Supplementary Information

1 Heteroscedasticity of learning modelling

When normal session scores were fit without log transformation, *heteroscedasticity* occurred, leading to a violation of model assumptions: that is, the error distribution of the model was not independent of the predicted score, Figure 1-left. Figure 1-right shows the data after log transform.

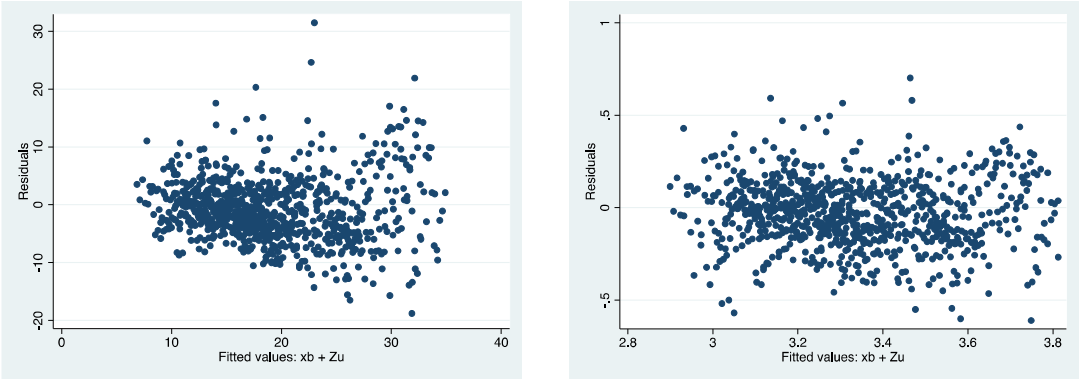

Figure SI.1. ***Left panel****: Error term as a function of the fitted value of the model when using plain normal session score as the outcome in a linear mixed model.* ***Right panel****: Error term as a function of the fitted value of the model when using of the logarithm of the normal session score as the outcome in a linear mixed model.*

2 Tabulated Results

Table SI.1. Linear mixed model of the interactions between DES score and the number of training sessions with the logarithm of normal SMR session scores as the outcome.

Predictor	*β*	*z*	*p* (Wald)	95% CI	
Session number	0.12	5.0	<.001	.08	.17
DES	0.06	3.3	.001	.02	.09
Session number x DES	-0.001	-2.3	.02	-.002	-.0002
Intercept	10.5	11.3	<.001	8.64	12.3

	Estimate	*se*	95 % CI	
Variance component	2.86	1.20	1.26	6.5
Residual	15.7	.95	14.0	17.7
Log likelihood	1589.92			
Observations	563			
Groups (patients)	15			

Table SI.2. Linear mixed model of the interactions between DES score and the number of training sessions with the logarithm of normal TBR session scores as the outcome.

	β	*z*	*p* (Wald)	95% CI	
Session number	0.43	6.0	<.001	.29	.57
DES	0.10	1.2	.24	-.07	.27
Session number x DES	-.005	-2.8	.005	-.009	-.002
Intercept	16.0	5.1	<.001	10.1	22.6

	Estimate	se	95 % CI	
Variance component	16.0	8.72	5.52	46.6
Residual	52.0	4.32	44.2	61.2
Log likelihood	-1021.6574			
Observations	298			
Groups (patients)	8			

To compare the outcome variables between learners, non-learners, and wait-list control group, we performed one-way ANOVA tests with independent variable Group and dependent variable corresponding to the Time 2 minus Time 1 delta of each variable of interest. For TOVA variables, Time 1 was the intake test and time 2 was the outtake test. For ASRS variables, Time 1 was pre-training; for every treatment subject Time 2 was ASRS at session 30 (as in the main text), while for wait list subjects Time 2 was the outtake test. We used Tukey's Honest Significant Differences post-hoc contrasts to examine pairwise differences for ANOVA tests with significant results. Results are summarised below in Table SI.3

Table SI.3. ANOVA models of the group differences between learners, non-learners, and wait list controls for TOVA and ASRS score deltas between pre- and post-treatment. In column headers, … In-table acronyms include: Lrn = Learner, Non = non-Learner, W.L. = wait list.

DV	df	F	p	Contrast	difference	95% CI	p.adj
D’	2	3.87	**0.03 ***	Non - Lrn	1.77	0.22 – 3.32	**0.02 ***
				W.L. - Lrn	0.89	-0.48 – 2.27	0.26
				W.L. - Non	-0.87	-2.25 – 0.50	0.28
Omission Error	2	5.28	**0.009 ****	Non - Lrn	-29.36	-53.7 – -5.02	**0.01 ***
				W.L. - Lrn	-24.82	-46.5 – -3.19	**0.02 ***
				W.L. - Non	4.54	-17.1 – 26.17	0.87
Commission Error	2	0.07	0.9	Non - Lrn	0.27	-7.57 – 8.12	0.99
				W.L. - Lrn	1.01	-5.96 – 7.98	0.93
				W.L. - Non	0.74	-6.23 – 7.71	0.96
Response Time	2	1.13	0.3	Non - Lrn	-28.96	-75.92 – 17.99	0.30
				W.L. - Lrn	-13.52	-55.24 – 28.19	0.71
				W.L. - Non	15.44	-26.28 – 57.16	0.64
Response Time Variability	2	4.73	**0.01 ***	Non - Lrn	-46.32	-83.27 – -9.37	**0.01 ***
				W.L. - Lrn	-19.33	-52.16 – 13.49	0.33
				W.L. - Non	26.99	-5.84 – 59.82	0.12
ASRS total	2	3.44	**0.04 ***	Non - Lrn	-1.10	-4.82 – 2.62	0.75
				W.L. - Lrn	2.26	-0.96 – 5.49	0.21
				W.L. - Non	3.36	0.03 – 6.69	**0.04 ***
ASRS inattention	2	4.31	**0.02 ***	Non - Lrn	-1.37	-3.28 – 0.53	0.19
				W.L. - Lrn	0.67	-0.98 – 2.37	0.58
				W.L. - Non	2.05	0.34 – 3.75	**0.01 ***
ASRS hyperactivity / impusivity	2	2.25	0.12	Non - Lrn	0.37	-1.91 – 2.65	0.92
				W.L. - Lrn	1.59	-0.39 – 3.57	0.14
				W.L. - Non	1.22	-0.82 – 3.26	0.32

3 Continuous Time Structural Equations Modeling

The Continuous Time Structural Equations Modeling (CTSEM) approach is a stochastic extension of the standard cross-lagged model:

$$(\mathbf{x}(t_i) - \mathbf{x}(t_{i-q}))/\Delta_i = \frac{(A(\Delta_i)-1)}{\Delta_i}\mathbf{x}(t_{i-1}) + \mathbf{w}(\Delta_i)/\Delta_i = A_*\mathbf{x}_{t_i} + \mathbf{w}(\Delta_i)/\Delta_i.$$

After replacing the discrete time intervals (Δ_i) with continuous time variable *t,* it can be described by a stochastic difference equation as described by Voelke et al (2012):

$$\frac{d(t)}{dt} = A_*(t) + G\frac{dW(t)}{dt} + MX(t),$$

where G is the Cholesky triangle of the diffusion matrix. This has a solution based on the stochastic integral

$$\mathbf{x}(t) = e^{A_*(t-t_0)}\mathbf{x}(t_0) + \int_{t_0}^{t} e^{A\cdot(t-s)}GdW(s) + \int_{t_0}^{t} e^{A\cdot(t-s)}MX(s)ds.$$

A stochastic integral is needed because *W* is a Wiener process and does not have a continuous derivative. The latter term in the solution is thus an integral of the diffusion process and it is possible to calculate its covariance matrix explicitly (Driver, Oud, & Voelkle, 2015: 4).

The model allows the separation of latent and measure models, but they were assumed to agree in this study because the data is produced in a relatively objective fashion (i.e. EEG scores, no questionnaire data, etc.). However, resolving the SEM with data consisting of several subjects, the individual level variance and the between person differences (the so-called *traits*) were separately estimated (cf. Oud & Jansen, 2000).

The *level change* due to individual training sessions was modeled by variables $X_{\text{norm, cum}}$ and $X_{\text{inv, cum}}$ that, for each session, stand for the cumulative number of the training blocks of a given type (normal or inverse) from the beginning of NF training until the end of the particular session. The *impact* associated with individual training sessions, by

contrast, was measured by a variable $X_{\text{norm, imp}}$ and $X_{\text{inv, imp}}$ standing for the number of the training blocks in a given training day (i.e. session).

The results based on the CTSEM approach were not subject to statistical testing so even if they apply to data, caution should be taken before seeking to generalise the results. In the context of this stucy, CTSEM methods should rather be viewed as being exploratory, providing addition insight invisible by more elementary statistical methods and as a way of generating hypotheses to be tested in future.