlates of training and transfer effects in working memory in older adults. *Neuroimage* 134:236-49

Holmes J, Gathercole SE, Dunning DL (2009) Adaptive training leads to sustained enhancement of poor working memory in children. *Dev Sci* 12:F9-15.

Howard DV (1979) *Category norms for adults between the ages of 20 and 80*. Washington D.C.

Hsu NS, Novick JM, Jaeggi SM (2014) The development and malleability of executive control abilities. *Front Behav Neurosci* 8:221.

- Jaeggi SM, Buschkuhl M, Perrig WJ, Meier B (2010a) The concurrent validity of the N-back task as a working memory measure. *Memory* 18:394-412.
- Jaeggi SM, Buschkuhl M, Shah P, Jonides J (2014) The role of individual differences in cognitive training and transfer. *Mem Cognit* 42:464-80.
- Jaeggi SM, Studer-Luethi B, Buschkuhl M, Su Y, Jonides J, Perrig WJ (2010b) The relationship between n-back performance and matrix reasoning — implications for training and transfer. *Intelligence* 38:625-35.
- Jonides J (2004) How does practice makes perfect? *Nat Neurosci* 7:10-1.
- Kelly AMC, Garavan H (2005) Human functional neuroimaging of brain changes associated with practice. *Cereb Cortex* 15:1089-1102.
- Klingberg T (2010) Training and plasticity of working memory. *Trends Cogn Sci* 14:317-24.
- Kühn S, Schmiedek F, Noack H, Wenger E, Bodammer NC, Lindenberger U, Lövdén M (2012) The dynamics of change in striatal activity following updating training. *Hum Brain Mapp* 34:1530-41.
- Luck SJ, Vogel EK (1997) The capacity of visual working memory for features and conjunctions. *Nature* 390:279-81.
- Lövdén M, Bäckman L, Lindenberger U, Schaefer S, Schmiedek F (2010) A theoretical framework for the study of adult cognitive plasticity. *Psychol Bull* 136:659-76
- McKendrick R, Ayaz H, Olmstead R, Parasuraman R (2014) Enhancing dual-task performance with verbal and spatial working memory training: Continuous monitoring of cerebral hemodynamics with NIRS. *Neuroimage* 85:1014-26.
- Melby-Lervåg M, Redick TS, Hulme C (2016) Working memory training does not improve performance on measures of intelligence or other measures of “far transfer”: Evidence from a meta-analytic review. *Perspect Psychol Sci* 11:512-34
- Mishra J, Gazzaley A (2014) Harnessing the neuroplastic potential of the human brain & the future of cognitive rehabilitation. *Front Hum Neurosci* 8:218.
- Miyake A, Friedman NP, Emerson MJ, Witzki AH, Howerter A, Wager TD (2000) The unity and diversity of executive functions and their contributions to complex “frontal lobe” tasks: A latent variable analysis. *Cognit Psychol* 41:49-100.
- Moody DE (2009) Can intelligence be increased by training on a task of working memory? *Intelligence* 37:327-28.

- Morris N, Jones DM (1990) Memory updating in working memory: The role of the central executive. *Br J Psychol* 81:111-21.
- Morrison AB, Chein JM (2011) Does working memory training work? The promise and challenges of enhancing cognition by training working memory. *Psychon Bull Rev* 18:46-60.
- Murdock BB (1976) Item and order information in short-term serial memory. *J Exp Psychol Gen* 105:191-216.
- Navon D (1977) Forest before trees: Precedence of global features in visual-perception. *Cognit Psychol* 9:353-83.
- Nee DE, Brown JW, Askren MK, Berman MG, Demiralp E, Krawitz A, Jonides J (2013) A meta-analysis of executive components of working memory. *Cereb Cortex* 23:264-82.
- Nee DE, Jonides J (2013) Trisecting representational states in short-term memory. *Front Hum Neurosci* 7:796.
- Noack H, Lövdén M, Schmiedek F (2014) On the validity and generality of transfer effects in cognitive training research. *Psychol Res* 78:773-89.
- Nyberg L, Karalija N, Salami A, Andersson M, Wåhlin A, Kaboovand N, Köhncke Y, Axelsson J, Rieckmann A, Papenberg G, Garrett DD, Riklund K, Lövdén M, Lindenberger U, Bäckman L (2016) Dopamine D2 receptor availability is linked to hippocampal-caudate functional connectivity and episodic memory. *Proc Natl Acad Sci USA* 113:7918-23.
- Packard MG, Knowlton BJ (2002) Learning and memory functions of the basal ganglia. *Annu Rev Neurosci* 25:563-93.
- Piekema C, Kessels RP, Mars RB, Petersson KM, Fernández G (2006) The right hippocampus participates in short-term memory maintenance of object-location associations. *Neuroimage* 33:374-82.
- Pollack I, Johnson LB, Knaff PR (1959) Running memory span. *J Exp Psychol* 57:137-46.
- Ranganath C, Blumenfeld RS (2005) Doubts about double dissociations between short- and long-term memory. *Trends Cogn Sci* 9:374-80.
- Ranganath C, Flegal KE, Kelly L (2011) Can cognitive training improve episodic memory? *Neuron* 72:688-91.
- Ranganath C, Ritchey M (2012) Two cortical systems for memory-guided behavior. *Nat Rev Neurosci* 13:713-26.
- Raven J, Raven JC, Court JH (1998) Manual for Raven's Progressive Matrices and Vocabulary Scales. San Antonio, TX: Harcourt Assessment.

- Redick TS, Shipstead Z, Harrison TL, Hicks KL, Fried DE, Hambrick DZ, Kane MJ, Engle RW (2013) No evidence of intelligence improvement after working memory training: A randomized, placebo-controlled study. *J Exp Psychol Gen* 142:359-79.
- Reuter M, Schmansky NJ, Rosas HD, Fischl B (2012) Within-subject template estimation for unbiased longitudinal image analysis. *Neuroimage* 61:1402-18.
- Rudebeck SR, Bor D, Ormond A, O'Reilly JX, Lee ACH (2012) A potential spatial working memory training task to improve both episodic memory and fluid intelligence. *PLoS One* 7:e50431.
- Sadeh T, Shohamy D, Levy DR, Reggev N, Maril A (2011) Cooperation between the hippocampus and the striatum during episodic encoding. *J Cogn Neurosci* 23:1597-608.
- Schmiedek F, Lövdén M, Lindenberger U (2010) Hundred days of cognitive training enhance broad cognitive abilities in adulthood: Findings from the COGITO study. *Front Aging Neurosci* 2:27.
- Schneiders JA, Opitz B, Tang H, Deng Y, Xie C, Li H, Mecklinger A (2012) The impact of auditory working memory training on the fronto-parietal working memory network. *Front Hum Neurosci* 6:173.
- Schweizer S, Grahn J, Hampshire A, Mobbs D, Dalgleish T (2013) Training the emotional brain: Improving affective control through emotional working memory training. *J Neurosci* 33:5301-11.
- Shattuck DW, Mirza M, Adisetiyo V, Hojatkashani C, Salamon G, Narr KL, Poldrack RA, Bilder RM, Toga AW (2008) Construction of a 3D probabilistic atlas of human cortical structures. *Neuroimage* 39:1064-80.
- Shipstead Z, Redick TS, Engle RW (2012) Is working memory training effective? *Psychol Bull* 138:628-54
- Spaniol J, Davidson PSR, Kim ASN, Han H, Moscovitch M, Grady CL (2009) Event-related fMRI studies of episodic encoding and retrieval: Meta-analyses using activation likelihood estimation. *Neuropsychologia*, 47:1765-79.
- Squire LR (2004) Memory systems of the brain: A brief history and current perspective. *Neurobiol Learn Mem* 82:171-77.
- Toril P, Reales J, Mayas J (2016) Video game training enhances visuospatial working memory and episodic memory in older adults. *Front Hum Neurosci* 10:206.
- Unsworth N, Heitz RP, Schrock JC, Engle RW (2005) An automated version of the operation span task. *Behav Res Methods* 37:498-505.

van Overschelde J, Rawson K, Dunlosky J (2004) Category norms: An updated and expanded version of the Battig and Montague (1969) norms. *J Mem Lang* 50:289-335.

Vinogradov S, Fisher M, de Villers-Sidani E (2012) Cognitive training for impaired neural systems in neuropsychiatric illness. *Neuropsychopharmacology* 37:43-76.

Voss MW, Erickson KI, Chaddock L, Prakash RS, Colcombe SJ, Morris KS, Doerksen S, Hu L, McAuley E, Kramer AF (2008) Dedifferentiation in the visual cortex: An fMRI investigation of individual differences in older adults. *Brain Res* 1244:121-31.

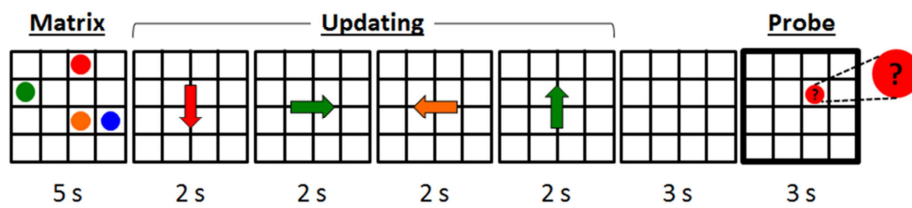
Wager TD, Smith EE (2003) Neuroimaging studies of working memory: A meta-analysis. *Cogn Affect Behav Neurosci* 3:255-74

Wechsler D (1997) Wechsler Adult Intelligence Scale – Third Edition. San Antonio, TX: Psychological Corporation.

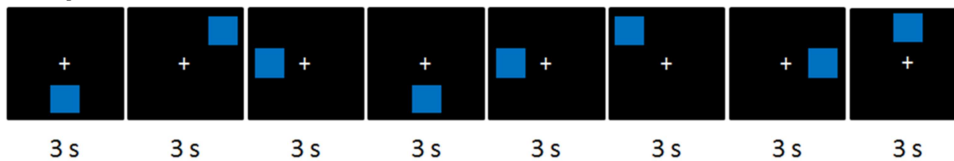
Yntema DB (1963) Keeping track of several things at once. *Hum Factors* 5:7-17.

Fig. 1: Scanned tasks. A. Matrix Updating was modified from the training task version for scanning as a working memory updating (WMU) criterion task; a 4-Updates trial type is depicted. B. Near transfer was assessed with Spatial N-Back, an untrained WMU task. C. Far transfer was assessed with Object-Location Association, an untrained episodic memory task; a 6-Associates trial type is depicted, including a full encoding phase and the first trial of a retrieval phase.

A. Criterion Task (Matrix Updating)



B. Spatial N-Back Task



C. Object-Location Association Task

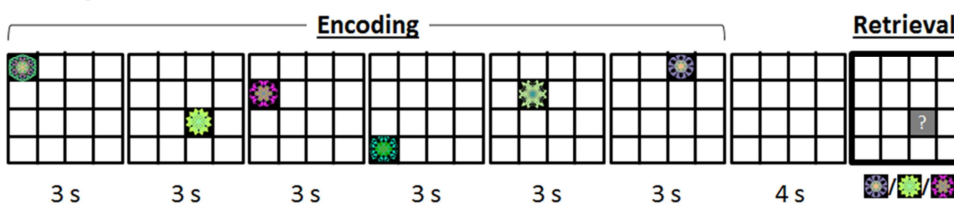


Fig. 2: Training trajectories for adaptive training participants on visuospatial and verbal working memory updating (WMU) tasks. Error bars denote standard error of the mean. See *Materials and Methods* for task descriptions. Non-adaptive active control participants performed same training tasks for the same number of sessions but at a fixed and relatively low, individualized level of difficulty.

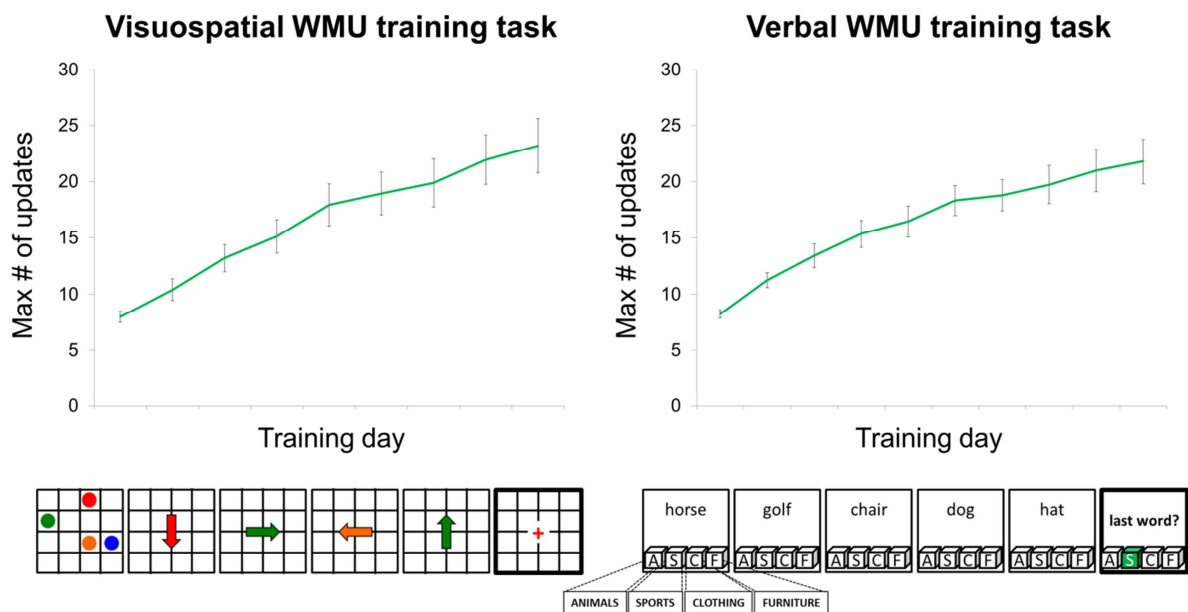


Fig. 3: Adaptive training enhances working memory updating (left) and results in far transfer to an episodic memory task (right). Pre- to post-training performance change is plotted separately for the adaptive training (AT; green), non-adaptive active control (NA; blue), and no-contact control (NCC; red) groups. Error bars denote standard error of the mean.

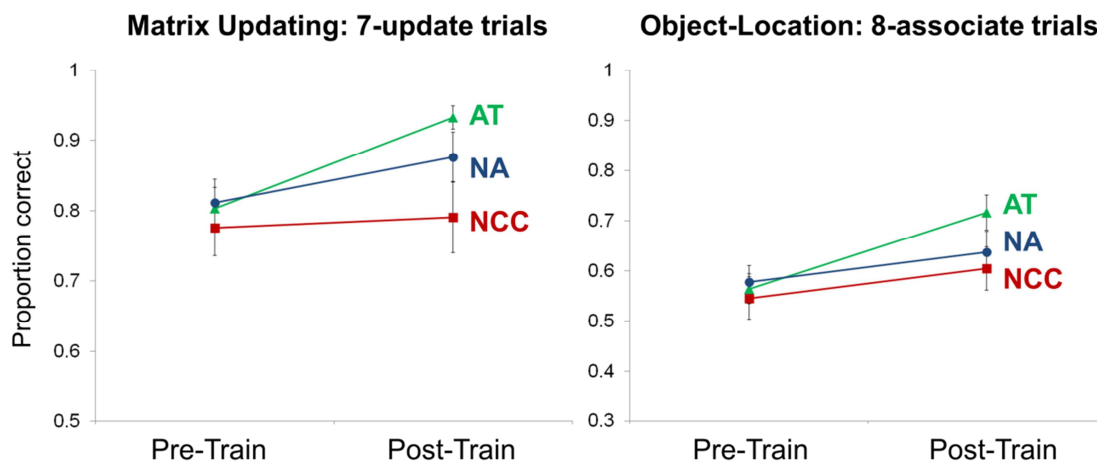


Fig. 4: Adaptive training decreases activation in task-related brain areas. Estimates of load-dependent activation changes (correct 7-update trials vs. correct 4-update trials) during the probe period of the Matrix Updating criterion task are shown for *a priori* striatal and hippocampal regions of interest. Pre-training (dark shading) and post-training (light shading) activation estimates are plotted separately for the adaptive training (AT; green) and non-adaptive active control (NA; blue) groups. Error bars denote standard error of the mean.

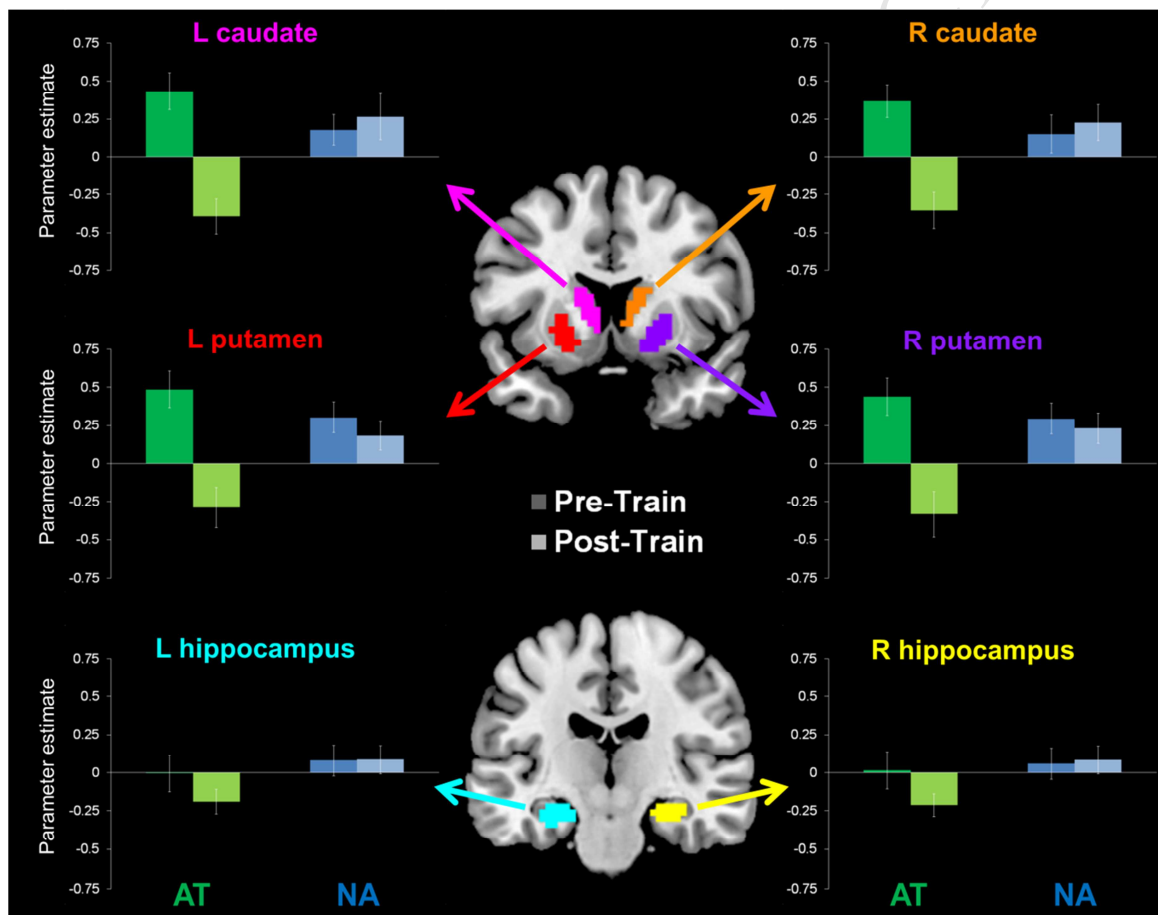


Fig. 5: Exploratory whole-brain analysis of group (AT/NA) by session (pre-training/post-training) interaction for primary contrast of interest in the Matrix Updating criterion task (probe-period activation on correct 7-update trials vs. correct 4-update trials). Activation decreases are greater for AT than NA participants in bilateral prefrontal, bilateral temporal, and left parietal clusters, in addition to bilateral striatum.

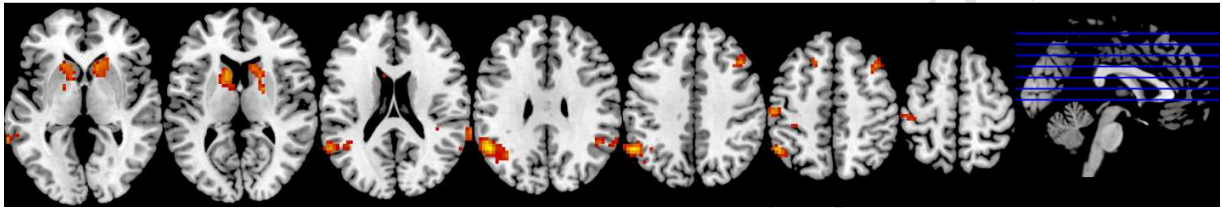


Table 1: Pre- and post-training scanned task behavioral data by group

	Adaptive Training (AT)		Non-Adaptive (NA)		No-Contact Control (NCC)		ANCOVA on post-training performance, controlling for pre-training performance
	Pre	Post	Pre	Post	Pre	Post	
	<i>M</i> (<i>SD</i>)	<i>M</i> (<i>SD</i>)	<i>M</i> (<i>SD</i>)	<i>M</i> (<i>SD</i>)	<i>M</i> (<i>SD</i>)	<i>M</i> (<i>SD</i>)	Effect of group
<u>Matrix Updating criterion task</u>							
7-Updates proportion correct	0.80 (0.13)	0.93 (0.07)	0.81 (0.14)	0.88 (0.15)	0.77 (0.17)	0.79 (0.22)	$F(2,52) = 4.50$, $p < .05$, $\eta_p^2 = .15$
4-Updates proportion correct	0.91 (0.08)	0.95 (0.07)	0.89 (0.11)	0.92 (0.08)	0.85 (0.15)	0.88 (0.16)	$F(2,52) = 1.49$, $p = .24$, $\eta_p^2 = .05$
0-Updates (maintenance only) proportion correct	0.93 (0.08)	0.97 (0.06)	0.93 (0.08)	0.92 (0.11)	0.91 (0.11)	0.89 (0.14)	$F(2,52) = 2.22$, $p = .12$, $\eta_p^2 = .08$
<u>Spatial N-Back</u>							
3-Back accuracy	0.89 (0.07)	0.93 (0.08)	0.91 (0.07)	0.94 (0.03)	0.89 (0.08)	0.94 (0.06)	$F(2,51) = 0.11$, $p = .90$, $\eta_p^2 < .01$
2-Back accuracy	0.94 (0.05)	0.96 (0.05)	0.93 (0.05)	0.97 (0.02)	0.92 (0.08)	0.96 (0.04)	$F(2,51) = 0.34$, $p = .71$, $\eta_p^2 = .01$
0-Back accuracy	0.98 (0.03)	0.98 (0.03)	0.98 (0.02)	0.98 (0.04)	0.98 (0.04)	0.98 (0.03)	$F(2,51) = 0.25$, $p = .78$, $\eta_p^2 = .01$
<u>Object-Location Association</u>							
8-Associates proportion correct	0.56 (0.13)	0.72 (0.15)	0.58 (0.15)	0.64 (0.18)	0.55 (0.18)	0.61 (0.18)	$F(2,52) = 3.75$, $p < .05$, $\eta_p^2 = .13$
6-Associates proportion correct	0.62 (0.17)	0.75 (0.17)	0.62 (0.16)	0.74 (0.16)	0.61 (0.18)	0.74 (0.20)	$F(2,52) = 0.05$, $p = .95$, $\eta_p^2 < .01$

Table 2: Pre- and post-training Matrix Updating criterion task fMRI data by group (as shown in Fig. 4): Parameter estimates extracted from anatomically-defined ROIs for the primary contrast of interest (probe-period activation on correct 7-update trials vs. correct 4-update trials)

	Adaptive Training (AT)		Non-Adaptive (NA)		ANCOVA on post-training activation, controlling for pre-training activation
	Pre	Post	Pre	Post	
Region of interest	<i>M (SD)</i>	<i>M (SD)</i>	<i>M (SD)</i>	<i>M (SD)</i>	Effect of group
L caudate	0.43 (0.53)	-0.39 (0.50)	0.18 (0.44)	0.27 (0.67)	$F(1,35) = 13.13, p < .001, \eta_p^2 = .27$
R caudate	0.37 (0.47)	-0.35 (0.52)	0.15 (0.55)	0.22 (0.52)	$F(1,35) = 10.88, p < .01, \eta_p^2 = .24$
L putamen	0.49 (0.50)	-0.29 (0.58)	0.30 (0.47)	0.18 (0.42)	$F(1,35) = 10.17, p < .01, \eta_p^2 = .23$
R putamen	0.44 (0.58)	-0.33 (0.66)	0.30 (0.52)	0.23 (0.43)	$F(1,35) = 11.09, p < .01, \eta_p^2 = .24$
L hippocampus	-0.01 (0.63)	-0.19 (0.35)	0.08 (0.45)	0.08 (0.40)	$F(1,35) = 4.66, p < .05, \eta_p^2 = .12$
R hippocampus	0.02 (0.73)	-0.21 (0.32)	0.06 (0.35)	0.08 (0.40)	$F(1,35) = 6.05, p < .05, \eta_p^2 = .15$

Table 3: Significant regions of activation from group by session interaction for primary contrast of interest in exploratory whole-brain analysis of the Matrix Updating criterion task fMRI data (as shown in Fig. 5). No suprathreshold clusters were identified in exploratory whole-brain interaction analysis for the Spatial N-Back task or the Object-Location Association task.

	cluster size (voxels)	MNI coordinates			<i>t</i>
		<i>x</i>	<i>y</i>	<i>z</i>	
L IPL	295	-57	-55	37	4.96
R striatum	119	12	17	4	4.87
R MTG	114	54	-19	-14	4.78
L striatum	162	-9	8	13	4.64
L MTG	61	-69	-46	-2	4.40
L SFG/FEF	53	-15	32	52	4.38
L postcentral gyrus	79	-51	-19	52	4.34
R MFG	79	42	23	40	4.33
R STG	109	69	-40	19	4.21

Appendix: Pearson's correlation coefficients for pre-training performance on all untrained tasks (n=56)

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
<u>Matrix Updating criterion task</u> (proportion correct)																
1 7-Updates																
2 4-Updates	.56															
3 0-Updates	.35	.24														
<u>Spatial N-Back</u> (accuracy)																
4 3-Back	.42	.30	.20													
5 2-Back	.45	.48	.48	.66												
6 0-Back	.53	.30	.41	.43	.52											
<u>Object-Location Association</u> (proportion correct)																
7 8-Associates	.10	.13	.06	.25	.19	.18										
8 6-Associates	.35	.46	-.01	.41	.34	.22	.56									
<u>Tasks administered outside of the scanner</u>																
9 Letter Running Memory (accuracy)	.26	.08	.11	.20	.22	.14	.05	-.01								
10 Counting Stroop (interference effect)	-.14	-.13	.01	-.13	-.21	-.29	-.27	-.22	-.36							
11 Global/Local (switch cost)	-.05	-.17	-.04	-.17	-.11	-.21	.06	-.01	.07	.02						
12 Operation Span (partial score)	-.01	.13	-.03	.28	.11	.09	-.11	.08	.35	-.32	-.10					
13 Change Detection (K)	.17	.16	.18	.30	.28	.23	.28	.24	.13	-.34	.16	.29				
14 HVL-T-R (percentage retention)	.05	.02	.01	.14	.17	.06	.31	.15	.27	-.30	.13	.15	.06			
15 Raven's APM (number correct)	.27	.09	.05	.29	.31	.07	.09	.20	.45	-.36	.05	.33	.26	.21		
16 PASAT (proportion correct)	.29	.18	.10	.41	.22	.27	.12	.21	.22	-.27	-.12	.58	.31	.12	.35	
17 Digit-Symbol Substitution (number correct)	.33	.10	.13	.25	.17	.29	.01	.19	.01	-.10	-.10	.15	.19	.03	.09	.32

Note: Significant correlations indicated in **bold** (* $p < .05$; 2-tailed).