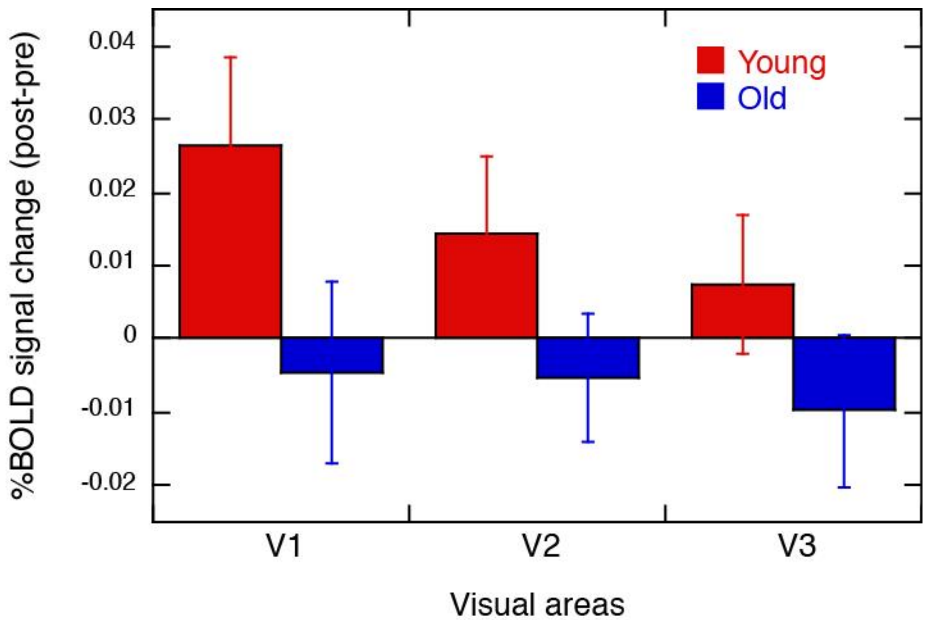
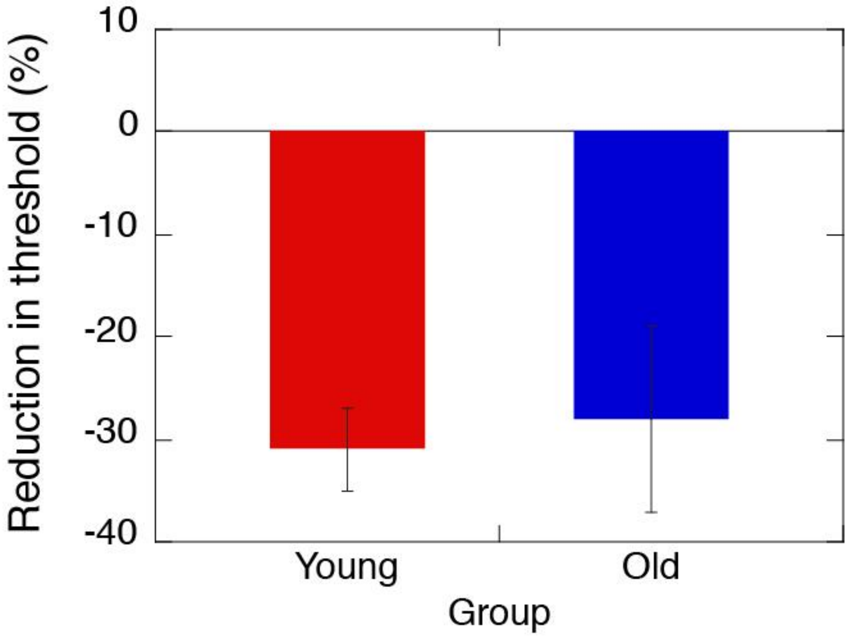
Supplementary Figure 1



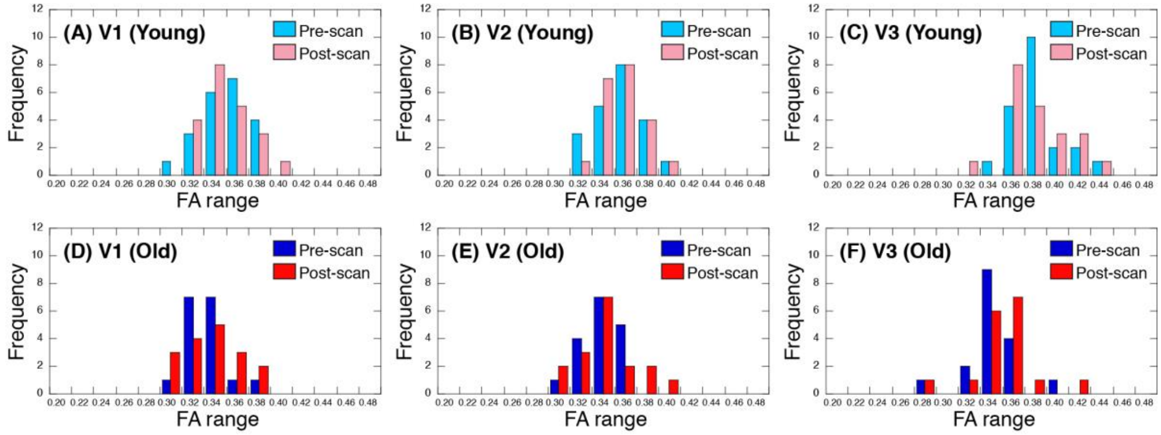
Supplementary Fig. 1. BOLD signal changes in association with perceptual learning in the trained regions of the early visual area. Average BOLD signal changes between pre- and post-training scans are shown (mean ± SE). Red bars show younger adults (n=21) and blue older (n=17). We applied 3-way ANOVA with factors being age (young vs old), visual areas (V1, V2 and V3) and location (trained vs untrained regions). The results indicate a significant interaction between age and location ($F(1,36) = 5.315$, $p = 0.027$). No significant main effect of age, visual areas, or location was found.

Supplementary Figure 2



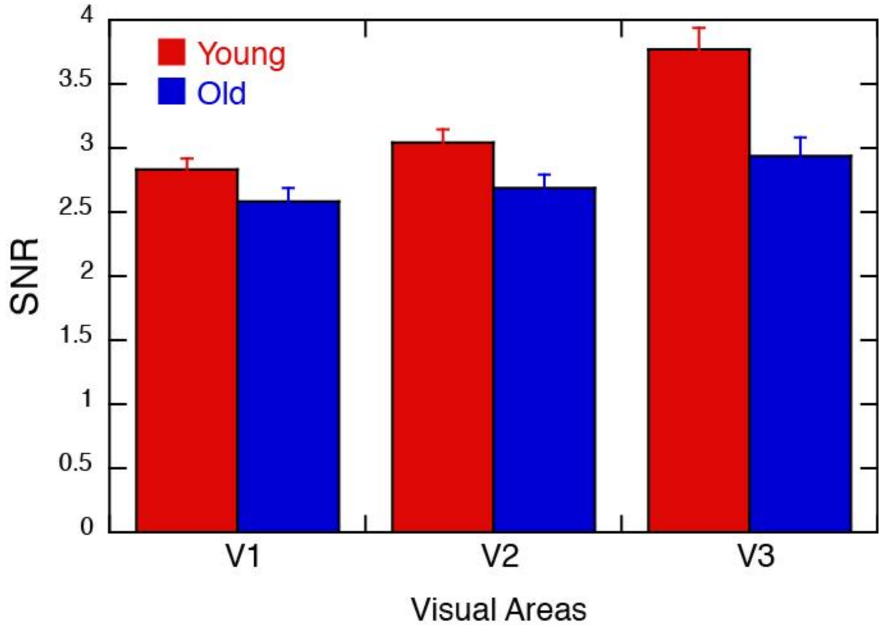
Supplementary Fig. 2. Performance improvement shown by reduction in the threshold SOA over 3 days. The overall performance improvement ratios (mean ± SE) were not statistically different between the younger (n=21) and older (n=17) groups.

Supplementary Figure 3



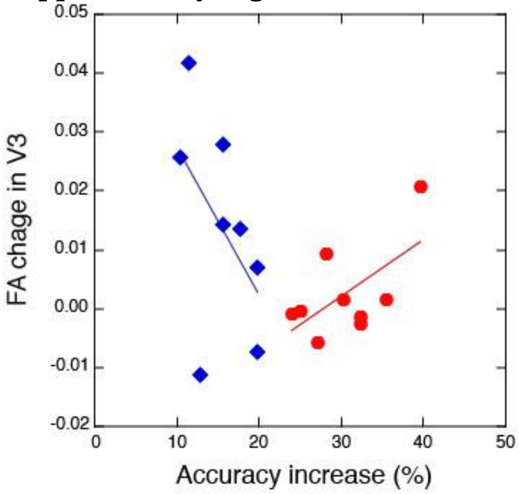
Supplementary Fig. 3. Distribution of FA values for each VOI for each age group. There was no ceiling effect in FA values in the younger (A-C) and older (D-F) groups.

Supplementary Figure 4



Supplementary Fig. 4. SNR of DTI data for each VOI for each age group. We applied 4-way ANOVA with factors being age (old vs young), VOI (V1, V2, and V3), location (trained vs untrained regions), and time (pre- vs post-training). Significant main effects of age ($F(1,36) = 12.777$, $p = 0.001$), VOI ($F(2,72) = 32.763$, $p < 0.0001$), and a significant interaction of age and VOI ($F(2,72) = 7.298$, $p = 0.0013$) were found. Bars (mean ± SE) in red show younger adults (n=21) and blue older (n=17).

Supplementary Figure 5



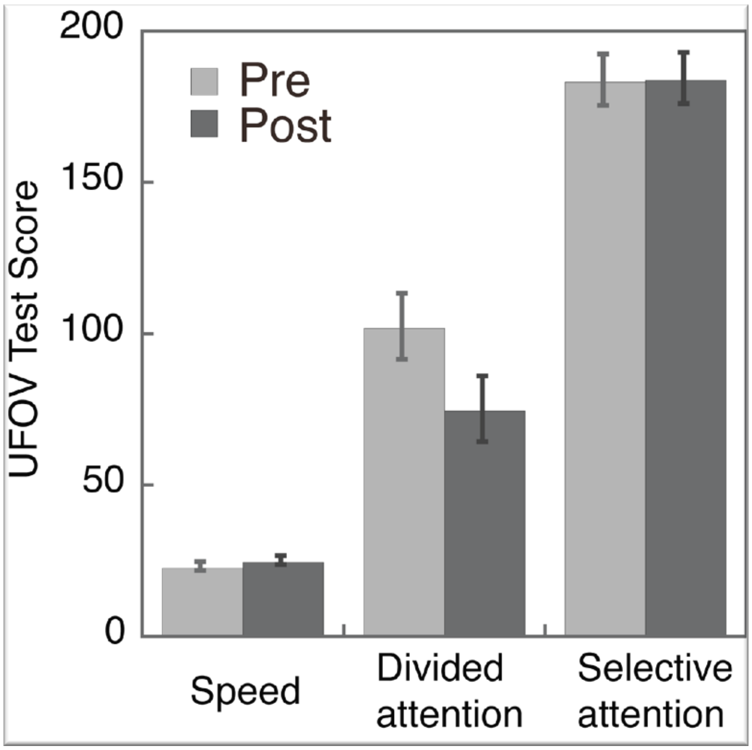
Supplementary Fig. 5. The piecewise linear regression model between the behavioral improvement and FA change beneath the trained quadrant of the V3 for older adults. Among all the linear piecewise models using different breakpoint in the accuracy increase, we chose the model that had the lowest residual sum of squares. A single liner model (R^2 = 0.1382, *df* = 15, *p* = 0.1416) fit worse than the chosen model. The piecewise linear regression model that we chose has the breakpoint of 22.8% the accuracy improvement and the residual sum of squares of 0.00199409, where the coefficient of determination is significant (R^2 = 0.3955, *df* = 15, *p* = 0.0068). Blue dots show the individual data whose accuracy increase was less than 22.8%, whereas the red more than 22.8%. The model can be written as follows:

FA change in V3 = -0.2521 * Accuracy increase + 0.0528 (for blue dots)

FA change in V3 = 0.0954 * Accuracy increase -0.0265 (for red dots).

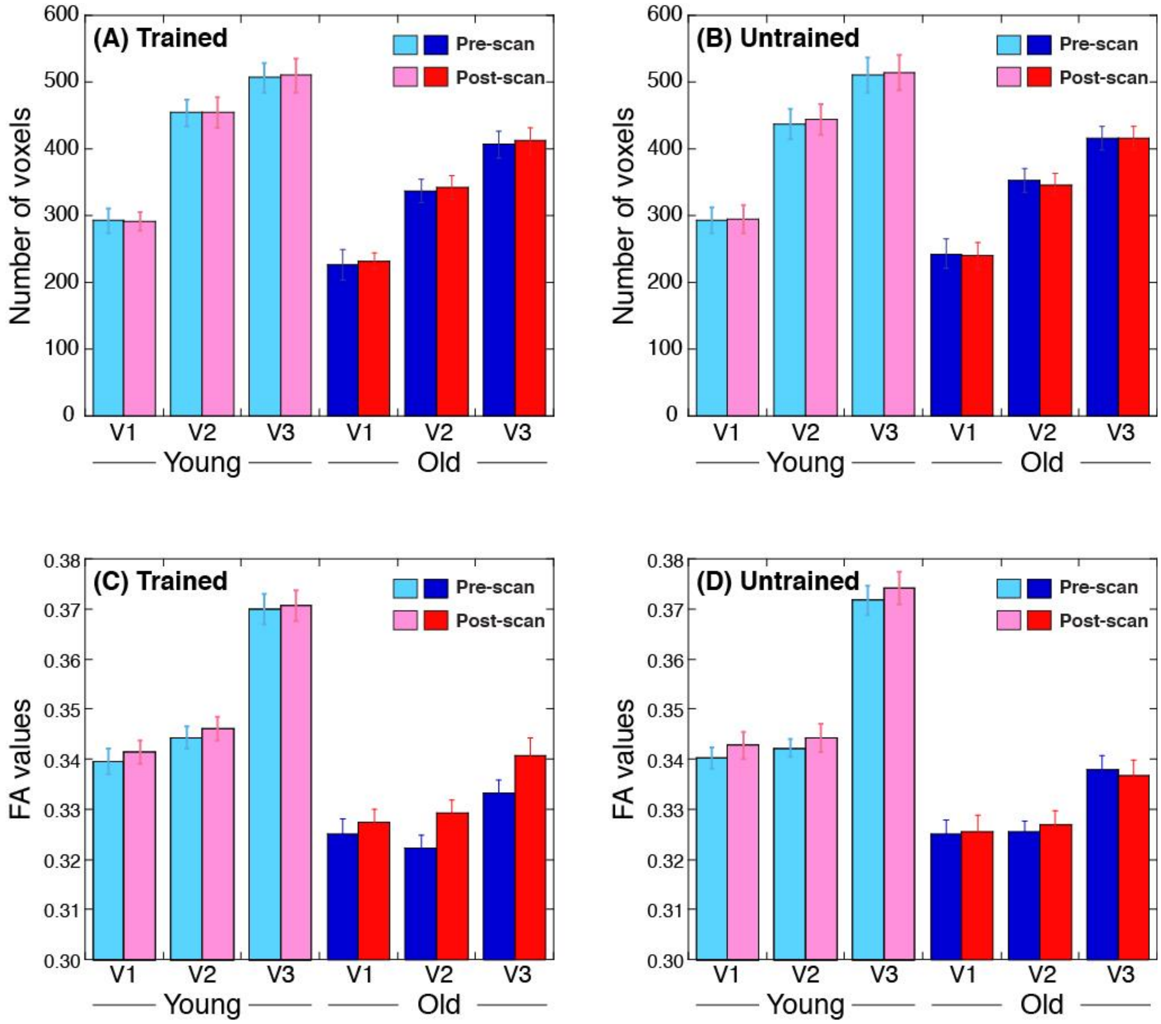
These results suggest that there are 2 types of subjects, poor learning group and good learning group. For the poor learning group whose accuracy increase after training was less than 22.8%, the FA increase was negatively correlated with the amount of accuracy increase. For the good learning group whose accuracy increase was larger than 22.8%, the FA increase was positively correlated with the amount of accuracy increase.

Supplementary Figure 6



Supplementary Fig. 6. The results of the UFOV test with the older group. Bars show mean (± SE) scores for 3 tests (n=18). There was no significant difference in processing speed, divided attention or selective attention between the sessions before and after training (for processing speed, $t(17) = 0.46$, ns; for divided attention, $t(17) = 1.99$, ns; for selective attention, $t(17) = -0.77$, ns).

Supplementary Figure 7



Supplementary Fig. 7. Information regarding voxels above the 0.2 threshold. The mean number of (± SE) voxels for (A) trained, and (B) untrained VOIs, and the mean absolute values (± SE) of FA for (C) trained, and (D) untrained VOIs for young (n=21) and old (n=17) groups.

Supplementary Table 1. Base Model Fixed Effects

	Coefficient	Std. Error	Z value
Intercept	1.1	0.16	6.69
Training Session	0.27	0.09	3.02
SOA	2.67	0.31	8.7
Training Session x SOA	0.34	0.15	2.3
SOA^2	-0.95	0.12	7.74
Training Session x SOA^2	-0.2	0.08	2.34

Supplementary Table 2. Base Model Random Effects

	1	2	3	4	5	6
1. Intercept	0.45	0.3	0.12	-0.16	-0.03	-0.01
2. SOA	0.37	1.52	-0.11	-0.57	-0.04	-0.01
3. Training Session	0.54	-0.26	0.12	0	0.03	-0.06
4. SOA^2	-0.48	-0.96	0.01	0.23	0	0.03
5. SOA x Training Session	-0.07	-0.06	0.17	-0.01	0.25	-0.11
6. Training Session x SOA^2	-0.03	-0.02	-0.59	0.2	-0.73	0.08

Note: For all tables of random effects, variances are provided on the diagonal. Taus are provided above the diagonal. Taus as correlations are provided below the diagonal.

Supplementary Table 3. Area-based Model Fixed Effects

	Coefficient	Std. Error	Z value
Intercept	1.07	0.16	6.6
Training Session	0.24	0.08	3.03
$\Delta_{V3\ Area}$ x Training Session	6.63	3.03	2.19
SOA	2.66	0.31	8.52
$\Delta_{V3\ Area}$ x SOA	2.15	11.97	0.18
Training Session x SOA	0.29	0.13	2.23
$\Delta_{V3\ Area}$ x Training Session x SOA	13.57	4.95	2.74
$\Delta_{V3\ Area}$ x SOA^2	-2.22	4.83	0.46
Training Session x SOA^2	-0.16	0.07	2.46
$\Delta_{V3\ Area}$ x Training Session x SOA^2	-8.58	2.79	3.08

Supplementary Table 4. Area-based Model Random Effects

	1	2	3	4	5	6
1. Intercept	0.43	0.29	0.1	-0.14	-0.07	0.02
2. SOA	0.36	1.52	-0.12	-0.57	-0.05	0
3. Training Session	0.52	-0.31	0.09	0.01	-0.03	-0.02
4. SOA^2	-0.46	-0.96	0.08	0.23	0.02	0.02
5. SOA x Training Session	-0.26	-0.1	-0.23	0.09	0.16	-0.04
6. Training Session x SOA^2	0.14	-0.01	-0.38	0.17	-0.52	0.04

Supplementary Table 5. Time-based Model Fixed Effects

	Coefficient	Std. Error	Z value
Intercept	1.06	0.19	5.61
$\Delta_{V3\ Time}$	5.83	11.81	0.49
Training Session	0.21	0.1	2.1
SOA	2.64	0.35	7.6
SOA^2	-0.93	0.14	6.64
$\Delta_{V3\ Time}$ x Training Session	8.97	6.2	1.45
$\Delta_{V3\ Time}$ x SOA	3.69	21.78	0.17
Training Session x SOA	0.25	0.16	1.61
$\Delta_{V3\ Time}$ x SOA^2	-5.95	8.93	0.67
Training Session x SOA^2	-0.12	0.07	1.66
$\Delta_{V3\ Time}$ x Training Session x SOA	15.94	10	1.59
$\Delta_{V3\ Time}$ x Training Session x SOA^2	-20.67	5.16	4

Supplementary Table 6. Time-based Model Random Effects

	1	2	3	4	5	6
1. Intercept	0.45	0.3	0.13	-0.17	-0.01	-0.06
2. SOA	0.37	1.46	-0.09	-0.55	-0.03	-0.06
3. Training Session	0.6	-0.22	0.11	-0.02	0.02	-0.05
4. SOA^2	-0.52	-0.94	-0.1	0.23	0	0.05
5. SOA x Training Session	-0.05	-0.06	0.11	0.01	0.18	-0.04
6. Training Session x SOA^2	-0.47	-0.24	-0.75	0.52	-0.51	0.04

Supplementary Note 1. Regression analysis for FA changes and performance

In order to evaluate and attempt to characterize the degree to which changes in FA were associated with improvements in task performance for older adults, we analyzed data from training sessions in three binomial linear mixed effect regression models: a base model (see Supplementary Tables 1 and 2 for details), an area-based model, and a time-based model. The area-based model includes the differences in between V3 FA in the trained location and V3 FA in the untrained location as a variable in the model (see Supplementary Tables 3 and 4 for details). The time-based model includes differences between V3 FA before training and V3 FA after training (in the trained locations) as a variable in the model (see Supplementary Tables 5 and 6 for details).

All of the models reported in this appendix predicted accuracy in the TDT judgment as a function of the centered session of training and the SOA of the texture. In the context of these analyses, one can expect to observe improvements in performance in two ways. First, improvement may occur as an overall shift in the probability of a correct answer across training sessions (represented by coefficients for the training session variable). Second, improvement may occur in the extent to which longer SOAs increase the probability of correct answers (represented by coefficients associated with the interaction between the linear SOA variable and training session variable).

For all analyses, SOA was coded in seconds, was mean-centered, and log transformed. As the logistic function was not a completely apt match for our observed data, we included a non-linear term for SOA to clarify the model output. Those coefficients are included here for the sake of completeness, but as they are not relevant to the research question being addressed, they were not interpreted.

The base model demonstrated that there were overall improvements in the TDT task as a function of training session ($b = .27$) and that increases in task performance as a function of SOA improved with training ($b = .34$). The area-based model demonstrated similar patterns of results. However, consistent with mediation, the magnitude of the improvements in performance as a function of training session ($b = .24$) and the improvements in task performance as a function of SOA ($b = .29$) were less than in the base model. Consistent with moderation, the improvements associated with training session were associated with FA ($b = 6.63$); the magnitude of improvements in task performance as a function of SOA with training was also associated with FA (b = 13.57). The time-based model, likewise, consistent with mediation, the magnitude of the improvements in performance as a function of training session (b = .21) and the improvements in task performance as a function of SOA associated with training (b = .25) were less than in the base model. Consistent with moderation, the improvements in task performance associated with the training session were also associated with the FA (b = 8.97). However, the magnitude of improvements in task performance as a function of SOA with training was not reliably associated with FA.

Supplementary Note 2. MD, RD and AXD

We have applied 4-way ANOVA with factors being age (young vs old), VOI (V1, V2, vs V3), location (trained vs untrained) and time (pre- vs post-training) to each of MD, RD, and AXD.

For MD, there were significant main effects of age ($F(1,36) = 7.983$, $p = 0.0077$), and VOI ($F(2,72) = 19.894$, $p<0.0001$), but no significant main effects of location ($F(1,36) = 4.056$, $p = 0.0515$) or time ($F(1,36) = 0.074$, $p = 0.7872$). There was no significant interaction. The overall MD was higher for the older group than the younger group. Neither significant main effect of time nor interaction of time and other factors indicates that the MD does not significantly change in association with perceptual learning in either group.

For RD, there were significant main effects of age ($F(1,36) = 14.969$, $p = 0.0004$), and VOI ($F(2,72) = 4.279$, $p = 0.0175$) but no significant main effect of location ($F(1,36) = 1.211$, $p = 0.2785$) or time ($F(1,36) = 0.162$, $p = 0.6900$). There was no significant interaction. The overall RD was higher for the older group than the younger group. Neither significant main effect of time nor interaction of time and other factors indicates that the RD did not significantly change in association with perceptual learning in either group.

For AXD, there was a significant main effect of VOI ($F(2,72) = 40.522$, $p < 0.0001$), but no significant main effects of age ($F(1,36) = 0.949$, $p = 0.3364$), location ($F(1,36) = 1811$, $p = 0.1868$), and time ($F(1,36) = 0.030$, $p = 0.8626$). Neither significant main effect of time nor interaction of time and other factors indicates that the AXD does not significantly in association with perceptual learning in either group.

Supplementary Note 3. Head motion

The sum of overall corrected motion was defined as the square root of the sum of (displacement-in-the-slice-direction (mm)^^{2} + displacement-in-the-column-direction (mm)^^{2} + displacement-in-the-row-direction (mm)^^{2}) from results of the motion correction procedure. We compared the sum of overall corrected motion between the younger and older groups, before and after the training. In the younger group, the averaged overall corrected motion amounts were 4.485 ± 1.97 mm before training and 4.536 ± 2.36 mm after training, whereas in the older subjects, the amounts were 3.31 ± 1.83 mm before training and 3.72 ± 2.61 mm after training. Two-way repeated measures ANOVA with factors being time (pre vs post) and age (young vs old) showed none of main effects of time ($F(1,37) = 0.492$, $p = 0.49$) and age ($F(1,37) = 2.47$, $p = 0.12$) factors, and no interaction of factors ($F(1,37) = 0.30$, $p = 0.59$).

Supplementary Note 4. Subjects screening

Means and standard deviations of participants' demographic information and results of perceptual and cognitive tests[1] are as follows. In the older group, the mean age was 72.16 ± 5.42 years old, with 17.94 ± 1.98 years of education. In the younger group, the mean age was 22.85 ± 2.94 years old, with 17.05 ± 1.82 years of education. While the age difference between the older and younger groups was significant ($p < 0.01$), the difference in the years of education was not. The Snellen letter acuity was 10/10.82 ± 1.81 in the older, and 10/10.09 ± 0.74 in the younger groups, with no significant group difference. The log contrast sensitivity, measured using the Pelli-Robson test[2], was 1.72 ± 0.18 in the older group, and 1.85 ± 0.1 in the younger group, with no significant group difference. Digit span forward, digit span backward, perceptual encoding manual, and MMSE score were only measured in the older group, and they were 10.94 ± 1.73, 7.06 ± 1.47, 64.67 ± 10.41, 29.28 ± 0.96, respectively.

Supplementary References:

1. Andersen GJ, Ni R, Bower JD, Watanabe T. Perceptual learning, aging, and improved visual performance in early stages of visual processing. *J Vis* **10**, 4 (2010).

2. Pelli DG, Robson JG, Wilkins AJ. The design of a new letter chart for measuring contrast sensitivity. *Clinical Vision Sciences* **2**, 187-199 (1998).