**Additional file 3.** Mann-Whitney U tests of the group differences between the teleology intervention and control groups are described below. Wilcoxon tests assessed within-group differences. Mean (SD) and effect size (Hedges’ g) are reported.

|  |  |  |  |
| --- | --- | --- | --- |
|  | Teleology Intervention  (n=48) | Control  (n=26) | Effect size |
| Teleological reasoning |  |  |  |
| Pre | 3.0 (0.80) | 3.31 (0.65) | 0.42 |
| Post | 2.00 (0.87) | 3.20 (0.81) | 1.43‡ |
| Effect size | -1.20‡ | -0.15 |  |
| Natural selection understanding |  |  |  |
| Pre | 9.00 (2.92) | 7.79 (2.51) | 0.44 |
| Post | 10.27 (2.35) | 8.81 (2.73) | 0.59\* |
| Effect size | 0.48‡ | 0.39 |  |
| I-SEA-total |  |  |  |
| Pre | 100.50 (15.55) | 98.52 (12.93) | -0.42 |
| Post | 106.70 (13.75) | 98.42 (14.95) | -0.58\* |
| Effect size | 0.57‡ | -0.0072 |  |
| I-SEA-Human Evolution |  |  |  |
| Pre | 32.65 (6.63) | 32.03 (5.61) | -0.098 |
| Post | 35.40 (5.38) | 31.92 (6.25) | -0.61\* |
| Effect size | 0.46‡ | -0.019 |  |

\**p* < 0.05; †*p* ≤ 0.01; ‡*p* ≤ 0.001