Supplementary Material

Gray Matter Responsiveness To Adaptive Working Memory Training: A Surface-Based Morphometry Study

Francisco J. Román[1], Lindsay B. Lewis[2], Chi-Hua Chen[3], Sherif Karama[2], Miguel Burgaleta[4], Kenia Martínez[1, 5], Claude Lepage[2], Susanne M. Jaeggi[6], Alan C. Evans[2], William S. Kremen[3], Roberto Colom[1*]

(1) Universidad Autónoma de Madrid, Madrid (Spain).

(2) Montreal Neurological Institute (MNI), McGill University, Montreal, Canada.

(3) University of California, San Diego (USA)

(4) Universidad Pompeu Fabra, Barcelona (Spain)

(5) Hospital Gregorio Marañon, Madrid (Spain)

(6) University of California at Irvine, Irvine (USA)

Supplementary 1. Tests and tasks administered.

Fluid ability (Gf) was measured by the Raven Advanced Progressive Matrices Test (RAPM)[1] (Raven, Raven & Court, 2004), the abstract reasoning subtests from the Differential Aptitude Test (DAT-AR) battery[2] (Bennett, Seashore & Wesman, 2005), and the inductive reasoning subtests from the Primary Mental Abilities (PMA-R) battery[3] (Thurstone, 1938). Gf was measured by screening versions (odd numbered items and even numbered items for the pretest and posttest evaluations, respectively). The RAPM comprises a matrix figure with three rows and three columns. Among eight possible alternatives the one completing the 3 x 3 matrix figure must be chosen (max. score = 18). DAT-AR is a series test based on abstract figures. Successive figures follow a given rule, so the one continuing the series must be chosen from several alternatives (max. score = 20). PMA-R comprises letters' series items. The rule (or rules) underlying a given sequence must be extracted for selecting the correct alternative (max. score = 15).

Crystallized ability (Gc) was measured by the verbal and numerical reasoning subtests from the DAT (VR and NR), along with the vocabulary subtests from the PMA (V). As in test of Gf, first odd numbered items were administered in pretest and even numbered items in posttest. DAT-VR is based on sentences stated like an analogy. The first and last words from the sentence are missing, and a pair of words completing the sentence must be selected. The screening version comprising odd items only was administered (max. score = 20). PMA-V is a synonym test based on the meaning of words that must be evaluated against a given model word (max. score = 25). DAT-NR consists of quantitative reasoning problems. The screening version comprising odd items only was administered (max. score = 20).

Working memory was measured by the reading span, computation span, and dot matrix tasks[4] (Colom et al., 2010). In the reading span task participants verify if a set of sentences sequentially displayed make or make no sense. Each display includes a sentence and a to-be remembered capital letter. Sentences are 10 – 15 words long. At the end of a given set, participants recall, in their correct serial order, each letter from the set. Set sizes range from 3 to 6 sentence/letter pairs per trial, for a total of 12 trials (4 levels x 3 trial = 12 trials total). The computation span task includes a verification task and a recall task. 6 seconds are allowed to see the math equation without a time limit for verifying its accuracy. The displayed solution, irrespective of its accuracy, must be serially remembered at the end of a given set. Each math equation includes two operations using digits from 1 to 10. The solutions are always single-digit numbers. Trials range from three to seven equation/solutions (5 levels x 3 trials each = 15 trials total). In the dot matrix task, a matrix equation must be verified and a dot location displayed in a five x five grid must be retained. The matrix equation is presented during a maximum of 4.5 seconds for adding or subtracting simple line drawings. Once the response is given, the grid comprising the to-be remembered dot is displayed for 1.5 s. After a given set of equation-grid pairs, the grid spaces that contained dots must be recalled clicking with the mouse on an empty grid. Trials increase in size from two to five equations and dots (4 levels x 3 trials = 12 trials total). The score for these three WM tasks is the number of hits in the verification and recalling tasks.

Attention was measured by verbal and numerical versions of the flanker task, along with a spatial variant of the Simon task[4] (Colom et al., 2010). The verbal and quantitative tasks require deciding, as fast as possible, if the letter/digit presented in the center of a set of three letters/digits is vowel/odd or consonant/even. The target (e.g. vowel/odd) can be surrounded by compatible (e.g. vowel/odd) or incompatible (e.g.

consonant/even) letters/digits. The spatial task (Simon task) requires deciding if an arrow (horizontally depicted) points to the left or to the right of a fixation point. The target arrow pointing to a given direction (e.g. to the left) can be presented at the left (e.g. compatible) or at the right (e.g. incompatible) of the fixation point. There are a total of 32 practice trials and 80 experimental trials. Half of the trials are compatible and they are randomly presented across the entire session. The mean reaction time for the incompatible trials is the dependent measure.

REFERENCES.

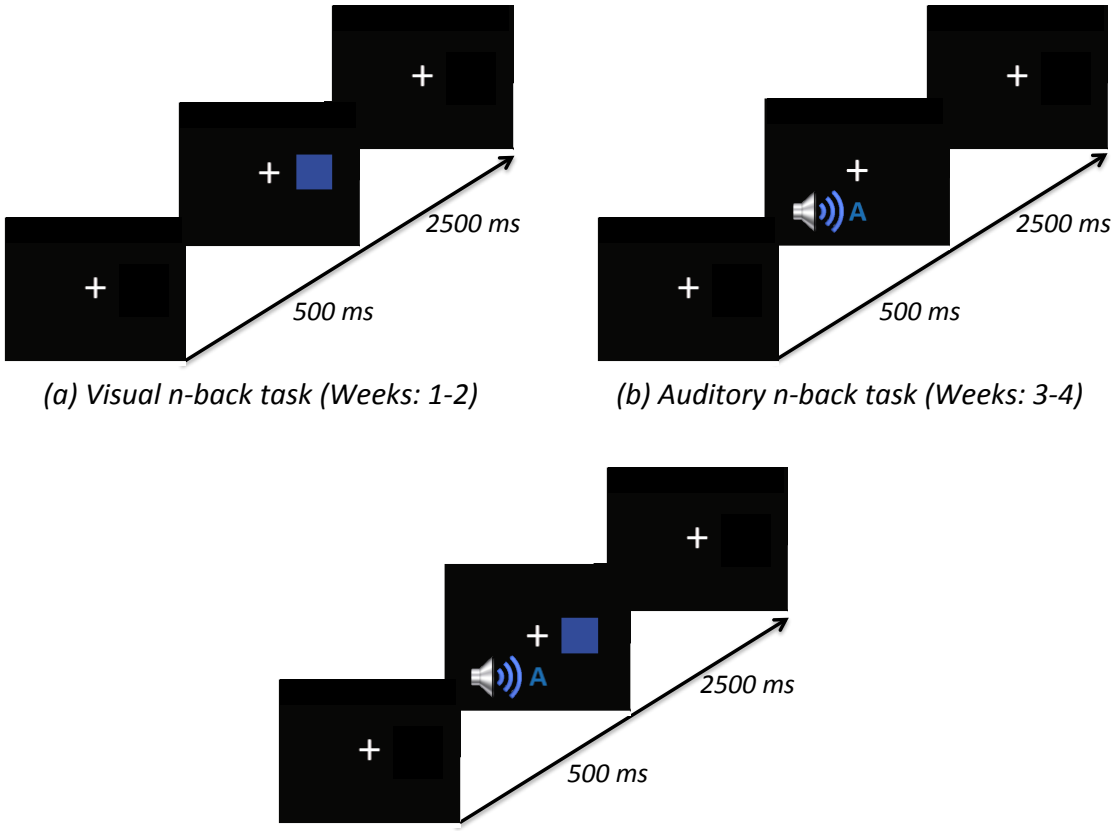
1. Raven, J., Raven, J. C., & Court, J. H. (2004). *Manual for Raven's Progressive Matrices and Vocabulary Scales.* Pearson Assessment: San Antonio, TX.
2. Bennett, G. K., Seashore, H. G., & Wesman, A. G. (2005). *Differential Aptitude Test (DAT-5).* Madrid: TEA, S.A.
3. Thurstone, L. L. (1938). *Primary mental abilities.* University of Chicago Press: Chicago.
4. Colom, R., Quiroga, M. Á., Shih, P. C., Martínez, K., Burgaleta, M., Martínez-Molina, A., ... & Ramírez, I. (2010). Improvement in working memory is not related to increased intelligence scores. *Intelligence, 38,* 497-505.

Supplementary 2. N-back task.

The training group (N = 28) completed the cognitive program based on the dual n-back task (Jaeggi et al., 2008) on individual cabins under strict supervision. However, we began with 8 sessions using the single n-back task, both in a visual and an auditory version. The completed program lasted for twelve weeks and twenty-four sessions (approx. 30 min per session). The first four sessions were devoted to the visual modality, and the subsequent four sessions to the auditory modality. Afterwards, participants completed sixteen dual sessions (visual + auditory). Data were analyzed every week for checking their progress at both the individual and the group level. Participants received systematic feedback regarding their performance.

2. 1. *Visual n-back (single task).*

Each session includes 12 blocks and each block comprises 20 trials + n trials depending of n-back level. (e.g. difficulty n-back level 4 = 24 trials). Blue squares were pseudo-randomly displayed in 8 different positions on the screen. Each square was presented during 500 ms and the participant can provide the answer once the square appears and during 2500 ms. Therefore, the inter-stimulus interval was 3000 ms (see Figure 2.1a).



(a) Visual n-back task (Weeks: 1-2)

(b) Auditory n-back task (Weeks: 3-4)

(c) Dual n-back task (Weeks: 5-12)

Figure 2.1. N-back training program: (a) Visual n-back, (b) Auditory n-back and (c) Dual n-back.

Targets were shown 6 times within each block. Participants must press "Z" when target was presented in the same position than "n" trials before. Hits, errors of commission (false alarms) and errors of omission were computed by block. The first block of the first session began at difficulty level 1 (press the "Z" when the square was presented in the same position than one before). From there, the level of difficulty was adapted according to participants' performance on each block. The rules for increasing, decreasing or keep the level of difficulty in the next block were:

- Increasing: less than 3 errors (commission + omissions) in the block.
- Decreasing: more than 5 errors (commission + omissions) in the block.
- Keep the level of difficulty: between 3 and 5 errors (commission + omissions) in the block.

The level of difficulty obtained in block 12 (the last block of each session), was the level of difficulty of the first block in the next session. Finally, at the end of each session, and at the beginning of a new session, the participant receives feedback regarding her performance.

2. 2. *Auditory n-back (single task).*

Each session includes 12 blocks and each block comprises 20 + n trials. The stimuli were the letters A, E, G, I, K, O, T, and Q. The participant listened to the sequence of letters via headphones. The sequence (see Figure 2.1b) and rules for increasing or decreasing the level were the same than for the visual version. Now participants must press "M" when they listen to the same letter than "n" trials before.

2. 3. *Dual n-back.*

Each session includes 12 blocks with 20 + n back trials per block, where visual and auditory stimuli were presented simultaneously (see Figure 2.1c). The n-back level for both tasks (visual and auditory) was the same for each level. It was not allowed to change the level for one task only. Stimuli were the same than for the single tasks (visual and auditory). The participant must press "Z" for visual stimuli and "M" for auditory stimuli when the square or the letter listened were the same than "n" trials before. Again, targets for both conditions were presented pseudo-randomly. It may or may not happen that targets for visual and auditory conditions showed-up in the same trial (See Figure 2.2).

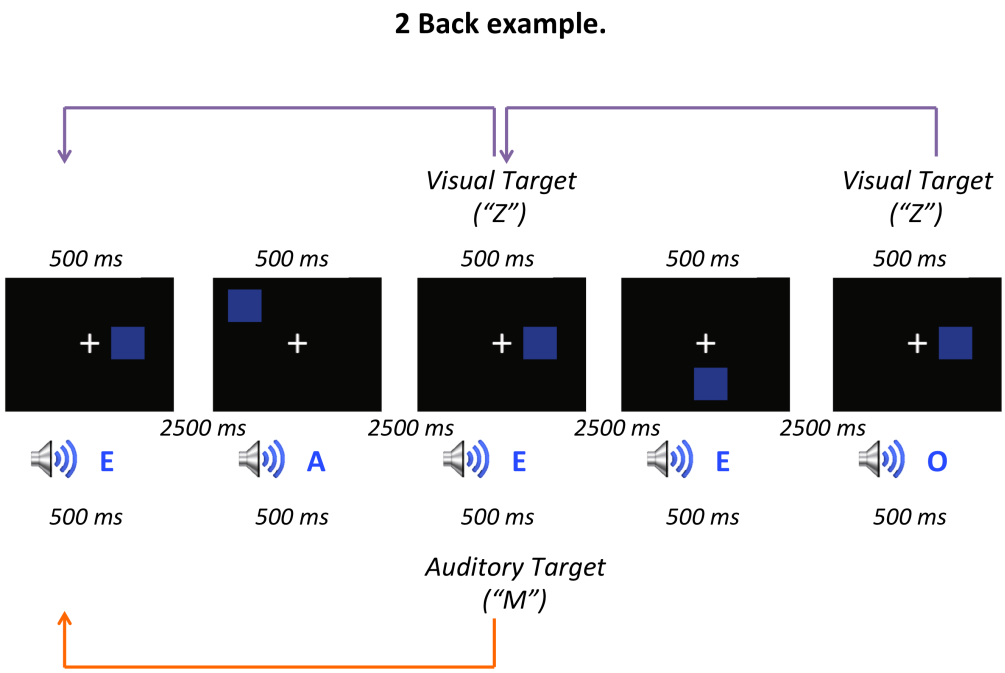


Figure 2.2. Example 2-back level in dual n-back task. The letters were auditory presented at the same rate as the spatial material was presented visually. Participant must press the "Z" for visual targets and "M" for auditory targets.

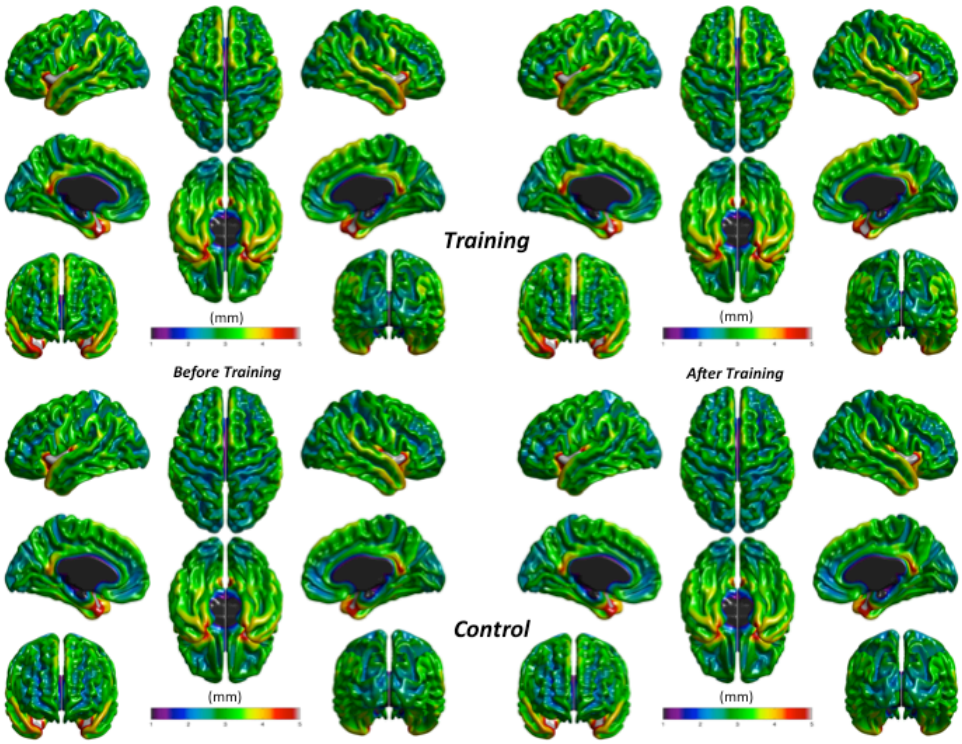
Targets were presented 6 times per block. The rules for increasing or decreasing the level after a block were the same than for the single n-back task. However, errors for both conditions (visual + auditory) were considered simultaneously.

REFERENCE.

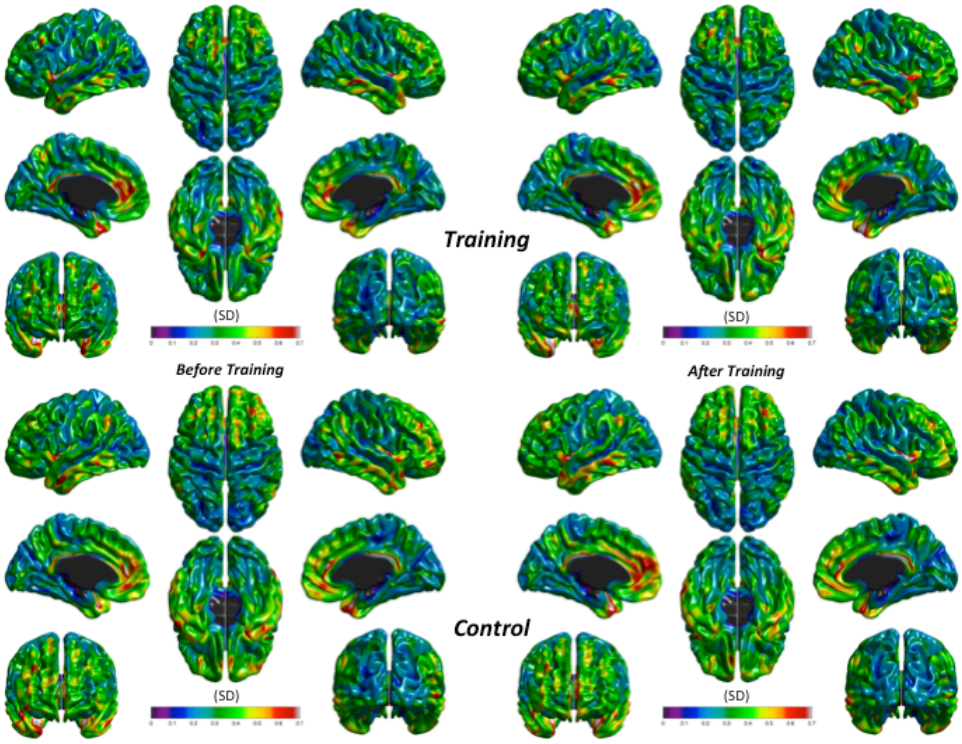
1. Jaeggi, S. M., Buschkuehl, M., Jonides, J., & Perrig, W. J. (2008). Improving fluid intelligence with training on working memory. *Proceedings of the National Academy of Sciences, 105,* 6829-6833.

Supplementary 3. Descriptive maps.

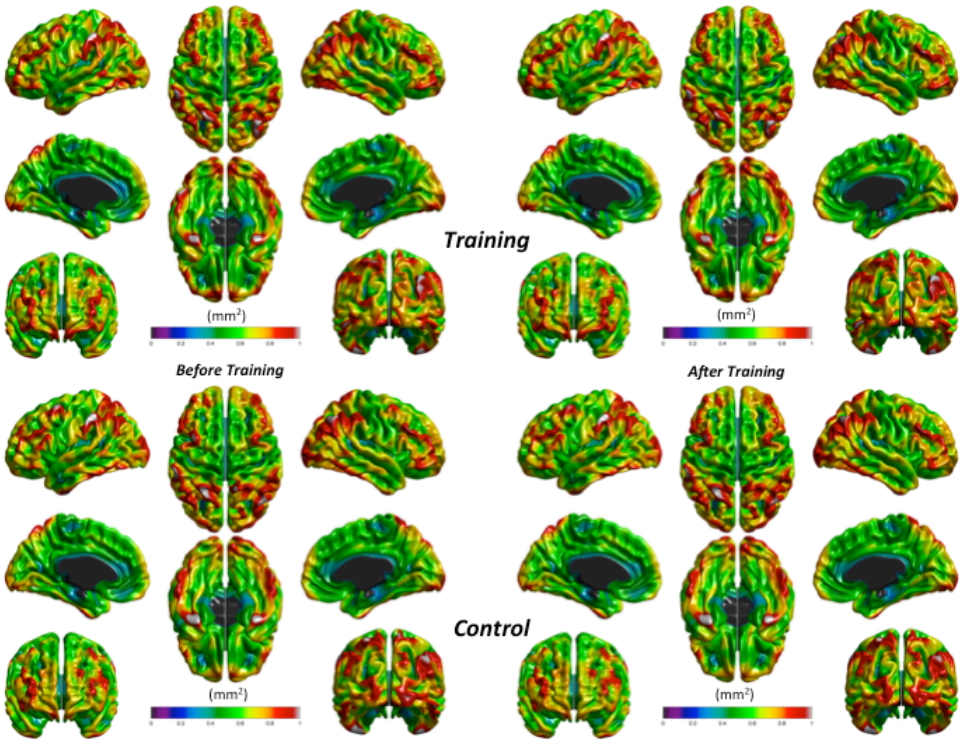
Supplementary Figures 1, 2, 3 and 4 show the distribution (means and standard deviations) for cortical thickness (CT) and cortical surface area (CSA) at each vertex before and after training in both groups. The maps were largely similar for both groups in both time points for cortical thickness; the insula was the region showing the highest mean thickness (> 5mm). Maps for cortical surface were also highly alike; parietal, frontal and temporal areas were the regions where mean surface area was highest.



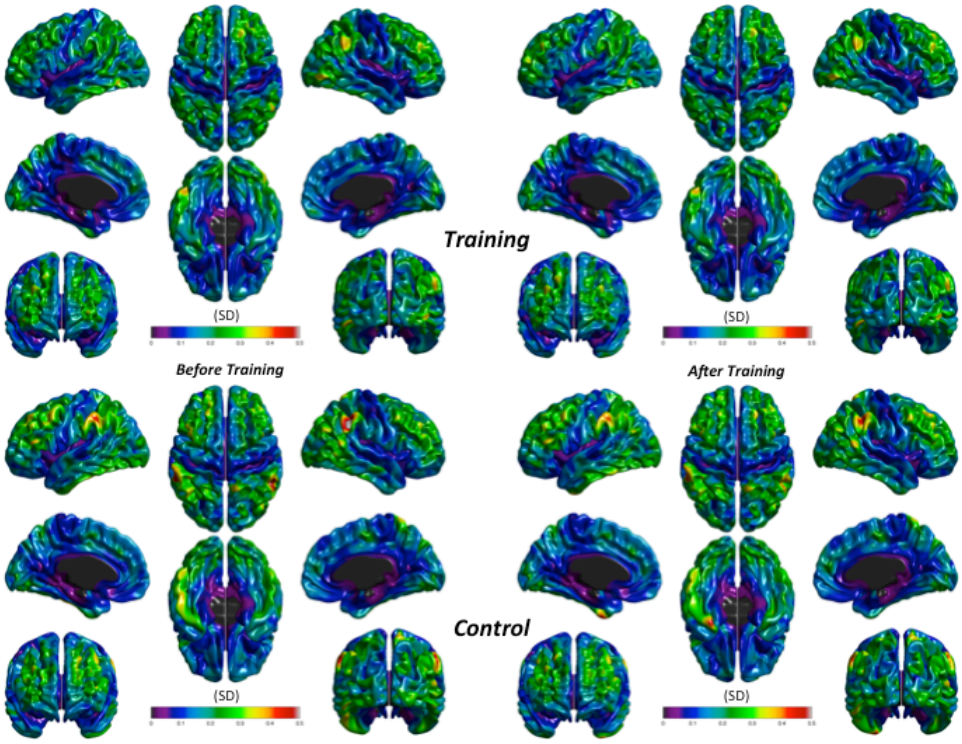
Supplementary Figure 1. Mean cortical thickness at the vertex level before and after training (Top Panel – training group, Bottom Panel – control group).



Supplementary Figure 2. Cortical thickness standard deviation at the vertex level before and after training (Top Panel – training group, Bottom Panel – control group).



Supplementary Figure 3. Mean cortical surface area at the vertex level before and after training (Top Panel – training group, Bottom Panel – control group).



Supplementary Figure 4. Cortical surface area standard deviation at the vertex level before and after training (Top Panel – training group, Bottom Panel – control group).

Table Supplementary 1. Standardized Change of Cortical Thickness and ANCOVA results.

	Standardized Change: Cortical Thickness						
	Training		Control		ANCOVA		
	M	*SD*	*M*	*SD*	*F*	*p*	η^2
Left ROI 1	-0.03	0.76	-0.04	0.93	1.54	0.111	0.03
Left ROI 2	0.08	0.67	-0.05	0.91	2.03	0.081	0.04
Left ROI 3	0.01	0.66	-0.07	0.96	1.54	0.111	0.03
Left ROI 4	0.10	0.68	-0.01	0.86	1.98	0.083	0.04
Left ROI 5	0.03	0.71	-0.15	1.09	0.86	0.179	0.02
Left ROI 6	-0.10	0.72	-0.12	0.86	0.92	0.171	0.02
Left ROI 7	-0.02	0.61	-0.12	0.96	1.12	0.147	0.02
Left ROI 8	0.08	0.80	0.08	0.96	1.07	0.153	0.02
Left ROI 9	0.03	0.75	-0.07	0.90	0.53	0.235	0.01
Left ROI 10	-0.14	0.52	-0.03	0.84	0.01	0.462	0.00
Left ROI 11	-0.11	0.68	-0.12	1.06	1.23	0.136	0.03
Left ROI 12	0.01	0.73	-0.03	0.97	0.50	0.242	0.01
Right ROI 1	0.04	0.64	0.04	0.85	1.87	0.089	0.04
Right ROI 2	0.02	0.70	-0.06	0.67	1.20	0.140	0.02
Right ROI 3	0.08	0.85	-0.01	0.86	0.96	0.166	0.02
Right ROI 4	0.08	0.81	-0.05	0.77	2.26	0.070	0.04
Right ROI 5	-0.09	0.77	-0.21	0.83	1.96	0.084	0.04
Right ROI 6	0.10	0.85	-0.10	0.78	1.35	0.126	0.03
Right ROI 7	**0.08**	**0.67**	**-0.20**	**0.84**	**4.55**	**0.019**	**0.09**
Right ROI 8	0.03	0.90	-0.11	0.84	1.67	0.102	0.03
Right ROI 9	**0.19**	**0.94**	**-0.17**	**0.91**	**4.02**	**0.026**	**0.08**
Right ROI 10	**0.12**	**0.77**	**-0.25**	**0.78**	**7.38**	**0.005**	**0.13**
Right ROI 11	**0.08**	**0.60**	**-0.05**	**0.83**	**3.79**	**0.029**	**0.07**
Right ROI 12	0.06	0.59	-0.07	0.85	2.47	0.061	0.05

Table Supplementary 2. Standardized Change of Cortical Thickness and ANCOVA results.

	Standardized Change: Cortical Surface Area						
	Training		Control		ANCOVA		
	M	*SD*	*M*	*SD*	*F*	*p*	η^2
Left ROI 1	0.00	0.29	0.00	0.15	0.02	0.452	0.00
Left ROI 2	0.03	0.29	0.01	0.25	0.12	0.367	0.00
Left ROI 3	-0.01	0.33	0.05	0.27	0.54	0.233	0.01
Left ROI 4	-0.02	0.47	0.02	0.39	0.02	0.451	0.00
Left ROI 5	-0.02	0.30	-0.06	0.20	0.34	0.280	0.01
Left ROI 6	-0.04	0.24	-0.08	0.44	0.06	0.406	0.00
Left ROI 7	-0.02	0.35	-0.03	0.30	0.04	0.419	0.00
Left ROI 8	-0.01	0.40	0.03	0.36	0.23	0.317	0.01
Left ROI 9	0.03	0.16	-0.04	0.16	2.52	0.060	0.05
Left ROI 10	-0.02	0.18	-0.03	0.15	0.04	0.417	0.00
Left ROI 11	0.02	0.25	-0.06	0.25	1.40	0.122	0.03
Left ROI 12	0.03	0.15	0.04	0.18	0.10	0.377	0.00
Right ROI 1	0.03	0.15	0.05	0.18	0.19	0.334	0.00
Right ROI 2	-0.02	0.29	-0.03	0.44	0.06	0.408	0.00
Right ROI 3	0.00	0.20	0.10	0.27	2.38	0.065	0.05
Right ROI 4	0.07	0.32	0.03	0.32	0.35	0.280	0.01
Right ROI 5	**0.01**	**0.17**	**-0.09**	**0.16**	**6.28**	**0.008**	**0.11**
Right ROI 6	0.00	0.21	-0.14	0.40	1.81	0.093	0.04
Right ROI 7	**0.10**	**0.26**	**-0.15**	**0.39**	**6.13**	**0.009**	**0.11**
Right ROI 8	-0.08	0.41	0.10	0.43	2.75	0.052	0.05
Right ROI 9	-0.01	0.14	0.01	0.15	0.12	0.364	0.00
Right ROI 10	-0.01	0.15	0.02	0.17	0.54	0.232	0.01
Right ROI 11	0.01	0.25	-0.05	0.30	0.50	0.241	0.01
Right ROI 12	0.03	0.24	0.00	0.18	0.28	0.301	0.01