**Supplementary Materials**

*RT analysis*



*Summary Tables*

*Exp 1*

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Repetition | #1 | #2 | #3 | #4 | #5 |
| Mean | 0.58 | 0.75 | 0.81 | 0.85 | 0.88 |
| SD | 0.19 | 0.18 | 0.18 | 0.18 | 0.17 |
| Skewness | -0.25 | -0.98 | -1.70 | -1.70 | -2.09 |
| Kurtosis | -0.44 | 0.59 | 1.82 | 2.60 | 4.31 |
| Reliability | 0.92 | 0.93 | 0.93 | 0.95 | 0.93 |
| Correlationswith VWM  | 0.26 | 0.35 | 0.36 | 0.35 | 0.37 |

*Exp 2*

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| Repetition | #1 | #2 | #3 | #4 | #5 | #6 |
| Mean | 0.63 | 0.70 | 0.79 | 0.85 | 0.88 | 0.91 |
| SD | 0.20 | 0.19 | 0.17 | 0.14 | 0.15 | 0.15 |
| Skewness | -0.59 | -0.68 | -0.94 | -1.18 | -2.42 | -3.11 |
| Kurtosis | 0.20 | 0.02 | 0.44 | 0.80 | 6.82 | 11.26 |
| Reliability | 0.996 | 0.997 | 0.997 | 0.997 | 0.998 | 0.998 |
| Correlationswith VWM  | 0.34 | 0.32 | 0.28 | 0.29 | 0.29 | 0.29 |

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| Repetition | #7 | #8 | #9 | #10 | #11 | #12 |
| Mean | 0.93 | 0.95 | 0.96 | 0.96 | 0.96 | 0.96 |
| SD | 0.13 | 0.11 | 0.09 | 0.09 | 0.12 | 0.12 |
| Skewness | -3.65 | -3.96 | -4.99 | -6.35 | -5.38 | -4.89 |
| Kurtosis | 14.74 | 19.36 | 32.29 | 48.78 | 30.44 | 24.27 |
| Reliability | 0.997 | 0.997 | 0.995 | 0.993 | 0.990 | 0.976 |
| Correlationswith VWM  | 0.24 | 0.23 | 0.20 | 0.13 | 0.07 | 0.09 |

*Exp 3*

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| Repetition | #1 | #2 | #3 | #4 | #5 | #6 |
| Mean | 0.58 | 0.64 | 0.74 | 0.81 | 0.85 | 0.88 |
| SD | 0.17 | 0.17 | 0.18 | 0.17 | 0.17 | 0.17 |
| Skewness | -0.19 | -0.14 | -0.81 | -1.17 | -1.87 | -2.08 |
| Kurtosis | -0.61 | -0.6 | 0.13 | 0.86 | 3.70 | 4.11 |
| Reliability | 0.990 | 0.990 | 0.989 | 0.987 | 0.987 | 0.987 |
| Correlationswith VWM  | 0.28 | 0.26 | 0.33 | 0.35 | 0.34 | 0.31 |

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| Repetition | #7 | #8 | #9 | #10 | #11 | #12 |
| Mean | 0.89 | 0.91 | 0.92 | 0.92 | 0.92 | 0.93 |
| SD | 0.15 | 0.15 | 0.14 | 0.15 | 0.14 | 0.14 |
| Skewness | -2.11 | -2.45 | -2.99 | -3.04 | -3.14 | -3.06 |
| Kurtosis | 4.56 | 5.94 | 9.84 | 9.85 | 11.19 | 9.93 |
| Reliability | 0.986 | 0.985 | 0.981 | 0.977 | 0.973 | 0.951 |
| Correlationswith VWM  | 0.32 | 0.29 | 0.28 | 0.27 | 0.27 | 0.24 |

*Exp 4*

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Repetition | #1 | #2 | #3 | #4 |
| Mean | 0.29 | 0.47 | 0.63 | 0.74 |
| SD | 0.24 | 0.25 | 0.25 | 0.23 |
| Skewness | 1.4 | 0.44 | -0.25 | -0.68 |
| Kurtosis | 1.25 | -0.71 | -1.00 | -0.56 |
| Reliability | 0.99 | 0.99 | 0.98 | 0.95 |
| Correlationswith VWM  | 0.14 | 0.21 | 0.25 | 0.29 |

*Exp 5*

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Repetition | #1 | #2 | #3 | #4 | #5 |
| Mean | 0.56 | 0.77 | 0.83 | 0.87 | 0.90 |
| SD | 0.18 | 0.17 | 0.16 | 0.15 | 0.13 |
| Skewness | 0.06 | -0.71 | -1.17 | -1.64 | -2.23 |
| Kurtosis | -0.96 | -0.18 | 1.00 | 2.38 | 6.19 |
| Reliability | 0.93 | 0.91 | 0.92 | 0.91 | 0.85 |
| Correlations with change localization  | 0.35 | 0.42 | 0.44 | 0.37 | 0.41 |
| Correlations with filtering change localization | 0.26 | 0.34 | 0.34 | 0.36 | 0.34 |
| Correlations with Simon Square | 0.26 | 0.31 | 0.29 | 0.26 | 0.25 |
| Correlations with Flanker Square | 0.29 | 0.30 | 0.34 | 0.32 | 0.33 |

**Open science statement** The data, code, and materials for all experiments will be publicly accessible on OSF upon publication. There is not a preregistration for our studies here.

*Power Estimation*

A power analysis was conducted using InteractionPoweR (Baranger et al., 2023) to determine the minimum sample size required to test the study hypothesis (i.e., aptitude x treatment interaction effect). We ran 1000 simulations for each of the power estimation hyperparameters and assumed that our working memory measures and long-term memory measures were both reliable (reliability of 0.8 and 0.9, respectively, according to Zhao & Vogel., 2024). Our plot above showed that with a small effect (interaction *r* = 0.2) and α = .05, the range of sample sizes used in our study (N = 100, 150, 200, 250, 300, and 700) all produced a robust power (>80%).