Electronic Supplementary Material 1

Appendix A

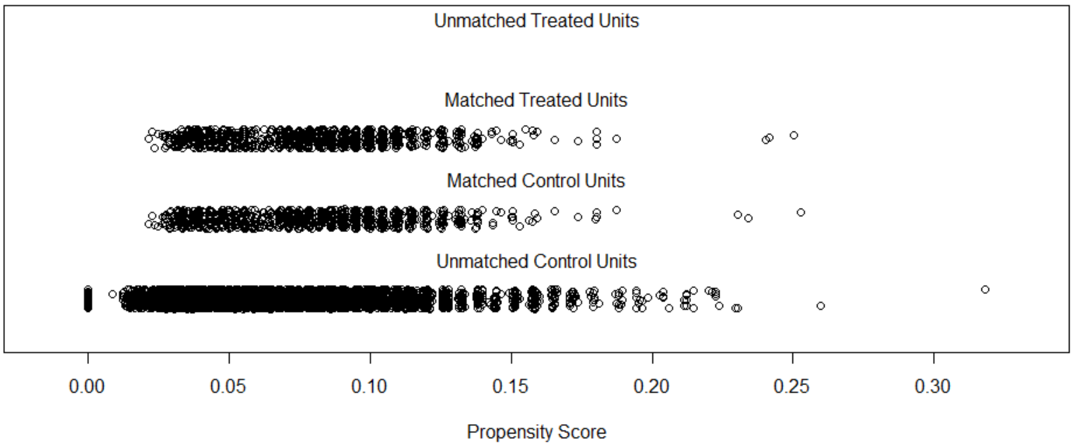

Figure A. Distribution of prospensity scores.

Compagnoni, M., Rechsteiner, B., Grob, U., Bayer, N., Wullschleger, A. & Maag Merki, K. (2023). No Loss, No Gain? COVID-19 school closures and Swiss fifth-graders' competencies and self-concept in mathematics. *Zeitschrift für Pädagogische Psychologie*.

https://doi.org/10.1024/1010-0652/a000366

Appendix B

Table B

Math competencies in the main sample (2019/20) and the reference sample (2018/19) with data before matching

	Main sample (19/20)			Reference sample (18/19)						
	n	*M*	*SD*	*n*	*M*	*SD*	d_{cohen}	*t*	*df*	*p*
Total	1193	-.688	1.246	13030	-.809	1.267	.096	3.177	14221	.000

Note. To compute d_{cohen}, differences in group size were included; *p* = two-tailed.

https://doi.org/10.1024/1010-0652/a000366

Appendix C

Table C

Intra-class correlations for math self-concepts and math competencies

	Variance components		ICC (b/b+w)				
	Variance within (w)	Variance between (b)	Estimate	*SD*	*p*-value	Lower 2.5%	Upper 2.5%
Self-concept t1	.276	.009	.032	.016	.000	.011	.075
Self-concept t2	.338	.010	.030	.017	.000	.008	.073
Self-concept t3	.317	.011	.035	.021	.000	.006	.084
Math competencies t1	1.568	.214	.117	.024	.000	.076	.172
Math competencies t3	1.494	.172	.102	.027	.000	.070	.177

Note. Self-concepts assessed as latent variables, math competencies as manifest variables.

https://doi.org/10.1024/1010-0652/a000366

Appendix D

A prerequisite for analyzing constructs over different time points is to confirm measurement invariance. Confirmatory factor analyses with cluster-robust standard errors to control for multilevel data structure were conducted for math self-concept (see Table C for model indices). Assuming that identically formulated items correlated across the three measurement time points, generally in addition to the stability of the shared variance of the factors, and more strongly for shorter periods of time, we imposed model constraints so that the error correlations of closer measurement time points were equaled for each indicator. As $\Delta\chi2$ is sensitive to sample size > 100, both $\Delta\chi2$ and ΔCFI were used as fit indices to compare model fit (Cheung & Rensvold, 2002). In a first model for all three measurement time points, the factors were configured from the same variables (configural measurement invariance). In the second model we imposed weak factorial measurement invariance by constraining the factor loadings to be invariant over time. Finding that weak factorial invariance held, we proceeded to test the third model (strong factorial invariance) in which we additionally constrained the intercepts to be invariant over time. In addition, we freed up the latent means from t2 onward to allow for mean change over time. This more restrictive model had a comparable model fit to model 1. Consequently, strong measurement invariance could be assumed for math self-concept. We performed the same stepwise invariance test for the subgroups (first language). Based on Cheung & Rensvold (2002), the ΔCFI of -.006 could be judged as not meaningful and strong factorial measurement invariance was assumed.

Cheung, G. W., & Rensvold, R. B. (2002). Evaluating goodness-of-fit indexes for testing measurement invariance. *Structural Equation Modeling*, *9*(2), 233–255.

https://doi.org/10.1024/1010-0652/a000366

Table D

Measurement invariance (MI) of math self-concept over time and subgroups

Models	χ^2	*df*	SCF	*p*	RMSEA	CFI	SRMR	S-B $\Delta\chi^2$	*p*
Math self-concept (n = 1,299)									
Configural MI	23.297	18	1.141	.179	.015	.999	.012		
Weak factorial MI	27.857	22	1.125	.181	.014	.999	.022	4.518	.340
Strong factorial MI	30.516	26	1.107	.247	.012	.999	.021	6.986	.538
Model with subgroups: German as first language (n = 857), not German as first language (n = 407)									
Configural MI	47.146	42	1.115	.270	.014	.999	.018		
Weak factorial MI (group)	55.328	51	1.107	.315	.012	.999	.049	8.115	.523
Weak factorial MI (group and time)	67.358	57	1.105	.164	.017	.998	.067	2.300	.161
Strong factorial MI (group)	74.646	64	1.089	.171	.016	.998	.070	27.634	.188
Strong factorial MI (group and time)	103.574	70	1.087	.006	.028	.992	.068	57.433	.001

Note. Type = complex, cluster = class, SCF = scaling factor, χ^2 = chi-square, RMSEA = root mean square error of approximation, CFI = comparative fit index, SRMR = standardized root mean square residual, S-B $\Delta\chi^2$ = Sartorra-Bentler scaled delta chi-square.

https://doi.org/10.1024/1010-0652/a000366

Appendix E

M-Plus Syntax for RI-CLPM

```
TITLE:        RI-CLPM (basic model math competencies and math self-concepts)
DATA:         FILE IS Covid.dat;
              !LISTWISE = ON
VARIABLE:     NAMES ARE[…];
              USEVARIABLES ARE
              t1_MSC_1      t1_MSC_2      t1_MSC_3
              t2_MSC_1      t2_MSC_2      t2_MSC_3
              t3_MSC_1      t3_MSC_2      t3_MSC_3
              t1_Math       t3_Math;
              CLUSTER = class_id;
              MISSING = all (-99)
 ANALYSIS:    type = complex;
              MODEL = NOCOV;
              *Auxiliary = (m) urbanity migratio eduparen firstlan gender age;
MODEL:        Mot1 by   t1_MSC_1      t1_MSC_2      t1_MSC_3 (L1 - L3);
              Mot2 by   t2_MSC_1      t2_MSC_2      t2_MSC_3 (L1 - L3);
              Mot3 by   t3_MSC_1      t3_MSC_2      t3_MSC_3 (L1 - L3);
              [t1_MSC_1     t1_MSC_2      t1_MSC_3] (i1-i3);
              [t2_MSC_1     t2_MSC_2      t2_MSC_3] (i1-i3);
              [t3_MSC_1     t3_MSC_2      t3_MSC_3] (i1-i3);
              [Mot2* Mot3*];
              t1_MSC_1 with t2_MSC_1 (a1);
              t1_MSC_2 with t2_MSC_2 (a2);
              t1_MSC_3 with t2_MSC_3 (a3);
              t2_MSC_1 with t3_MSC_1 (b1);
              t2_MSC_2 with t3_MSC_2 (b2);
              t2_MSC_3 with t3_MSC_3 (b3);
              t1_MSC_1 with t3_MSC_1 (c1);
              t1_MSC_2 with t3_MSC_2 (c2);
              t1_MSC_3 with t3_MSC_3 (c3);
              RIMot by Mot1@1 Mot2@1 Mot3@1;
              WMot1 by Mot1@1;
              WMot2 by Mot2@1;
              WMot3 by Mot3@1;
              Mot1@0;
              Mot2@0;
              Mot3@0;
              WMot2 on WMot1 (n);
              Wmot2 on  t1_Math;
              WMot3 on WMot2 (m);
              t3_Math on  t1_Math WMot2;
              RIMot with  t3_Math  t1_Math;
              t1_Math with WMot1;
              t3_Math with WMot3;
              RIMot with WMot1@0;

              Model Constraints:
              a1 = b1;
              a2 = b2;
              a3 = b3;
              !n = m;
OUTPUT:       TECH1 TECH4 STDYX SAMPSTAT CINTERVAL;
```

https://doi.org/10.1024/1010-0652/a000366

Appendix F

Table F

Standardized parameter estimates of RI-CLPM for math competencies and math self-concept including auxiliary variables

		CI 95 %				Two-tailed
	EST	lower	upper	*SE*	EST/*SE*	*p*-value
Lagged effect State SC t1 - t2	.053	-0.29	.396	.175	.301	.763
Lagged effect State SC t2 - t3	.413	.211	.615	.103	4.004	.000
Lagged effect Math t1 - t3	.684	.648	.721	.019	36.721	.000
Cross-lagged effect State SC t2 on Math t1	.158	-.008	.324	.085	1.861	.063
Cross-lagged effect Math t3 on State SC t2	.081	.002	.160	.040	2.002	.045
Correlation Trait SC with Math t3	.225	.155	.295	.036	6.282	.000
Correlation Trait SC with Math t1	.508	.418	.598	.046	11.052	.000
Cross-sectional correlation State SC t1 with Math	-.118	-.299	.064	.093	-1.270	.204
Cross-sectional correlation State SC t3 with Math	.123	.034	.212	.045	2.704	.007

Note. N = 1,299; SC = math self-concept, Math = math competencies, Model fit: $\chi2(38) = 55.925$, $p < .031$, RMSEA = .019, CFI = .997, SRMR = .023.

https://doi.org/10.1024/1010-0652/a000366

Appendix G

Table G

Multiple group RI-CLPM to test for differential effects based on children's first language

Models	χ^2	*df*	SCF	*p*	RMSEA	CFI	SRMR	str. pos. S-B $\Delta\chi^2$	*p*
German first language (n = 857*)* *Not German first language (n* = 407*)*									
Multiple group RI-CLPM 1	138.774	92	0.997	.001	.028	.992	.042		
Multiple-group RI-CLPM 2 (constrained [a])	134.997	101	1.061	.014	.023	.994	.046	3.231	.954

Notes. N = 1,264, str. pos. S-B $\Delta\chi^2$ = strictly positive Satorra-Bentler Chi2 Test (Asparouhov & Muthén, 2010); [a] all effects invariant across groups.

Asparouhov, T., & Muthén, B. (2010). Computing the strictly positive Satorra-Bentler chi-square test in Mplus. *Mplus Web Notes*, *12*, 1-12.

https://doi.org/10.1024/1010-0652/a000366

Appendix H

Table H

Standardized lagged, cross-lagged paths, and cross–sectional correlations for RI-CLPM with first language, gender, and age as covariates

		CI 95 %				Two-tailed
Parameter	EST	low	up	*SE*	EST/*SE*	*p*-value
Lagged effect State SC t1 - t2	.095	-.310	.499	.206	0.459	.646
Lagged effect State SC t2 - t3	.434	.209	.660	.115	3.773	.000
Lagged effect Math t1 - t3	.664	.627	.701	.019	34.150	.000
Cross-lagged effect Math t1 - State SC t2	.134	-.036	.305	.087	1.548	.122
Cross-lagged effect State SC t2 - Math t3	.074	-.008	.156	.042	1.765	.076
Correlation Trait SC with Math t3	.228	-.153	.303	.038	5.941	.000
Correlation Trait SC with Math t1	.496	.390	.585	.054	9.175	.000
Cross-sectional correlation State SC t1 with Math t1	-.112	-.301	.076	.096	-1.172	.241
Cross-sectional correlation State SC t3 with Math t3	.118	.031	.205	.044	2.668	.008
First language - Trait SC	-.014	-.075	.046	.031	-0.464	.642
First language - Math t1	.100	.042	.159	.030	3.367	.001
First language - Math t3	.037	-.008	.082	.023	1.592	.111
Gender - Trait SC	.214	.143	.286	.036	5.872	.000
Gender - Math t1	.122	.068	.176	.028	4.398	.000
Gender - Math t3	.033	-.010	.076	.022	1.490	.136
Age - Trait SC	-.155	-.231	-.080	.038	-4.044	.000
Age - Math t1	-.158	-.211	-.104	.027	-5.799	.000
Age - Math t3	-.074	-.110	-.040	.018	-4.183	.000

Note. N = 1,156; SC = math self-concept, Math = math competencies, First language (1 = first language German, 2 = first language not German). Covariates were included as predictors of Trait SC, Math t1, and Math t3. Model fit: $\chi2(62) = 91.946$, $p < .008$, RMSEA = .020, CFI = .994, SRMR = .024.

https://doi.org/10.1024/1010-0652/a000366

Appendix I

Table I 1

Comparison of RI-CLPMs with gender as time-invariant covariate (constant vs varying effect)

	Unconditional				Gender (constant effect)				Gender (varying effect)			
		CI 95 %				CI 95 %				CI 95 %		
Parameter	EST	low	up	*p*	EST	low	up	*p*	EST	low	up	*p*
Lagged effects												
State SC t1 - t2	.06	-.27	.40	.71	.04	-.30	.39	.80	.03	-.33	.38	.89
State SC t2 - t3	.42	.22	.62	.00	.41	.21	.62	.00	.41	.21	.61	.00
Math t1 - t3	.68	.65	.72	.00	.68	.65	.72	.00	.68	.65	.72	.00
Cross-lagged effects												
Math t1 - State SC t2	.16	-.01	.32	.06	.16	-.01	.33	.06	.14	-.04	.31	.12
State SC t2 - Math t3	.08	.00	.16	.04	.08	.00	.16	.04	.08	.00	.16	.04
Correlations												
Trait SC with Math t3	.22	.16	.30	.00	.23	.16	.30	.00	.23	.16	.30	.00
Trait SC with Math t1	.51	.42	.60	.00	.50	.41	.59	.00	.51	.41	.60	.00
State SC t1 with Math t1	-.12	-.30	.06	.19	-.12	-.30	.06	.19	-.14	-.32	.05	.15
State SC t3 with Math t3	.13	.04	.21	.01	.13	.04	.21	.01	.12	.04	.21	.01

Note. N = 1,299; SC = math self-concept, Math = math competencies. Gender was included as predictor of Trait SC, State SC (constant vs varying effect), Math t1, and Math t3. All models fitted the data well. Standardized model results are reported.

https://doi.org/10.1024/1010-0652/a000366

Table I 2

Comparison of RI-CLPMs with age as time-invariant covariate (constant vs varying effect)

	Unconditional				Age (constant effect)				Age (varying effect)			
		CI 95 %				CI 95 %				CI 95 %		
Parameter	EST	low	up	*p*	EST	low	up	*p*	EST	low	up	*p*
Lagged effects												
State SC t1 - t2	.06	-.27	.40	.71	.07	-.29	.42	.72	.06	-.30	.43	.73
State SC t2 - t3	.42	.22	.62	.00	.41	.20	.61	.00	.41	.20	.61	.00
Math t1 - t3	.68	.65	.72	.00	.67	.64	.71	.00	.68	.64	.71	.00
Cross-lagged effects												
Math t1 - State SC t2	.16	-.01	.32	.06	.15	-.02	.32	.09	.14	-.03	.32	.10
State SC t2 - Math t3	.08	.00	.16	.04	.08	.00	.16	.05	.08	.00	.16	.05
Correlations												
Trait SC with Math t3	.22	.16	.30	.00	.22	.15	.30	.00	.22	.15	.29	.00
Trait SC with Math t1	.51	.42	.60	.00	.50	.41	.60	.00	.50	.41	.60	.00
State SC t1 with Math t1	-.12	-.30	.06	.19	-.12	-.31	.07	.21	-.12	-.31	.07	.20
State SC t3 with Math t3	.13	.04	.21	.01	.12	.03	.21	.01	.12	.03	.21	.01

Note. N = 1,299*,* SC = math self-concept, Math = math competencies. Age was included as predictor of Trait SC, State SC (constant vs varying effect), Math t1, and Math t3; N = 1,170 for the models including time-invariant covariate. All models fitted the data well. Standardized model results are reported.

https://doi.org/10.1024/1010-0652/a000366

Table I 3

Comparison of RI-CLPMs with first language as time-invariant covariate (constant vs varying effect)

	Unconditional				First language (constant effect)				First language (varying effect)			
		CI 95 %				CI 95 %				CI 95 %		
Parameter	EST	low	up	*p*	EST	low	up	*p*	EST	low	up	*p*
Lagged effects												
State SC t1 - t2	.06	-.27	.40	.71	.10	-.27	.47	.60	.10	-.27	.47	.60
State SC t2 - t3	.42	.22	.62	.00	.44	.23	.65	.00	.45	.24	.66	.00
Math t1 - t3	.68	.65	.72	.00	.68	.64	.71	.00	.68	.64	.72	.00
Cross-lagged effects												
Math t1 - State SC t2	.16	-.01	.32	.06	.13	-.03	.29	.12	.12	-.05	.29	.16
State SC t2 - Math t3	.08	.00	.16	.04	.08	-.00	.16	.06	.08	-.00	.16	.06
Correlations												
Trait SC with Math t3	.22	.16	.30	.00	.23	.16	.31	.00	.23	.16	.31	.00
Trait SC with Math t1	.51	.42	.60	.00	.52	.42	.62	.00	.52	.42	.63	.00
State SC t1 with Math t1	-.12	-.30	.06	.19	-.13	-.31	.06	.18	-.14	-.33	.06	.18
State SC t3 with Math t3	.13	.04	.21	.01	.12	.03	.21	.01	.12	.03	.21	.01

Note. N = 1,299, SC = math self-concept, Math = math competencies. First language was included as predictor of Trait SC, State SC (constant vs varying effect), Math t1, and Math t3. N = 1,264 for the models including time-invariant covariate. All models fitted the data well. Standardized model results are reported.

https://doi.org/10.1024/1010-0652/a000366