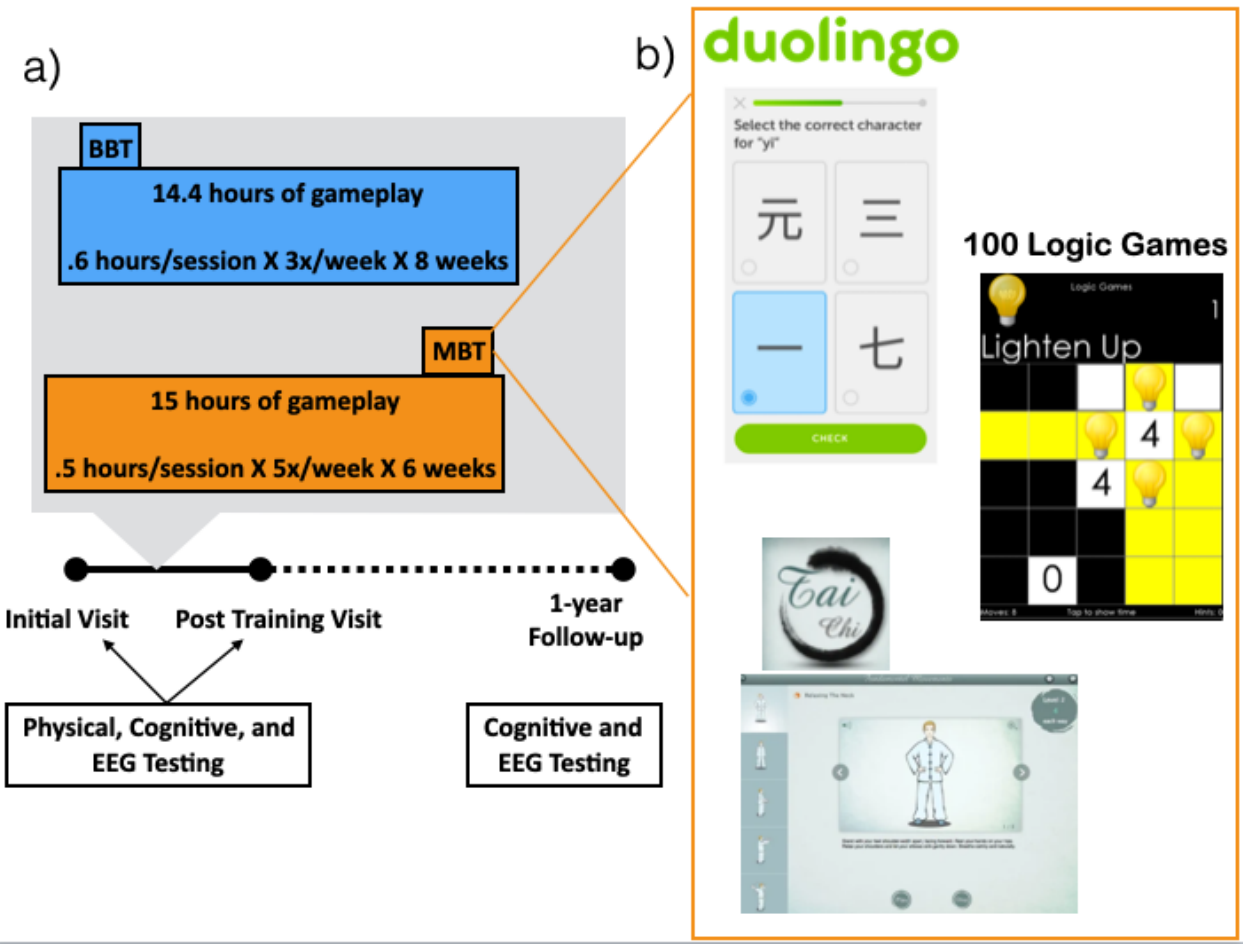
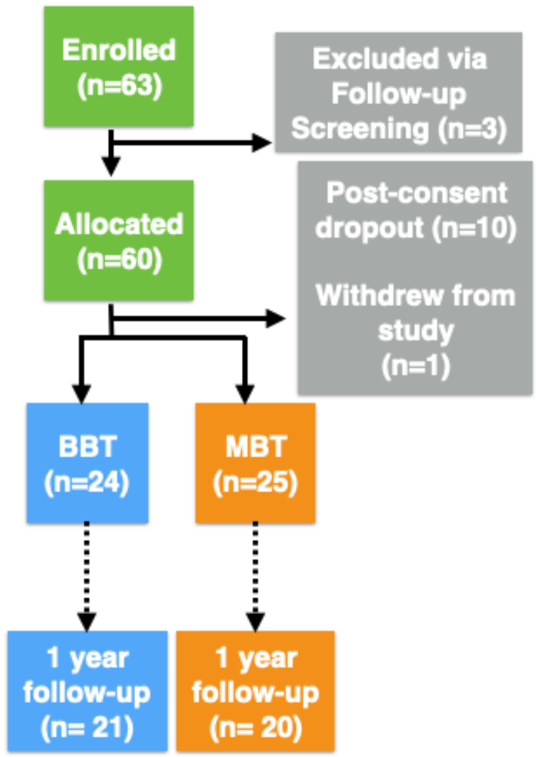
Supplementary Materials



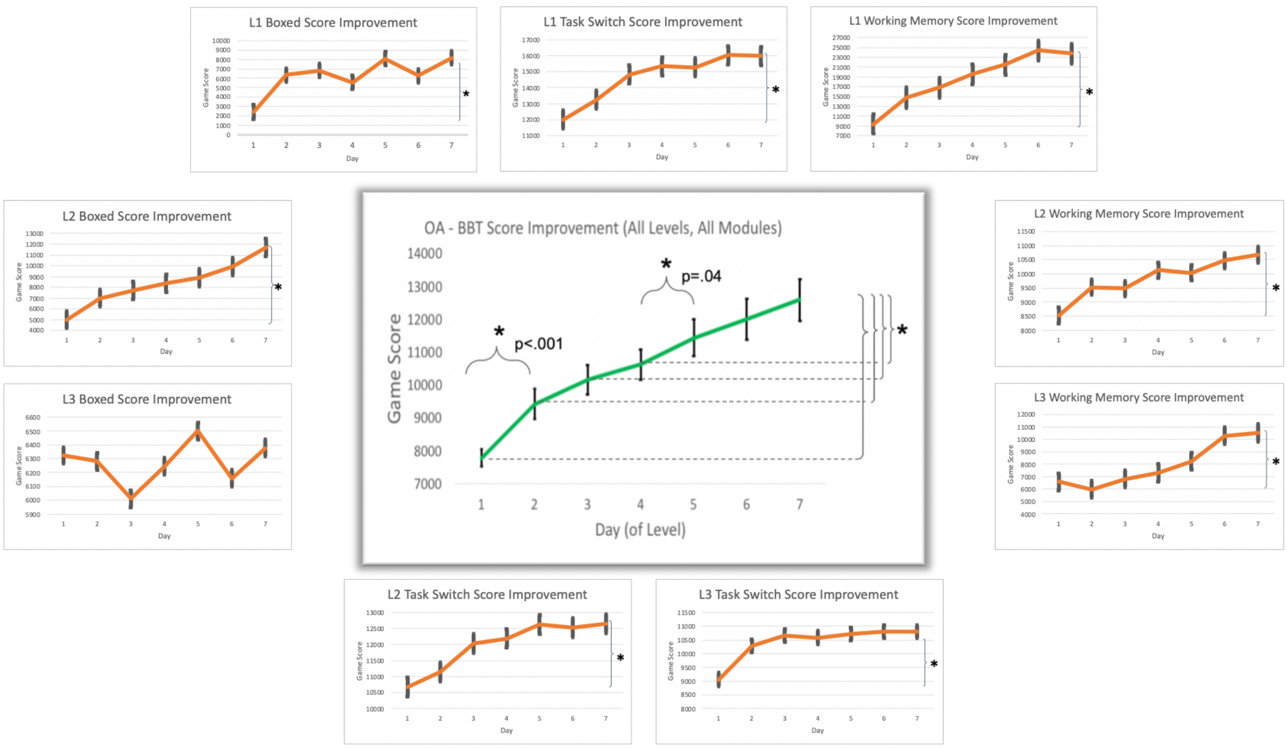
Supplemental Figure 1. Trial Design and Mind-Body Trainer modules. a. An overview of the study timeline as well as the breakdown in gameplay for Body-Brain Trainer (BBT) and Mind-Body Trainer (MBT), the active control. b. We identified a set of three expectancy matched apps to be used as an active control group for BBT: duolingo™, 100 Logic Games™, and Tai Chi™. Participants played each game for ten minutes a day, five times a week, for six weeks.



__**Supplementary Figure 2. CONSORT diagram.**__ Enrolled= how many individuals passed neuropsychological assessment. Excluded via follow-up screening = When confirming an individual wanted to participate, these (3) individuals stated that they could not or no longer wanted to participate. Allocated: Individuals who were consented and confirmed participation interest. Post-consent dropout = after consenting to be in the study, these (10) individuals decided to drop out of the study. 5 individuals dropped out during the preassessment tasks, while the other 5 dropped out during training. Withdrew from study = had a non-study related adverse event precluding continued participation. BBT = Body-Brain Training group, MBT = Mind-Body Training group.

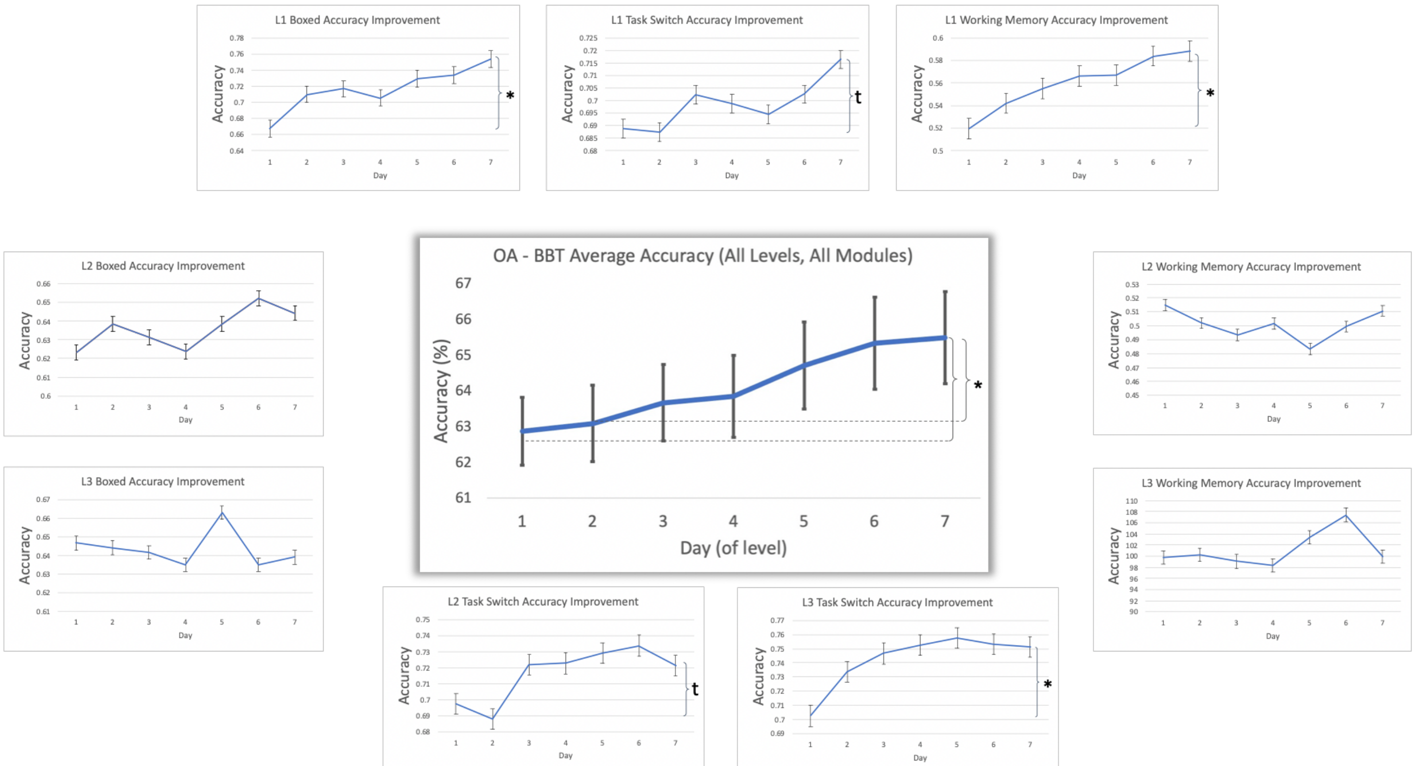
BBT training improvements:

Here we describe the BBT training improvement on the BBT platform itself as a quality control check to demonstrate that participants improved over time on both the cognitive and physical aspects associated with this intervention. With respect to BBT training itself, we found that participants in the BBT group demonstrated increased performance across their training sessions as measured by the amount of points accrued ($t_{21} = 9.40$, $p < 0.001$, **Supplementary Figure 3**), with these points accumulating as a function of training at or above their predetermined cognitive and physiological baseline thresholds (see **Methods** for details). Looking deeper into the individual components contributing to these points, we observed both improved performance accuracy ($t_{21} = -3.34$, $p = 0.003$, **Supplemental Figure 4**) as well as increased heart rate ($t_{21} = -2.31$, $p = 0.031$, **Supplemental Figure 5**) over the course of training, suggesting that, unsurprisingly, cognitive and physiological performance on the intervention itself improved over time. These findings begged the question of whether those individuals who put the most effort into their physical fitness training, in terms of having an elevated heart rate, also showed the greatest improvement in performance accuracy across their training modules over time. This was indeed the case regardless of the stage of the intervention, as we observed a significant correlation between accuracy and heart rate averaged over the initial two days of training ($r_{22} = 0.49$, $p = 0.02$; **Supplemental Figure 6a**) as well as over the final two days of training ($r_{22} = 0.70$, $p = 0.0001$; **Supplemental Figure 6b**).

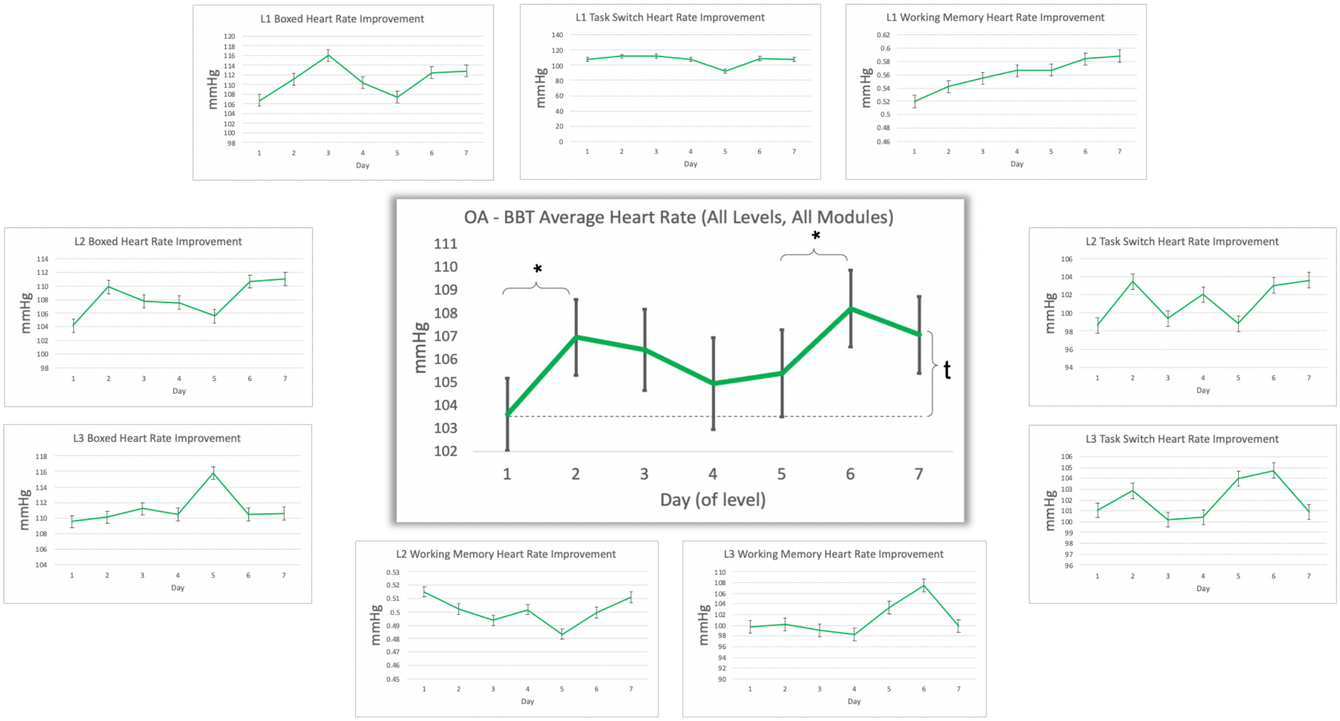


Supplemental Figure 3: Intra-intervention training improvement

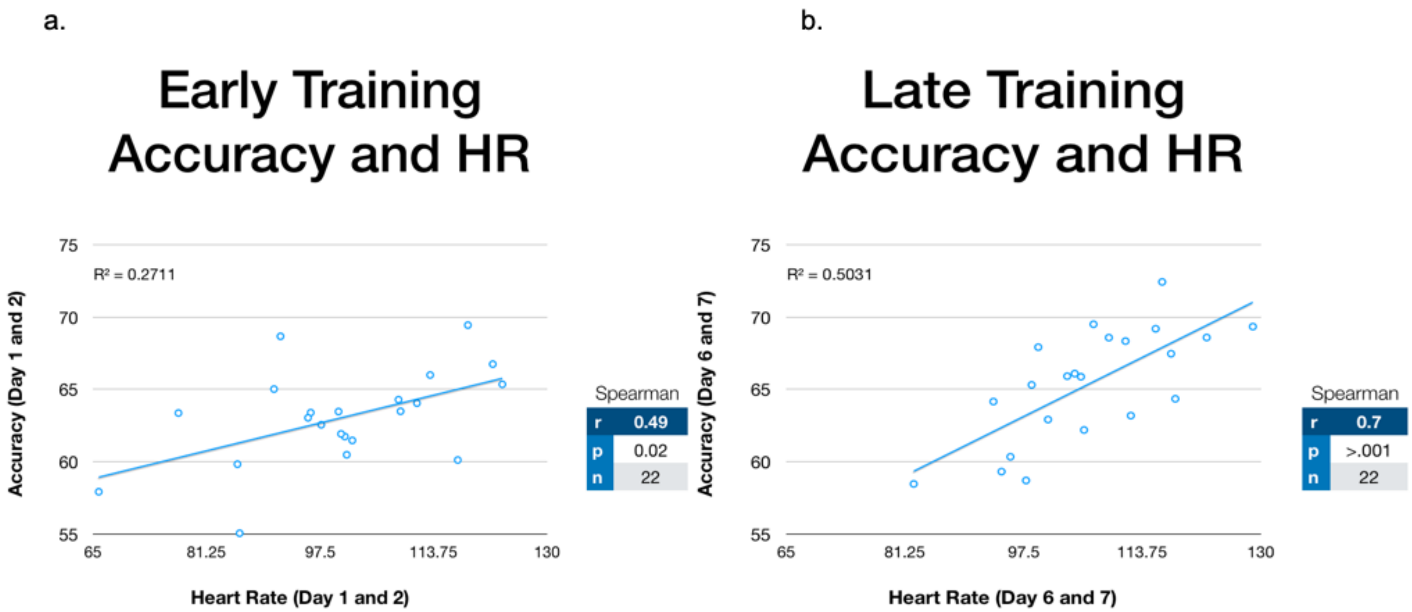
During training on BBT, participants significantly increased their game score from the start to the end of each module for 8 out of 9 modules. Note: game score is dependent on correct answers, difficulty of task, time to reach the correct target, and the distance of those targets to the center of the screen. Taking the average of the nine BBT modules (center figure), participants significantly increased their game score from the start of training (Days 1 and 2, mean= 8640.63 points, s.e. = 492.25) to the end (Days 6 and 7, mean= 12297.48 points, s.e. = 739.50, Δ = 3656.86 points, two-tailed paired t_{21} = -9.396, p = < 0.001, 95% CI: 2847.52 to 4466.19). * = p< 0.05



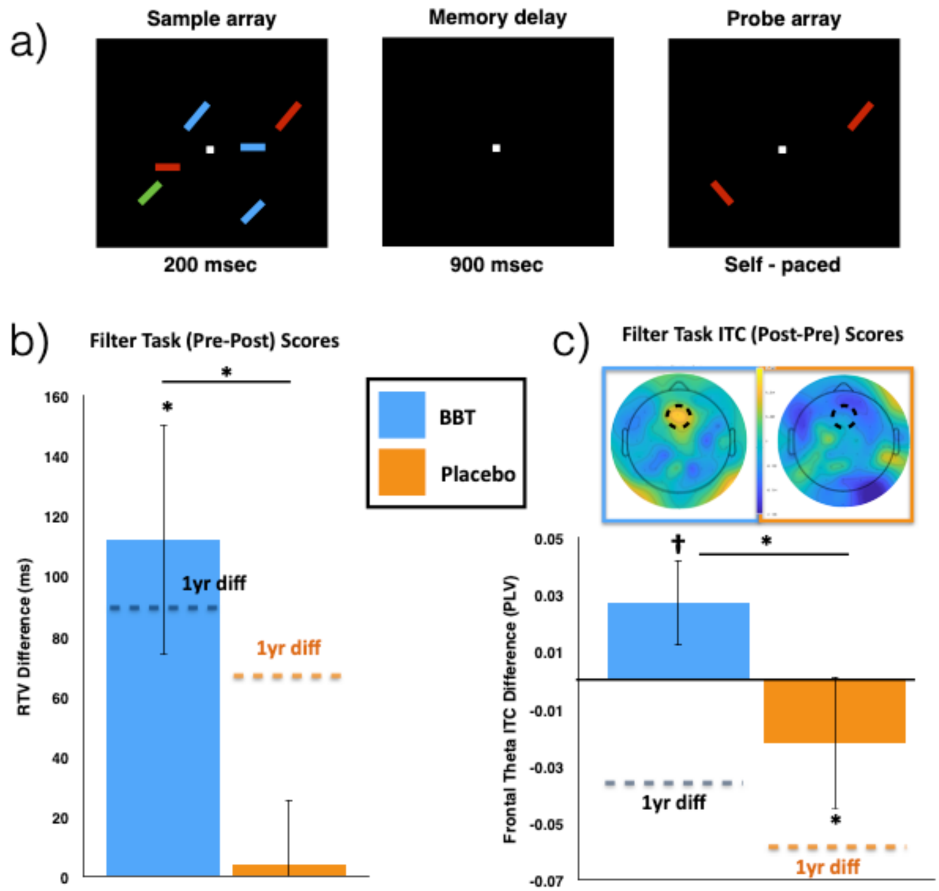
__Supplemental Figure 4: Accuracy across BBT training modules over time.__ During training on BBT, participants significantly increased their accuracy from the start to the end of each module for 3 out of 9 modules. Taking the average of the nine BBT modules (center figure), participants significantly increased their accuracy from the start of training (Days 1 and 2, mean= 63.0%, s.e. = 0.7) to the end (Days 6 and 7, mean= 65.5%, s.e. = 0.7, Δ = 2.5%, two-tailed paired t_{22} = -3.583, p = 0.002, 95% CI: -1.204 to -0.277). * = p< 0.05.



<u>Supplemental Figure 5: Heart Rate across BBT training modules over time</u>: During training on BBT, participants numerically increased their heart rate from the start to the end of each module for 8 out of 9 modules, but none of the changes were significant. Taking the average of the nine BBT modules (center figure), participants showed a trending increase of their heart rate from the start of training (Day 1, mean= 103.6, s.e. = 0.919) to the end (Day 7, mean= 107.1, s.e. = 0.997, Δ = 3.44 mmHG, two-tailed paired t_{22} = -1.925, p = 0.056, 95% CI: -0.306 to 0.004). * = p< 0.05



Supplemental Figure 6: Correlation between heart rate and training accuracy. There was a significant correlation between training accuracy and heart rate, such that the participants who had the higher heart rate, gained a higher cognitive benefit during training. This was the case both early in training (The average of Days 1 and 2, Spearman $r_{22} = 0.49$, $P = 0.02$) as well as late in training (The average of Days 6 and 7, Spearman $r_{22} = 0.7$, $P < 0.001$).



Supplementary Figure 7. Attention with Distraction Task (Filter). a, Stimuli and protocol for the attention with distraction (Filter) task. b, Bar graphs illustrating the group mean change (pre - post) in response time variability (RTV) for each group, with the dashed line (1-year diff = pre - 1-year) illustrating the change in RTV at the 1-year mark. c, Bar graphs and topographic plots illustrating the group mean change (post - pre) in inter-trial coherence (ITC) for each group, with the dashed line on the bar graphs illustrating the change in ITC at the 1-year mark (1-year - pre). The dashed circle on the topographic plot illustrates the electrodes where statistical analyses took place. * = p< .05. † = p= .08. Error bars represent s.e.m.

Attention with Distraction (Filter task):

Given that this task was not a part of the preregistration battery, we performed an exploratory analysis with a more challenging complex visual discrimination task with varying levels of distraction

(**Supplementary Figure 7a**; see **Methods** for task details) to see if comparable improvements (as on the CPT task) might emerge. We chose to evaluate a similar metric to ex-gaussian tau, response time variance (RTV), on this task given that we have previously used this particular measure to assess cognitive enhancement following an intervention[1]. A repeated measures ANOVA evaluating RTV across trials revealed a significant set size (1 or 3 items to be remembered) X time (Pre, Post) X intervention group (BBT, Control) interaction ($F_{(1,42)}$=5.70, p=0.022, Cohen's d = .73). For set size one, follow-up tests revealed performance-related improvements (i.e., less variability) for both the BBT group (Δ = +77.3 msec improvement, t_{24} = 3.26, p = 0.003) and the MBT group (Δ = +61.6 msec improvement, t_{23} = 2.97, p = 0.007). However, for set size three, follow-up tests evidenced performance-related improvements for only the BBT group (Δ = +112.55 msec improvement, t_{20} = 2.93, p = 0.008 for set size of 3), and not for the MBT group (Δ = +4.20 msec improvement, t_{22} = 0.20, p = 0.85 for set size of 3 (**Supplementary Figure 7b**)). With respect to set size three RTV performance one year later, paired t-tests revealed that the BBT group (t_{12} = 1.01, p= .32) showed comparable performance versus their post-training levels. While unexpected, the positive gains by the MBT group at the 1-year mark (t_{12} = 3.25, p= .006) is consistent with previous work demonstrating heightened practice effects in older adults on especially demanding tasks[2-5]. When compared to the group of young adults, the BBT group showed lower performance (greater pre-training RTV) (t_{74} = 2.93, p= .003; see **Supplementary Figure 8**). However, after the intervention period, the BBT group's post-training performance on set size 3 reached an equivalent level of performance (t_{71} = .08, p= .94; **Supplementary Figure 8**).

Neural Correlates of Attention with Distraction:

We also assessed neural correlates associated with training-related improvements in the complex visual discrimination task. Here we examined frontal midline theta inter-trial coherence (ITC), a measurement of trial by trial phase consistency as ITC has been shown to correlate with RTV, the behavioral metric examined on this task[1,6] and is sensitive to intervention-based changes[1,6]. ITC is a measure that reflects the extent to which synchronization occurs from trial to trial in EEG at a particular frequency and latency[7]. In other words, it is a measure of electrophysiological response consistency comparable to the previously described behavioral metrics of ex-gaussian tau and RTV. ITC is quantified by the unit "phase locking value" (PLV), which ranges anywhere between 0 and 1, with a value of 0 indicating that the phase synchrony is completely random, and a value of 1 indicating that the phase-locking is perfectly synchronized across trials. ITC is defined as: ITC(f,t) = 1n∑k = 1nF$_k$(f,t)|F$_k$(f,t)|.

Theta ITC analyses were conducted across all trial types and specifically during the set size of 3 condition of the filter task in accord with the behavioral outcome findings. Based on previous work examining theta ITC during cognitive control tasks[8-14], we focused our analyses on the Fz electrode. All calculated ITC PLVs were controlled for individual state differences at each session by baseline correcting each individual's PLVs using their -200 to 0 period prior to stimulus onset (thus, a relative PLV score was calculated for each subject), as in our previous work[15,16]. To examine a particular time window of interest, we first segmented the resolved ITC time series into 100 msec bins following stimulus onset. To select a particular time window of interest for interrogation, we first conducted a repeated measures ANOVA with session (pre and post) and time window (0-600 in 100msec bins) as the within subjects factors, and group (BBT and MBT) as the between subjects factor. The session by time window by group interaction approached significance ($F_{(4,120)}$= 2.20, p=0.09), with the session by group interaction showing a significant effect ($F_{(1,30)}$= 4.83, p=0.036) supporting a differential improvement in ITC for the BBT group beyond that of the MBT group over all time windows. Given the strong trend for the session by time window by group interaction and our desire to avoid collapsing across all time windows at the individual subject level for correlational analyses, we chose to conduct follow-up repeated measure ANOVAs at each time window (0-100, 100-200, etc) with time (pre and post) as the within subjects factor and group (BBT or MBT) as the between subjects factor. The p-values from each group x time interaction underwent an FDR correction for multiple comparisons, and the time window nearest to stimulus onset that survived the FDR correction (300-400 msec; corrected p = 0.016) was used to select our time window of interest here.

Given that the behavioral results revealed that the BBT group selectively improved beyond the MBT group on the more difficult conditions of this task (set size 3), we focused the neural analyses during this condition. Here, a repeated measures ANOVA with a within-subject factor of time (Pre, Post) and a between-subject factor of Study Group (BBT, Control) showed a significant Group by Time interaction ($F_{(1,39}$=7.98, p=0.007, Cohen's d = .91; **Supplementary Figure 7c**). Post-hoc t-tests revealed an improvement that was trending toward significance (i.e., higher ITC) in the BBT group (Δ = 0.027, paired t_{20} = 1.87, p = 0.08), and a significant decrease in the MBT group (Δ = -0.022, paired t_{19} = -2.40, p = 0.03). With respect to maintenance of any neural changes 1 year later, follow-up tests revealed a both groups regressed from their post-training state (BBT: t_{13}= 2.15, p= .05; MBT: t_{13}= -2.89, p= .01; see **Table 1** for values). When compared to young adults, both BBT and MBT post-training ITC did not differ to young adults at baseline (BBT: t_{54} = 50, p= .62; MBT: t_{53} = 1.00, p= .34) or after the intervention period (BBT: t_{48} = .65, p= .52; MBT: t_{51} = .74, p= .46; **Supplementary Figure 8**).

Supplementary Figure 8. Comparison of BBT and MBT performance versus younger adults.

	BBT Pre-Training	**MBT Pre-Training**	**BBT Post-Training**	**MBT Post-Training**
Filter (RTV)	⬇	=	=	=
Filter (mfTheta ITC)	=	=	=	=

⬇ => performance significantly lesser than that of young adults

⬆ => performance significantly better than that of young adults

= => performance equivalent to that of young adults

Green = Significant improvement from pre-training

Sustained Attention (CPT) RT and RTV results

For response time, using the same repeated measures ANOVA approach as in the main manuscript to evaluate the change in performance from pre-intervention to post-intervention, we found no significant time (Pre, Post) by intervention group (BBT, MBT) interaction ($F_{1,42}$= 1.91, p = 0.17), nor a main effect of session ($F_{1,42}$= 0.86, p = 0.771). Using the same statistical approach, a similar pattern of effects was observed for response time variance with respect to the interrogated interaction ($F_{1,42}$= 1.23, p = 0.27) and the main effect of session ($F_{1,42}$= 0.59, p = 0.44).

BBT cognitive and physical training calibration

For each cognitive task, difficulty was initially determined through a pre-training assessment to determine an optimal training threshold, then scaled on a trial-by-trial basis throughout the training experience. The associated physical challenge for each cognitive task was driven via motion capture of one's hands and feet, where the distance that an individual had to reach/kick to respond during the cognitive task was directly tied to a participant's predetermined optimal training heart rate. For example, if an individual's heart rate was below their training window, then the distance an individual had to traverse to respond on a given trial increased.

This continuous cognitive scaling occurred in parallel with an adaptive cardiovascular challenge, with a pre-training VO_2 max assessment determining an optimal training window with respect to one's heart rate using indirect calorimetry[17] to calculate a more precise level of cardiovascular fitness for each individual. More specifically, this involved participants completing a treadmill graded exercise VO_2 Max test to subsequently determine an individual's maximal heart rate, which subsequently was used to determine the heart rate percentiles (60-70% of one's max HR, 70-80% of one's max HR, etc.) at which participants would train at during the BBT training experience. This predetermined heart rate window was used as a baseline comparator against one's current heart rate, with the responses needed for a given cognitive task scaled to be closer or further away than a previous trial without impeding the ability to perform the cognitive task.

BBT Module Descriptions

(i) BBT Attention. This module demands an active scan of the screen in search of a target, much like traditional visual search tasks[18]. This module involves a constantly evolving amount of cued information as well as number of incongruent distracting elements, as based on our recent work[19], such that participants experience less cued information while experiencing more and more distracting elements as they advance. Participants are required to quickly identify the direction of a probe target that is facing at a right angle (up, down, left, right), and are aided by the presence of directional cue indicating in which location of the screen the target will appear amongst distracting elements. Responses are made by reaching their hands to indicate the direction of the probe, with the additional physical challenge of running in place if the target is up or down. Prior to each level, participants completed a thresholding session to determine the optimal starting point from both a cognitive and physical perspective. After completing their initial 7 training sessions, participants advance to LEVEL 2 of this module which entailed facing a greater challenge: here they encountered an increase in the number and salience of distracting elements, including the presence of congruent distractors, as based on their performance on the previous trial. After completing 14 training sessions, participants moved on to LEVEL 3 of this module: here participants performed the same task as before, but now without the aid of a directional cue. Participants only receive game points when they correctly perform a given trial faster than the predetermined, personalized threshold determined at the beginning of each level to optimize the attentional engagement.

(ii) BBT Working Memory. This module engages spatial working memory resources similar to the Corsi block task[20-22], requiring individuals to memorize an additional stimulus following two consecutive correct responses, with two consecutive incorrect trials leading to one element being subtracted. Participants memorize the location of objects on screen followed by a 5-7 second delay period during which the participants perform a directed physical movement, with a correct response leading to a greater number of potential targets to be memorized on the next trial (and vice versa). Responses are made with both hands and feet by reaching/kicking targets, with additional physical challenges (making a woodchopping motion) occurring during the delay period. Prior to each level, participants completed a thresholding session to determine the optimal starting point from both a cognitive and physical perspective. After completing 7 training sessions, participants are asked to also memorize and report the sequential order in which the targets originally appeared on the screen (LEVEL 2), thus increasing the spatial working memory load. After 14 training sessions, participants perform a working memory/multiple object tracking task that requires memorizing and tracking the targets as they become invisible and move amongst a sea of moving objects (LEVEL 3). Participants receive game points when they correctly complete a working memory trial faster than a predetermined, personalized threshold so as to challenge the underlying cognitive working memory circuitry.

(iii) BBT Task Switching. This module challenges cognitive flexibility resources by requiring participants to rapidly switch their focus based on distinct rules, much like a traditional task-switching paradigm [23-25]. Here a morphing algorithm is used to titrate the perceptual similarity of the target presented, such that a correct trial makes a subsequent exemplar morph more similar to the probe presented (and vice versa). Participants are presented with exemplar objects along with a target, and move their hands to the target object that is most similar to the exemplar presented. For example, when a greenish-blue target appears, participants decide whether the image is more GREEN or more BLUE. The target changes its degree of likeness to each exemplar following each trial, with each correct response morphing the probe towards an indistinguishable 50/50 ratio of each exemplar (and vice versa). Prior to each level, participants completed a thresholding session to determine the optimal starting point from both a cognitive and physical perspective. After completing 7 training sessions, the presented probes now have features that integrate two rule bases (Level 2, e.g. both Color and Shape, so a BLUE square), creating greater cognitive demands, similar to interference generated by a Stroop task. Finally, after 14 training sessions, participants perform the same task, but the exemplars now spawn in random locations across the screen, heightening the cognitive demands further by requiring visual search (LEVEL 3). Participants receive game points when they perform a trial as fast or faster than a predetermined, personalized threshold to pressure underlying goal-management circuitry.

MBT Compliance Tracking:

Because these apps were commercial apps and not designed to send data to our server, we took additional steps to track compliance in the MBT group. On the day they were given their iPad to begin training, participants were also provided an instructional binder with instructions on how to play, a calendar for recording notes and comments throughout their training experience, and were given e-mail support throughout the treatment period. They were instructed to try to split their time equally among the three apps. Upon completion of the training, researchers confirmed the participants' logged training times.

Physical outcome measures:

Here we describe other exploratory measures collected and assessed beyond the two primary measures of diastolic blood pressure and limits of stability. When we evaluated systolic blood pressure, we observed the BBT group showed a trend towards improvement following training (t_{21}= 1.94, p= 0.065), unlike the MBT group (t_{23} = -0.27, p= 0.79). However, the repeated measures ANOVA analysis did not evidence a significant group by time interaction ($F_{(1,45)}$= 2.28, p= 0.13), and at baseline, the BBT group showed a trend towards having greater systolic blood pressure than the MBT group (t_{45}= 1.78, p= .08). For systolic blood pressure, the ANCOVA approach replicated the previously described null interaction effect ($F_{(1,44}$= .40, p= 0.53).

As exploratory measures, we performed physical assessments from the Senior Fitness Test (SFT) to assess global physical performance[26]. Assessments from the SFT battery included: Chair Stand (lower body strength), Arm Curl (upper body strength), Chair Sit-and-Reach Left and Right (lower body flexibility), Back Scratch Left and Right (upper body flexibility), the (8-foot) Timed Up and Go (agility), and Gait Speed (preferred and fast). Note that none of these exploratory measures showed a group X time interaction (Chair Stands: $F_{(1,45)}$= 2.5, p= .12, Arm Curls: $F_{(1,45)}$= 1.72, p= .20, Chair Sit-and-Reach

Left: $F_{(1,45)}$= 1.18, p= .28, Chair Sit-and-Reach Right: $F_{(1,45)}$= 0.34, p= .56, Back Scratch Left: $F_{(1,45)}$= 0.07, p= .80, Back Scratch Right: $F_{(1,45)}$= 0.03, p= .86, Timed Up and Go: $F_{(1,45)}$= 0.37, p= .55, Gait Speed preferred: $F_{(1,45)}$= 0.056, p= .82, Gait Speed fast: $F_{(1,45)}$= 1.32, p= .26).

Neuropsychological Battery

Prior to experimental testing, all participants were evaluated on 3 measures probing for cognitive impairments and depression (the Montreal Cognitive Assessment (MoCA[27]; minimum score of 26; the Geriatric Depression Scale (GDS)[28]; the PHQ-9[29]) as well as nine neuropsychological tests. These nine tests were subdivided into related domains and composite scores of each were calculated for each of the following:

1) Immediate Memory – consisted of the five immediate recall trials from the California Verbal Learning Test (CVLT-II) [30]
2) Delayed Memory – the long delay free and cued recall trials as well as the Recognition measure from the CVLT-II
3) Processing Speed – Digit Symbol task (citation), Executive Composite-DKEFS Trails Condition 2 (numbers only) [31]
4) Cognitive Flexibility (Task Switching) – DKEFS Trails Condition 4 (number-letter switch)
5) Fluency - Verbal Fluency (Animals [32], D- Words (MoCA)

Color vision deficiency was assessed with Ishihara's Tests for Colour Deficiency[33]. To be included in the study, all individuals were required to be within 2 standard deviations (SD) of age-matched controls on all five of the composite scores. Participants were also excluded if two or more composite scores exceeded 1.5 SDs. This procedure provided a thorough characterization of the cognitive status of each OA participant in multiple domains while simultaneously ensuring that their cognitive faculties were comparable to that of their age-matched peers. All participants tested within two SDs of the normative values established for each of these measures.

References

1 Ziegler, D. A. *et al.* Closed-loop digital meditation improves sustained attention in young adults. *Nat Hum Behav* **3**, 746-757, doi:10.1038/s41562-019-0611-9 (2019).

2 Wais, P. E. & Gazzaley, A. External distraction impairs categorization performance in older adults. *Psychol Aging* **29**, 666-671, doi:10.1037/a0037617 (2014).

3 McCarley, J. S., Yamani, Y., Kramer, A. F. & Mounts, J. R. Age, clutter, and competitive selection. *Psychol Aging* **27**, 616-626, doi:10.1037/a0026705 (2012).

4 Becic, E., Boot, W. R. & Kramer, A. F. Training older adults to search more effectively: scanning strategy and visual search in dynamic displays. *Psychol Aging* **23**, 461-466, doi:10.1037/0882-7974.23.2.461 (2008).

5 Rolle, C. E., Anguera, J. A., Skinner, S. N., Voytek, B. & Gazzaley, A. Enhancing Spatial Attention and Working Memory in Younger and Older Adults. *Journal of Cognitive Neuroscience* **29**, 1483-1497, doi:10.1162/jocn_a_01159 (2017).

6 Lutz, A. *et al.* Mental training enhances attentional stability: neural and behavioral evidence. *J Neurosci* **29**, 13418-13427, doi:10.1523/JNEUROSCI.1614-09.2009 (2009).

7 Tallon-Baudry, C., Bertrand, O., Delpuech, C. & Pernier, J. Stimulus specificity of phase-locked and non-phase-locked 40 Hz visual responses in human. *J Neurosci.* **16**, 4240-4249. (1996).

8 Bishop, D. V. & Hardiman, M. J. Measurement of mismatch negativity in individuals: a study using single-trial analysis. *Psychophysiology* **47**, 697-705, doi:10.1111/j.1469-8986.2009.00970.x (2010).

9 Busch, N. A., Dubois, J. & VanRullen, R. The phase of ongoing EEG oscillations predicts visual perception. *J Neurosci* **29**, 7869-7876, doi:10.1523/JNEUROSCI.0113-09.2009 (2009).

10 Caravaglios, G., Castro, G., Muscoso, E. G., Crivelli, D. & Balconi, M. Beta Responses in Healthy Elderly and in Patients With Amnestic Mild Cognitive Impairment During a Task of Temporal Orientation of Attention. *Clin EEG Neurosci* **49**, 258-271, doi:10.1177/1550059416676144 (2018).

11 Ko, D. *et al.* Theta oscillation related to the auditory discrimination process in mismatch negativity: oddball versus control paradigm. *J Clin Neurol* **8**, 35-42, doi:10.3988/jcn.2012.8.1.35 (2012).

12 Muller, V. & Anokhin, A. P. Neural synchrony during response production and inhibition. *PLoS One* **7**, e38931, doi:10.1371/journal.pone.0038931 (2012).

13 Wang, W., Viswanathan, S., Lee, T. & Grafton, S. T. Coupling between Theta Oscillations and Cognitive Control Network during Cross-Modal Visual and Auditory Attention: Supramodal vs Modality-Specific Mechanisms. *PLoS One* **11**, e0158465, doi:10.1371/journal.pone.0158465 (2016).

14 Lutz, A. *et al.* Mental training enhances attentional stability: neural and behavioral evidence. *The Journal of neuroscience : the official journal of the Society for Neuroscience* **29**, 13418-13427 (2009).

15 Anguera, J. A. *et al.* Video game training enhances cognitive control in older adults. *Nature* **501**, 97-101, doi:10.1038/nature12486 (2013).

16 Anguera, J. A. *et al.* Long-term maintenance of multitasking abilities following video game training in older adults. *Neurobiol Aging* **103**, 22-30, doi:10.1016/j.neurobiolaging.2021.02.023 (2021).
17 Shephard, R. J. *et al.* The maximum oxygen intake. An international reference standard of cardiorespiratory fitness. *Bull World Health Organ* **38**, 757-764 (1968).
18 Treisman, A. M. & Gelade, G. A feature-integration theory of attention. *Cognitive Psychology* **12**, 97-136 (1980).
19 Rolle, C. E., Voytek, B. & Gazzaley, A. Exploring the Potential of the iPad and Xbox Kinect for Cognitive Science Research. *Games for Health Journal* **4**, 221-224 (2015).
20 Chechlacz, M., Rotshtein, P. & Humphreys, G. W. Neuronal substrates of Corsi Block span: Lesion symptom mapping analyses in relation to attentional competition and spatial bias. *Neuropsychologia* **64C**, 240-251, doi:10.1016/j.neuropsychologia.2014.09.038 (2014).
21 Claessen, M. H., van der Ham, I. J. & van Zandvoort, M. J. Computerization of the standard corsi block-tapping task affects its underlying cognitive concepts: a pilot study. *Appl Neuropsychol Adult* **22**, 180-188, doi:10.1080/23279095.2014.892488 (2015).
22 Kessels, R. P., van den Berg, E., Ruis, C. & Brands, A. M. The backward span of the Corsi Block-Tapping Task and its association with the WAIS-III Digit Span. *Assessment* **15**, 426-434, doi:10.1177/1073191108315611 (2008).
23 Nieuwenhuis, S. & Monsell, S. Residual costs in task switching: testing the failure-to-engage hypothesis. *Psychon Bull Rev* **9**, 86-92 (2002).
24 Kray, J., Karbach, J. & Blaye, A. The influence of stimulus-set size on developmental changes in cognitive control and conflict adaptation. *Acta Psychol (Amst)* **140**, 119-128, doi:S0001-6918(12)00051-0 [pii]
10.1016/j.actpsy.2012.03.005 (2012).
25 Mari-Beffa, P., Cooper, S. & Houghton, G. Unmixing the mixing cost: contributions from dimensional relevance and stimulus-response suppression. *J Exp Psychol Hum Percept Perform* **38**, 478-488, doi:2011-25208-001 [pii]
10.1037/a0025979 (2012).
26 Jones, C. J. & Rikli, R. E. Measuring functional fitness of older adults. *The Journal on active aging* **1**, 24-30 (2002).
27 Nasreddine, Z. S. *et al.* The Montreal Cognitive Assessment, MoCA: a brief screening tool for mild cognitive impairment. *J Am Geriatr Soc* **53**, 695-699, doi:10.1111/j.1532-5415.2005.53221.x (2005).
28 Yesavage, J. A. *et al.* Development and validation of a geriatric depression screening scale: A preliminary report. *Journal of Psychiatric Research* **17**, 37-49, doi:10.1016/0022-3956(82)90033-4.
29 Kroenke, K., Spitzer, R. L. & Williams, J. B. The PHQ-9: validity of a brief depression severity measure. *Journal of general internal medicine* **16**, 606-613 (2001).
30 Delis, D. C., Kramer, J. H., Kaplan, E. & Ober, B. A. Manual for the California Verbal Learning Test—Second Edition (CVLT–II). *San Antonio, TX: The Psychological Corporation* (2000).
31 Tombaugh, T. N. Trail Making Test A and B: normative data stratified by age and education. *Archives of clinical neuropsychology* **19**, 203-214 (2004).
32 Goodglass, H. & Kaplan, E. *The assessment of aphasia and related disorders*. (Lea & Febiger, 1983).

33 Ishihara, S. *Ishihara's Tests for Colour Deficiency: 24 Plates*. (Kanehara Shuppan, 2010).