**Supplementary Materials**

**EEG Data Acquisition and Reduction**

EEG data were recorded from a 33 electrode actiCap (Brain Products, GmbH; Munich, Germany) arranged according to the 10/20 system. Electrooculogram activity was recorded from an electrode placed 2 cm next to the left eye and another electrode placed 2 cm below the right eye. Data were recorded using an Electrical Geodesics, Inc. (EGI; Eugene, OR, USA) amplifier system (20,000 gain, bandpass=0.10–100 Hz) with Cz as the online reference. Data were digitized at 500 Hz with a 24-bit analog-to-digital converter and impedances were kept below 20 kΩ throughout recording.

Data were exported to EEGLAB version 14.1.1 (Delorme & Makeig, 2004) for offline analyses. Data were bandpass filtered using a 2nd-order Butterworth filter of 0.10-30 Hz and adjusted for DC offset. All continuous EEG data were visually inspected to identify and remove segments containing large muscle-related artifacts or extreme offsets of activity. Data were then referenced offline to the mean of the mastoids (TP9, TP10). For the doors task, feedback-locked segments were extracted using a -200-800 ms time window, while response-locked segments for the flankers task were extracted using a -400-800 ms time window. Oculomotor and eye blink artifacts were removed from the segmented waveforms using independent component analysis (ICA) blink templates generated by the author of the ERP PCA toolkit version 2.66 (Dien, 2010) and from the segmented data. ICA components that were highly correlated (i.e., *r* ≥ .90) with topographies of the blink templates provided were removed during this step. Segments were rejected if: (1) there was at least a 100 µV voltage difference within a segment; (2) if channels differed by more than 50 µV, which was measured from the neighboring 6 channels; or (3) > 15% of channels were marked bad. Remaining bad channels were corrected through spherical spline interpolation obtained from good channels of the scalp voltage field within each data segment. Segments were averaged separately by trial type for each task (doors: gain, loss; flankers: correct, error) and baseline corrected using the -200-0 ms pre-feedback interval for the doors task and the -400 to -200 ms pre-response interval for the flankers task.

**Supplementary References**

Delorme, A., & Makeig, S. (2004). EEGLAB: an open source toolbox for analysis of single-trial EEG dynamics including independent component analysis. *Journal of Neuroscience Methods*, *134*(1), 9–21.

Dien, J. (2010). The ERP PCA Toolkit: An open source program for advanced statistical analysis of event-related potential data. *Journal of Neuroscience Methods*, *187*(1), 138–145.

Lakens, D. (2013). Calculating and reporting effect sizes to facilitate cumulative science: A practical primer for t-tests and ANOVAs. *Frontiers in Psychology, 4*, 863.

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Table S1** |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| *Pre- and Post-Intervention Measures and Effect Size Estimates by Condition* | | | | | | | | | | | | | | | |
|  | Exercise (*n* = 35) | | | | | | |  | Stretching (*n* = 31) | | | | | | |
|  | Pre | |  | Post | | | |  | Pre | |  | Post | | | |
| Variable | *M* | 95% CI |  | *M* | 95% CI |  | *g*av |  | *M* | 95% CI |  | *M* | 95% CI |  | *g*av |
| BDI-II Score | 22.06 | [19.30, 24.82] |  | 9.83 | [7.01, 12.64] |  | 1.47 |  | 20.35 | [17.77, 22.94] |  | 13.97 | [11.77, 16.17] |  | 0.95 |
| BAI Scorea | 13.73 | [10.22, 17.24] |  | 10.42 | [7.35, 13.49] |  | 0.35 |  | 12.47 | [9.45, 15.48] |  | 11.17 | [7.09, 15.25] |  | 0.13 |
| VO2 peak (ml/kg/min) | 38.06 | [34.66, 41.46] |  | 39.69 | [36.49, 42.88] |  | 0.17 |  | 38.84 | [35.31, 42.36] |  | 39.36 | [36.06, 42.65] |  | 0.05 |
| IPAQ (MET\*min/wk) | 508.71 | [337.47, 679.94] |  | 597.87 | [440.20, 755.54] |  | 0.18 |  | 472.07 | [322.69, 621.46] |  | 480.30 | [311.95, 648.66] |  | 0.02 |
| RewP (μV) | 3.48 | [2.40, 4.56] |  | 3.22 | [2.12, 4.32] |  | 0.08 |  | 2.67 | [1.43, 3.92] |  | 2.48 | [1.32, 3.65] |  | 0.06 |
| Flanker Accuracy (%) | 84.70 | [78.59, 90.80] |  | 85.60 | [79.24, 91.94] |  | 0.05 |  | 86.91 | [80.96, 92.87] |  | 90.98 | [87.32, 94.65] |  | 0.29 |
| Flanker RT (ms) | 641.24 | [581.93, 700.54] |  | 636.70 | [588.90, 684.50] |  | 0.03 |  | 649.45 | [580.86, 718.04] |  | 652.64 | [584.35, 720.93] |  | 0.02 |
| ERN (μV)b | -2.31 | [-3.45, -1.17] |  | -3.23 | [-5.14, -1.31] |  | 0.21 |  | -3.78 | [-5.44, -2.12] |  | -5.22 | [-7.62, -2.82] |  | 0.29 |
| *Note.* CI = confidence interval;  *g*av = Hedges’s *g* adjusted effect size (see Lakens, 2013 for its computation); BDI-II = Beck Depression Inventory, Second Edition; BAI = Beck Anxiety Inventory; IPAQ = International Physical Activity Questionnaire; MET min/wk = metabolic equivalent minutes per week; RewP = reward positivity; μV = microvolts; % = percentage correct; RT = reaction time; ms = milliseconds; ERN = error-related negativity. | | | | | | | | | | | | | | | |
| a BAI Score Analyses: *N* = 63 (Exercise: *n* = 33; Stretching: *n* = 30). | | | | | | | | | | | | | | | |
| b ERN Analyses: *N* = 55 (Exercise: *n* = 31; Stretching: *n* = 24). | | | | | | | | | | | | | | | |

|  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Table S2** |  |  |  |  |  |  |  |  |  |  |  |
| *Means and 95% Confidence Intervals of Number of Trials Contributing to Each ERP Waveform by Condition Across the Intervention* | | | | | | | | | | | |
|  | Exercise | | | | |  | Stretching | | | | |
|  | Pre | |  | Post | |  | Pre | |  | Post | |
| Variable | *M* | 95% CI |  | *M* | 95% CI |  | *M* | 95% CI |  | *M* | 95% CI |
| *Doors ERPs* |  |  |  |  |  |  |  |  |  |  |  |
| ERP to Gains | 46.8 | [45.0, 48.7] |  | 43.9 | [41.3, 46.4] |  | 46.5 | [44.1, 48.9] |  | 44.3 | [40.9, 47.7] |
| ERP to Losses | 46.7 | [44.6, 48.8] |  | 43.8 | [41.2, 46.4] |  | 46.6 | [44.4, 48.7] |  | 43.7 | [40.4, 47.1] |
| RewP | 93.5 | [89.6, 97.4] |  | 87.7 | [82.5, 92.8] |  | 93.0 | [88.6, 97.5] |  | 88.0 | [81.3, 94.7] |
| *Flankers ERPs* |  |  |  |  |  |  |  |  |  |  |  |
| ERP to Errors | 37.6 | [23.4, 51.9] |  | 31.5 | [18.1, 44.8] |  | 23.3 | [14.9, 31.8] |  | 17.5 | [12.4, 22.7] |
| ERP to Correct Trials | 187.3 | [170.7, 203.8] |  | 180.2 | [160.1, 200.3] |  | 201.4 | [189.4, 213.4] |  | 208.9 | [198.5, 219.3] |
| ERN | 224.9 | [217.6, 232.2] |  | 211.7 | [194.5, 228.9] |  | 224.8 | [215.1, 234.4] |  | 226.4 | [216.0, 236.9] |
| *Note.* CI = confidence interval; RewP = reward positivity, calculated as ERP to Rewards minus ERP to Losses; ERN = error-related negativity, calculated as ERP to Errors minus ERP to Correct Trials. | | | | | | | | | | | |

|  |  |  |  |
| --- | --- | --- | --- |
|  |  |  |  |
| **Table S3** |  |  |  |
| *Pre- and Post-Intervention Reliability Estimates* | | | |
|  | Pre |  | Post |
| Measure | Reliability |  | Reliability |
| *Self-Reports* |  |  |  |
| BDI-II | .771 |  | .873 |
| BAI | .893 |  | .922 |
| *ERPs* |  |  |  |
| ERP to Rewards | .978 |  | .969 |
| ERP to Losses | .916 |  | .923 |
| RewP | .705 |  | .672 |
| ERP to Errors | .797 |  | .967 |
| ERP to Correct Trials | .989 |  | .997 |
| ERN | .716 |  | .610 |
| *Note.* Self-report reliability estimates are calculated as Cronbach's alpha; ERP reliability estimates are calculated as the Spearman-Brown adjusted coefficient. | | | |

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **Table S4**  *Multilevel Model for Intervention Effects on Depressive Symptom Severity* | | | | | | |
|  | *b* | SE | *t* | df | *p* | 95% CI |
| *Depressive Symptom Severitya* |  |  |  |  |  |  |
| Intercept | 20.17 | 1.22 | 16.52 | 104.41 | <.001 | [17.79, 22.55] |
| Time | -1.65 | 0.33 | -5.08 | 141.72 | <.001 | [-2.29, -1.01] |
| Treatment Group | 0.65 | 1.68 | 0.39 | 105.68 | .699 | [-2.62, 3.94] |
| Time x Treatment Group | -1.23 | 0.45 | -2.74 | 141.76 | .007 | [-2.11, -0.35] |

*Note.* *b* = unstandardized regression coefficient; df = Satterthwaite-

approximated degrees of freedom; ICC = intraclass correlation coefficient of the

unconditional model.

a ICC = .25.

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| **Table S5** | | | | | | | |
| *Predictive Accuracy of Baseline RewP Predicting Responder Status Across the Whole Sample* | | | | | | | |
| Baseline RewP (Whole Sample) | | | | | | | |
| AUC | Cutoff SD Values | Cutoff RewP (µV) Values | Sens | Spec | PPV | NPV | Acc |
| 0.66 | -2.0 | -3.42 | 1.00 | 0.02 | 0.50 | 0.43 | 0.50 |
|  | -1.5 | -1.79 | 0.96 | 0.07 | 0.50 | 0.53 | 0.50 |
|  | -1.0 | -0.16 | 0.96 | 0.12 | 0.51 | 0.59 | 0.52 |
|  | -0.5 | 1.47 | 0.76 | 0.34 | 0.55 | 0.62 | 0.57 |
|  | 0 | 3.10 | 0.60 | 0.68 | 0.61 | 0.62 | 0.61 |
|  | +0.5 | 4.73 | 0.40 | 0.85 | 0.70 | 0.60 | 0.63 |
|  | +1.0 | 6.36 | 0.28 | 0.90 | 0.81 | 0.57 | 0.61 |
|  | +1.5 | 7.99 | 0.20 | 0.98 | 0.89 | 0.54 | 0.57 |
|  | +2.0 | 9.62 | 0.10 | 1.00 | 0.96 | 0.52 | 0.54 |
| *Note.* AUC = area under the curve; Sens = sensitivity; Spec = specificity; PPV = positive predictive value; NPV = negative predictive value;Acc = accuracy; *n* = 66. | | | | | | | |

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| **Table S6** | | | | | | | |
| *Predictive Accuracy of Baseline RewP Predicting Responder Status Among Individuals Assigned to the Exercise Condition* | | | | | | | |
| Baseline RewP (Exercise Condition) | | | | | | | |
| AUC | Cutoff SD Values | Cutoff RewP (µV) Values | Sens | Spec | PPV | NPV | Acc |
| 0.70 | -2.0 | -2.84 | 0.98 | 0.01 | 0.49 | 0.09 | 0.49 |
|  | -1.5 | -1.26 | 0.94 | 0.02 | 0.49 | 0.28 | 0.48 |
|  | -1.0 | 0.32 | 0.87 | 0.15 | 0.50 | 0.52 | 0.51 |
|  | -0.5 | 1.90 | 0.76 | 0.41 | 0.56 | 0.63 | 0.59 |
|  | 0 | 3.48 | 0.61 | 0.75 | 0.71 | 0.66 | 0.68 |
|  | +0.5 | 5.06 | 0.44 | 0.94 | 0.88 | 0.62 | 0.69 |
|  | +1.0 | 6.64 | 0.29 | 0.99 | 0.97 | 0.58 | 0.64 |
|  | +1.5 | 8.22 | 0.16 | 1.00 | 1.00 | 0.54 | 0.58 |
|  | +2.0 | 9.80 | 0.08 | 1.00 | 1.00 | 0.52 | 0.54 |
| *Note.* AUC = area under the curve;Sens = sensitivity; Spec = specificity; PPV = positive predictive value; NPV = negative predictive value; Acc = accuracy; *n* = 35. | | | | | | | |

**Figure S1**

*Timeline of Study Procedures*



*Note.* Eligible participants were randomized into either eight-weeks of aerobic exercise or stretching. BDI-II = Beck Depression Inventory, Second Edition; BAI = Beck Anxiety Inventory; IPAQ = International Physical Activity Questionnaire; ERP = event-related potential.

**Figure S2**

*CONSORT Study Flow Diagram*

Analyzed (n=35)  
 Excluded from ERN analysis due to not enough usable EEG data (n=4)

## Analysis

Analyzed (n=31)  
 Excluded from ERN analysis due to not enough usable EEG data (n=7)

## Follow-Up

## Allocation

Lost to follow-up (n=0)

Discontinued intervention (n=6)

Lost to follow-up (n=0)

Discontinued intervention (n=9)

Assessed for eligibility (n=81)

Excluded (n=15)

  Not meeting inclusion criteria (n=12)

  Declined to participate due to time (n=3)

Randomized (n=66)

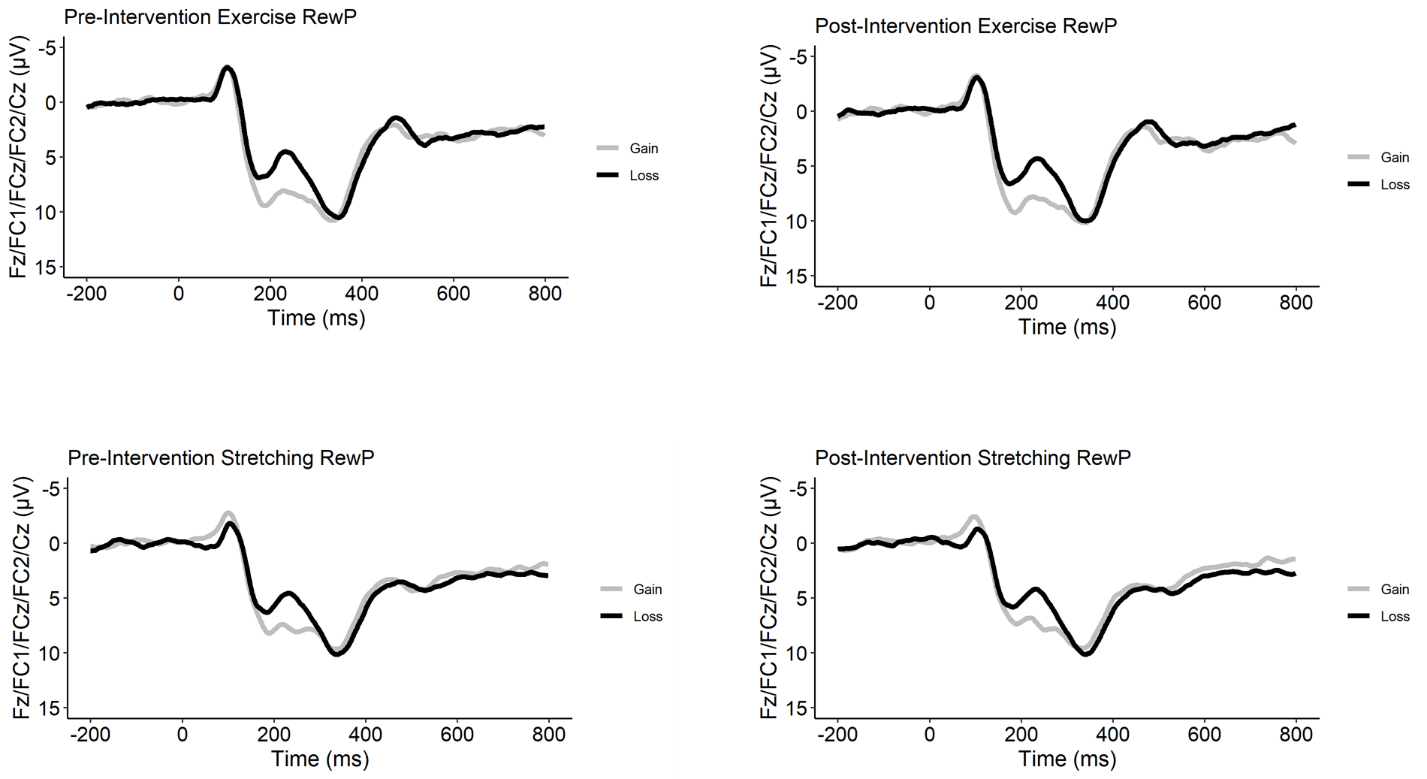
## Enrollment

Moderate-Intensity Aerobic Exercise (n=35)

Light-Intensity Stretching (n=31)

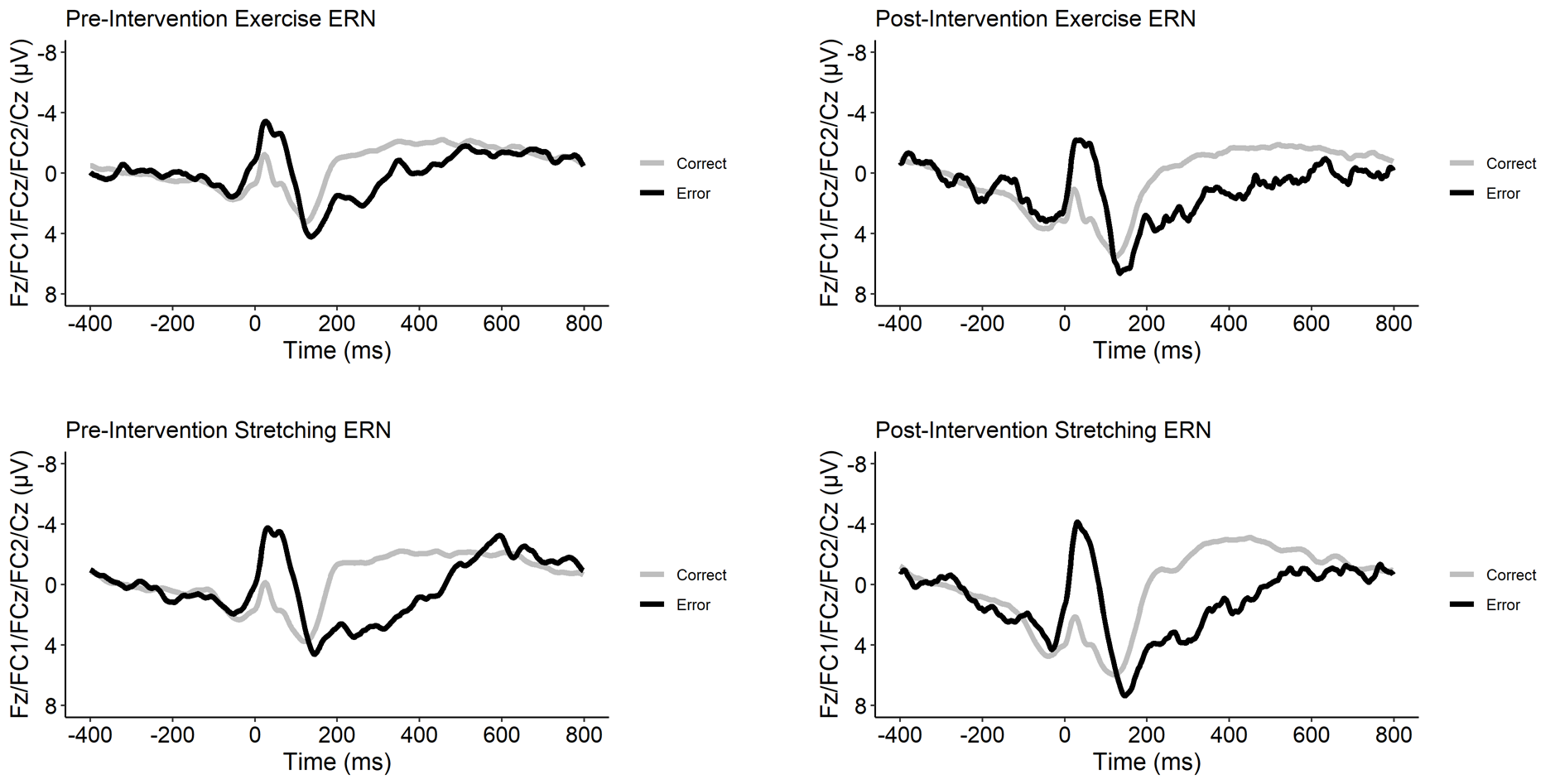
**Figure S3**

*Grand-Averaged RewP ERP Parent Waveforms Across a Frontocentral Region-of-Interest (Fz, FC1, FCz, FC2, Cz) by Treatment (Exercise, top; Stretching, bottom) Before and After the Intervention*



**Figure S4**

*Grand-Averaged ERN ERP Parent Waveforms Across a Frontocentral Region-of-Interest (Fz, FC1, FCz, FC2, Cz) by Treatment (Exercise, top; Stretching, bottom) Before and After the Intervention*

**