Intraindividual Structural Equation Models (ISEM)

Lars-Erik Malmberg

University of Oxford, UK

QRM Keynote, 10 June 2024,
University of Gothenburg, Sweden

Outline

Outline

1. Intraindividual SEM (ISEM)
   - Lagged variables
2. Multiple reporter models
   - Cross-classified models
3. Dynamic SEM (DSEM)
   - Stationarity and equidistance
   - Individual differences in residuals
   - Admissible range priors (McNeish)

Why intraindividual (intrapersonal, micro-longitudinal, within-person, ideographic) research?

A state is an experience in a situation
A process is a

- sequence of states, or
- a transformation of one state into another

The important questions to ask are (Schmitz, 2006)

- What do teaching and learning processes look like?, and
- What effect does teaching have on learning processes?

We cannot draw conclusions about processes based on cross-sectional data!

- Intensive longitudinal data needed

Many theoretical perspectives pose process models

Why?

There is a growing interest in intraindividual research

- Subjective reports collected via diaries, experience sampling and ecological momentary assessments
  - Retrospection bias reduced, contextual closeness enhanced
- Multiple reporters, e.g., teachers, observers, parents, peers
  - Self-reporter bias reduced
- Objective measures collected using brief standardized tests and tasks, computer traces, ambulatory devices for physiological measures
  - facial expressions, heart-rate (HR), electrodermal activity (EDA), physical activity (PA)

Toward a theory of individualized / personalized learning

Educational Psychology Review

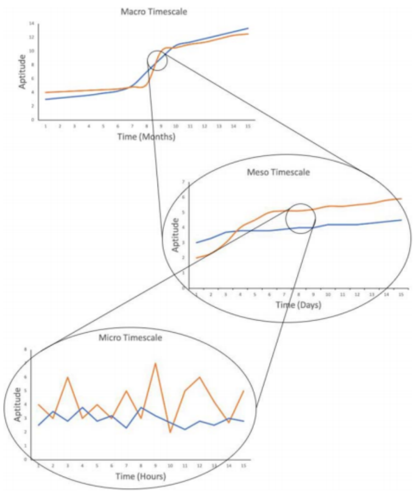


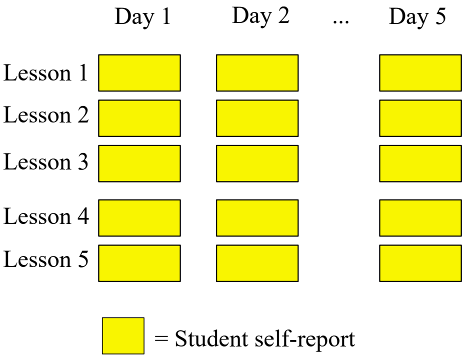
Fig. 2 Development of two fictional learners in a single aptitude over the course of months, days, and hours

(Tetzlaff, Schmiedek, & Brod, 2021)

Part I: Intraindividual SEM (ISEM)

"Standard" intraindividual design

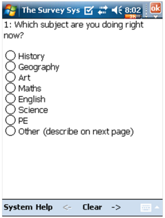
The Learning Every Lesson (LEL) study

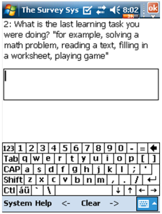


Learning Experience Questionnaire, LEQ
Teacher reports: performance, task-focus, involvement
(Malmberg, Woolgar, & Martin, 2013; Malmberg & Martin, 2019)

Learning Every Lesson

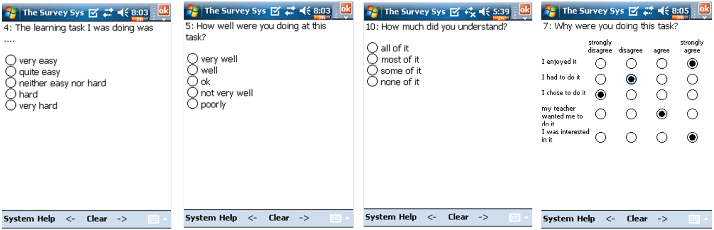
The Learning Every Lesson (LEL) study 2007-08





Mixtures of single-item and multiple item constructs

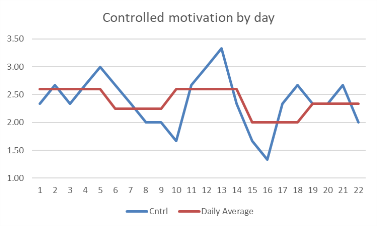
Learning Every Lesson (LEL) 2007-08 (more items....)

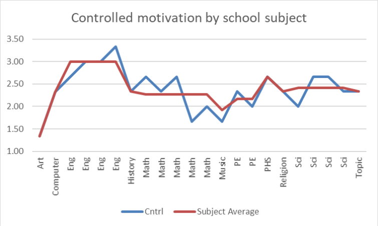


Let's get an intuition for intraindividual data

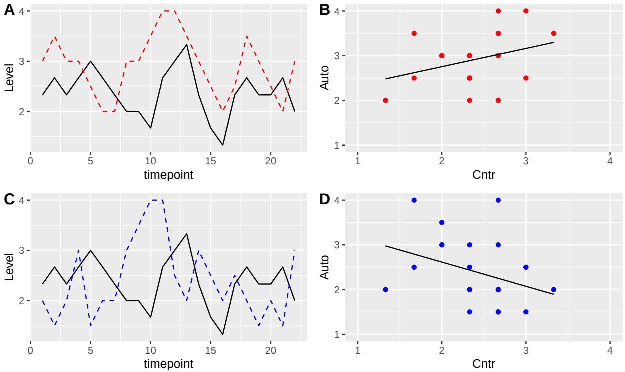
Fictive cases

Learning experiences by rank-order time and by school-subject





Learning experiences by rank-order time (left) and intraindividual association (right)



Note: Red is Kamal's autonomous motivation and blue is Lisa's autonomous motivation. Black is controlled motivation

Modelling sequence

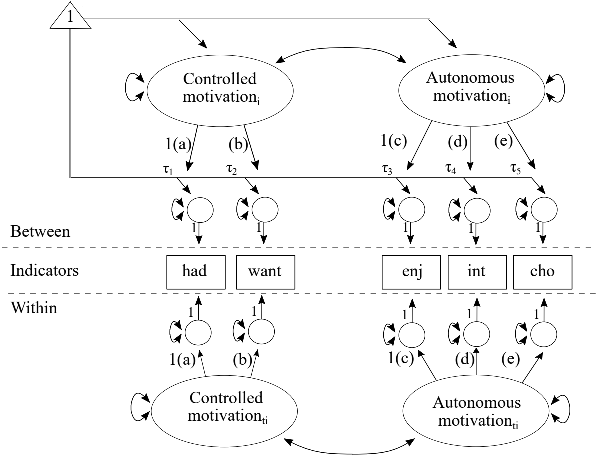
1. Factor structure
   - Multilevel Confirmatory Factor Analysis (MCFA)
2. Fixed and random effects models, cross-level interactions moderation
   - Multilevel SEM (MSEM)
3. Autoregressive and cross-lagged models
   - Multilevel SEM (MSEM)

Table: Descriptives for all data, and level specific data

	enj	cho	int	had	Min	Max	M	SD	Skew	Kurt
Raw										
enj					1	4	3.00	1.04	-0.72	-0.70
cho	0.55				1	4	2.48	1.13	0.05	-1.38
int	0.77	0.61			1	4	2.88	1.09	-0.53	-1.04
had	0.16	0.09	0.17		1	4	3.21	0.92	-1.09	0.40
tea	0.20	0.16	0.23	0.74	1	4	3.28	0.88	-1.25	0.91
CWC										
enj					-2.82	2.73	0.00	0.85	-0.43	0.24
cho	0.52				-2.65	2.73	0.00	0.80	0.10	0.63
int	0.72	0.56			-2.82	2.73	0.00	0.89	-0.30	-0.07
had	0.07	0.03	0.08		-2.82	2.73	0.00	0.66	-1.23	3.58
tea	0.11	0.07	0.11	0.61	-2.82	2.73	0.00	0.62	-1.33	4.38
GMC										
enj					1.13	4.00	2.98	0.61	-0.31	-0.17
cho	0.62				1.00	4.00	2.47	0.78	0.14	-0.71
int	0.86	0.70			1.00	4.00	2.86	0.63	-0.13	-0.35
had	0.29	0.14	0.30		1.00	4.00	3.20	0.63	-1.06	1.49
tea	0.33	0.24	0.38	0.88	1.00	4.00	3.27	0.62	-1.40	2.58

Note: The top section is estimated based on the raw single-level data ($n_{ti} = 3{,}017$), the mid section on the within-cluster centred data ($n_{ti} = 3{,}017$), and the bottom section is based on the grand-mean centred data ($n_i = 202$).

(1) Multilevel Confirmatory Factor Analysis (MCFA)

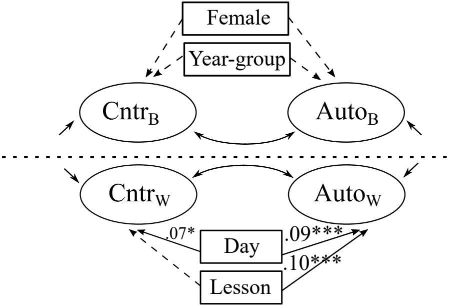


(Malmberg, 2020)

Challenges for model fit for MCFA

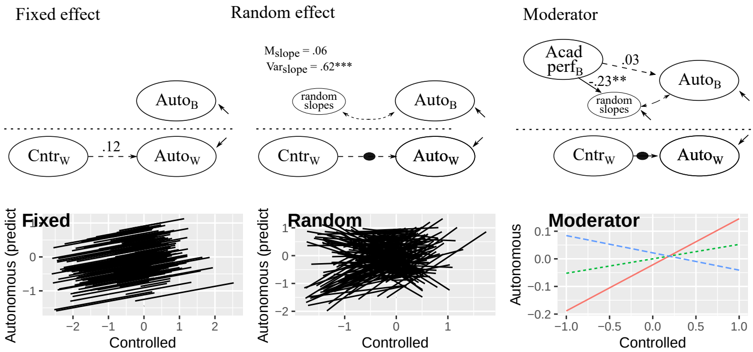
- Goodness of fit measures CFI and RMSEA are not level-specific $\rightarrow$ Can be handcalculated following (Ryu & West, 2009)
- SRMR are level-specific SRMR_B and SRMR_W

(1b) MCFA with level-specific covariates



(Malmberg, 2020)

(2) Fixed and random effects models, cross-level interactions



Note: Blue line = +1SD academic performance, red line = -1SD academic performance

The starting point, back in 1993 (Schmitz & Skinner, 1993)

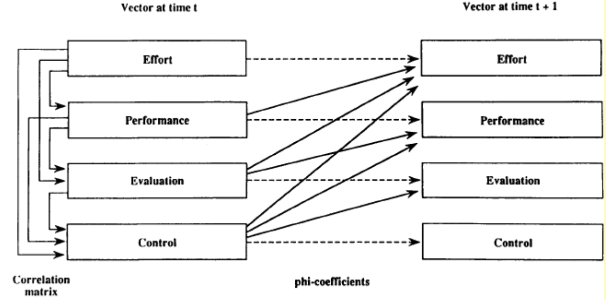
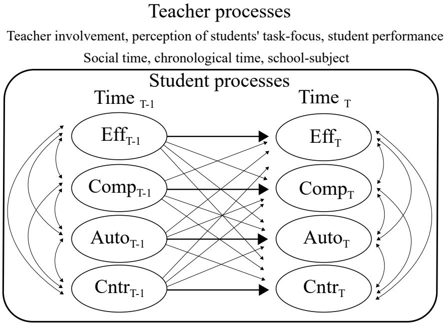


Figure A3. The relations between the components of the time-series vector (contained in the correlation matrix) and the links to the vector components at the next point in time (contained in the phi matrix).

Multivariate time-series models of cases

Cyclic and dynamic model following Schmitz and Skinner (1993)



- Time-points: Time_T and Time_{T-1}
- Autoregressive paths (cyclic) = self-sustaining
- Cross-lagged paths (dynamic)
  - Positive = self-enhancing
  - Negative = self-diminishing

Creating contextually appropriate lags

- Lags within days (Auto_{T-D1})
- Lags within school subjects (Auto_{T-S1})

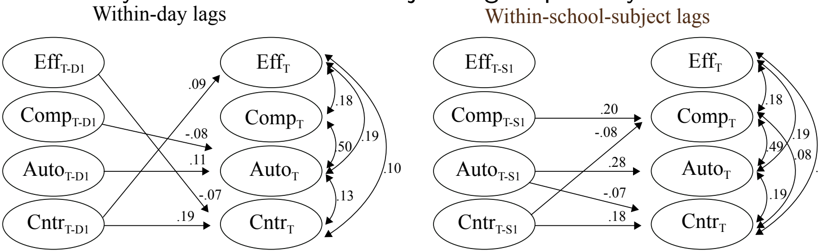
Day	Lesson	Subject	Auto	Auto_{T-D1}	Auto_{T-S1}
2	1	English	3.00	*	2.67
2	2	Maths	1.67	3.00	2.00
2	3	Arts	2.00	1.67	*
2	4	PE	1.67	2.00	*
2	5	History	2.33	1.67	*
3	1	Maths	2.00	*	1.67
3	2	Language	3.00	2.00	*
3	3	English	2.67	3.00	3.00
3	4	Science	2.00	2.67	2.33
3	5	Computer	2.67	2.00	*

Creating contextually appropriate lags

- Lags within days (Auto_{T-D1})
- Lags within school subjects (Auto_{T-S1})

Day	Lesson	Subject	Auto	Auto_{T-D1}	Auto_{T-S1}
2	1	English	$\boxed{3.00}$	*	2.67
2	2	Maths	$\boxed{1.67}$	3.00	2.00
2	3	Arts	2.00	1.67	*
2	4	PE	1.67	2.00	*
2	5	History	2.33	1.67	*
3	1	Maths	2.00	*	$\boxed{1.67}$
3	2	Language	3.00	2.00	*
3	3	English	2.67	3.00	$\boxed{3.00}$
3	4	Science	2.00	2.67	2.33
3	5	Computer	2.67	2.00	*

Within-day and within-school-subject lags separately



- Effort not sustained
- Competence and autonomous motivation sustained in subjects
- Controlled motivation sustained *both* in days and in subjects

Challenges with MSEM?

- Small variance at between level
  - Adjust s.e. for clustering
  - Use MCFA for test of structural validity, then run MLMs
- Parameters that misbehave
  - Haywood-cases with negative variances $\rightarrow$ fix at > 0?
  - Correlations very close to or > 1
  - Massive variance $\rightarrow$ adjust scaling of variable
  - Is model misspecified?
- Model doesn't converge
  - Allow to run for longer, more iterations
  - Set lower convergence criteria
  - Change estimator $\rightarrow$ Weighted Least Squares or Bayesian

Part II: Multiple reporter models

1. Teacher-reports of student characteristics (one-off)
2. Students *and* teacher-reported emotions the same lesson
3. Teacher-reports of four target students per lesson

Measures

Teacher reports (once)

- Students' task focus (6 items from Zhang et al., 2011) e.g., "actively attempts to solve even difficult tasks", "gives up easily" ($\alpha = .94$; $\omega_B = .95$)
- Teacher's involvement: Dedication of resources (aid, time, energy; e.g., Wellborn et al., 1988): (6 items) "I spend one-to-one time with him / her", "I give him / her a lot of attention" (0 = not at all, 4 = very often), $\alpha = .84$; $\omega_B = .84$; λ_Bs = .43 to .89 .
- Students' academic performance; Coverage of curriculum levels in maths, English and science ($\alpha = .90$, and $\omega_B = .90$) .

Teacher reports of student characteristics

How are teacher perceptions and involvement associated with students' learning experiences?

Learning
Every
Lesson

Table: Between-level (students) correlation matrix

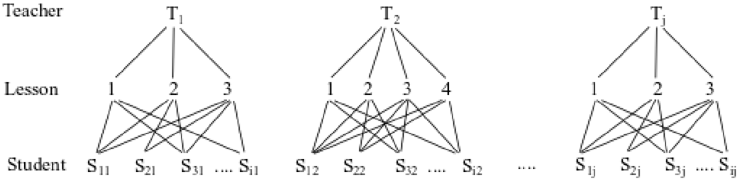
	Effort_i	Comp_i	Auto_i	Cntr_i	Perf	Focus
Effort_i						
Comp_i	.23					
Auto_i	.29	.50				
Cntr_i	.20	.30	.44			
Perf	*.07*	.49	*.08*	.24		
Focus	.14	.41	.21	.21	.71	
Inv	*-.21*	-.28	*-.03*	*.01*	-.45	-.42

Note: Eff = effort exertion, Comp = competence belief, Auto = autonomous motivation, Cntr = controlled motivation; i = student average during the week; Perf = student academic performance, Focus = student task-focus, Inv = teacher involvement with each student.

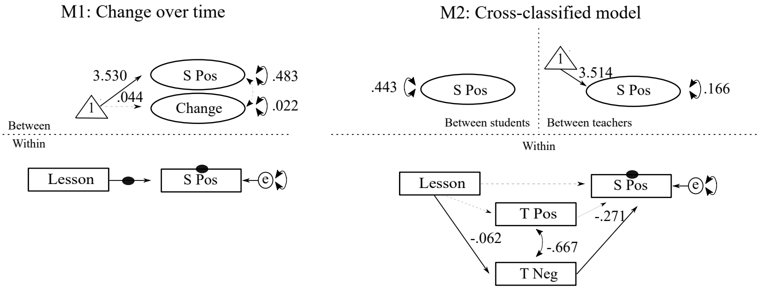
How do teacher emotions predict students' emotions during a school day (Sorkkila et al, submitted)?

- Primary school teachers ($n_j = 20$) in Finland reported their emotions during 2-4 lesson in a school day ($n_{tj} = 61$)
- Students ($n_i = 258$) reported on their emotions during those same lessons ($n_{ti} = 672$)

Lessons were nested in both students and teachers



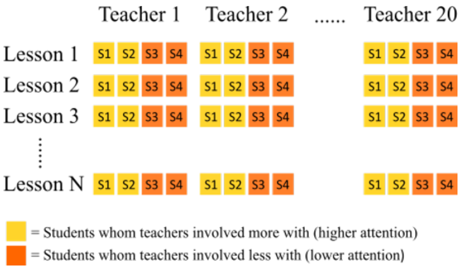
Teachers negative emotions decreased during the day, and predicted lower level of positive emotions in students



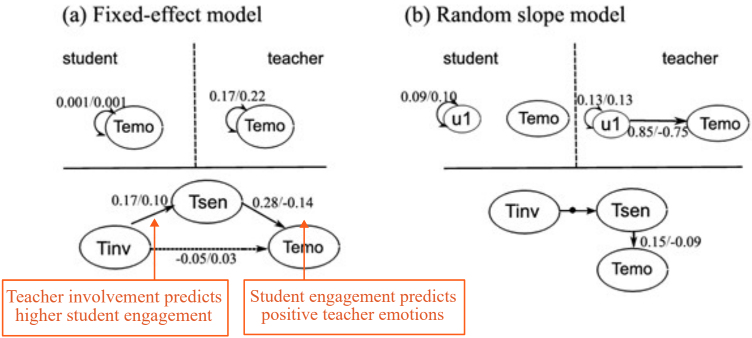
Teachers' reports of target students

Do teachers differ in their interaction with individual students during one week's time? (Li, Mayer, & Malmberg, 2024).

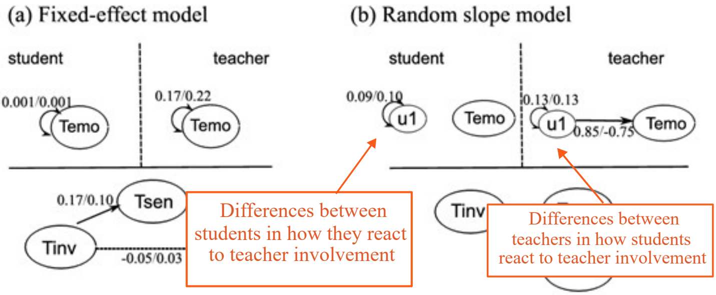
- Primary school teachers ($n_t = 20$) in Taiwan reported on four target students ($n_{st} = 80$) each lesson ($n_{lst} =$ 981-996)
- Does student-reported teacher's instructional support differentially predict teacher-reported student behavioral engagement (two student Ts' "pay relatively more" and two "relatively less" attention to?)



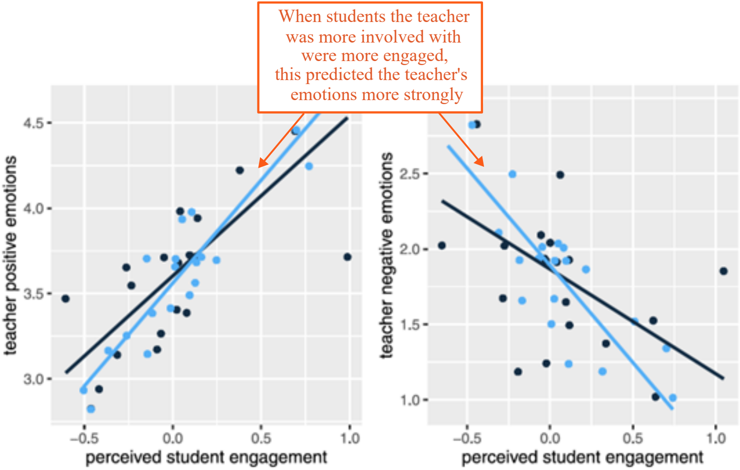
Teachers' reports of target students



Teachers' reports of target students



Teachers' reports of target students



Note: Blue lines represent two target students the teacher was relatively more involved with, and black lines represent two target students the teacher was relatively less involved with.

Optimal study designs for educational settings?

- Bayesian estimator
  - Complex models possible to specify
  - Does not rely on large sample theory
  - Handles small variances and between-level sizes
- Challenges
  - Use the Bayesian estimator only as an "estimator" rather than adopting the Bayesian philosophy (i.e., specify priors)
  - Posterior estimates misbehave (trace plots, MCMC autocorrelations, Potential Scale Reduction > 1.05)

Part III: Dynamic SEM (DSEM)

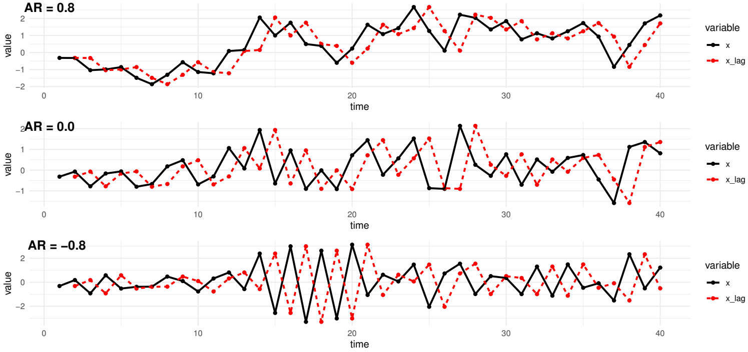
Model process-parameters
"Adjust" unequal time-lags
Individual residuals
Semi-informative priors

Advantages of DSEM

- Combines multilevel models and time-series models within SEM
  - "Multiple" time-points per person
  - Close equal spacing between time-points
  - Stationarity assumed (i.e., no mean trend)
  - There is variability around a stable mean
  - We want to predict fluctuations (within-person variation)
- Many features of DSEM
  - Adjustment of time-lags ("TINTERVAL"simulation)
  - Individually varying residuals
  - Process variables modelled at the between-level

AR examples

Examples of stationary time-series with different auto-regressive parameters (0.8, 0.0, -0.8)



(Multilevel Multivariate) Vector Auto-Regressive (VAR) model with individual (random) residuals (Hamaker, Asparouhov, Brose, Schmiedek, & Muthén, 2018)

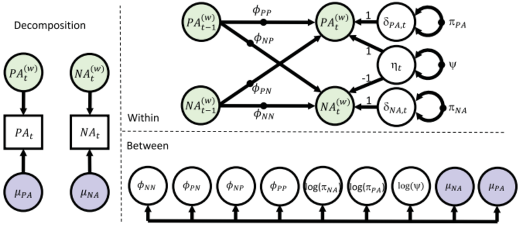


Figure 4. Representation of the multilevel VAR(1) model with random innovation variances and covariance. Left part contains the decomposition into within-person (time-varying) and between-person (time-invariant) components. Top right contains the within-person level model, which is a VAR(1) model. Bottom right contains the between-person level model, which includes the between-person components of the observed data as well as all the random effects of the model, corresponding to the solid black circles in the within-person level model. It shows that random variances at the within-person level are modeled using their log at the between-person level (which are then assumed to come from a multivariate normal distribution); this is done to ensure that the individual variances are positive for all individuals.

Associations between process-parameters (at level 2)

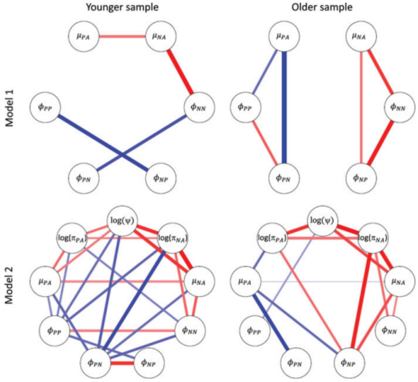
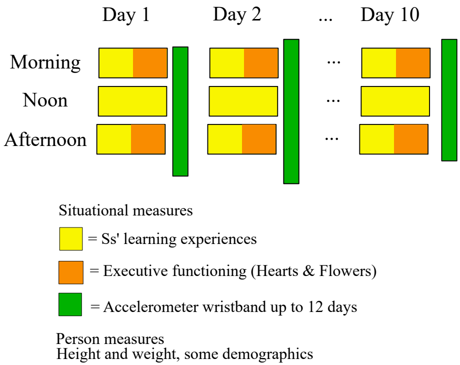


Figure 3. Visual representations of the correlation structure of the random effects of the multilevel VAR(1) models in the two samples. Upper part contains the correlations between the six random effects in model 1, while the bottom part contains the correlations between the nine random effects of model 2. Only correlations whose 95% credible interval did not include zero are included here. Red connections represent positive correlations, blue connections represent negative correlations. Strength of the correlations is represented by the thickness of the connection.

PP, NN are auto-regressive (inertia, stability) paths
NP, PN are cross-lagged paths

The Physical activity, Engagement and Cognition (PEC) study

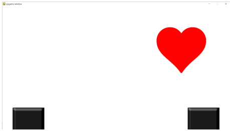
(Malmberg, Heemskerk, Lo, Esser, Dawes, Kodzhabashev, Roebers)



43 children (21 boys) from 8:11 and 11:4 years old (M_{Age} = 9:11 years, SD_{Age} = 9 months), in grades 3,4 and 5 in England completed a modified version of the Hearts and Flowers (HF) trials (Davidson, Amso, Anderson, & Diamond, 2006) .

The Hearts and Flowers tasks

- The Common EF perspective (Friedman & Miyake, 2017) poses that working memory, inhibition and shifting "might be a collection of processes" (p. 187).
- "Hearts" block with congruent responses



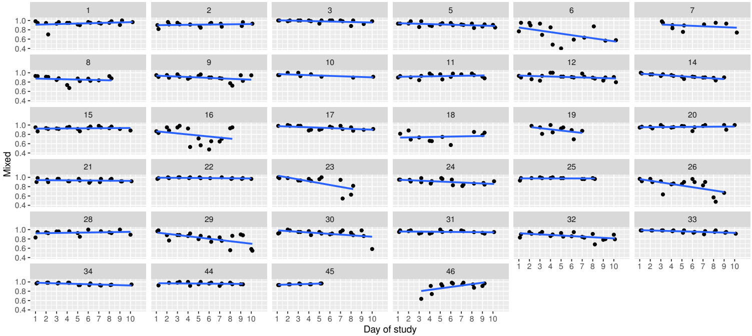
Press button on the same side as the heart (to the right)

- "Flowers" block with *in*congruent responses

Press button on the *opposite* side of the flower (to the left)

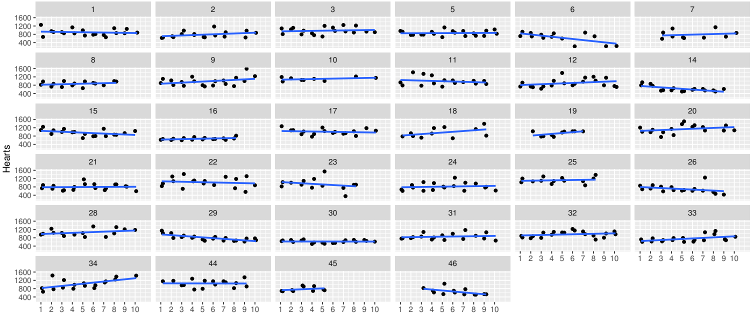
- "Mixed" session with congruent *and* incongruent responses

Accuracy for the mixed block ("Mixed") during 10 days, by child



Some possible non-stationary trends? Removing cases did not enhance model fit or convergence. We included time-of-day and day-or-week in some models as covariates.

Response-time for the mixed block ("Mixed") during 10 days, by child



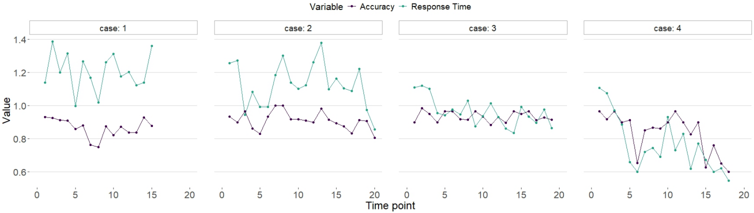


Figure: Four case-studies
Note: Values for accuracy were 0.56-1.00 indicating proportion correct, and response-time were 0.54-1.55 in seconds.

.

Table: Descriptives of executive functioning tasks (n_{ti} = 651) of students (n_i = 43)

Within	1.	2.	3.	4.	5.	M	SD	Range	
1. State-Acc						0.92	0.07	0.56 - 1.00	
2. State-Acc lag[a]	.18					0.92	0.07	0.56 - 1.00	
3. State-RT	.37	.13				1.08	0.20	0.54 - 1.55	
4. State-RT lag[a]	.11	.36	.28			1.08	0.20	0.54 - 1.55	
5. Time[b]	-.11	.06	.01	.02		0.12	0.92	-1 to 1	
6. Day[c]	-.14	-.11	-.27	-.14	-.04	-0.09	1.38	-2 to 2	
Between	7.	8.	9.	10.		M/%	SD	Range	ICC
7. Trait-Acc						0.92	0.05	0.75 - 0.99	0.43
8. Trait-RT	.30					1.09	0.16	0.74 - 1.35	0.55
9. Boy	-.18	-.21				48.8%			
10. Age	.09	-.22	.19			9.95	0.72	8.90 - 11.35	

Note: "Acc" = Accuracy and "RT" = response-time for the Mixed tasks. Within-level variables are referred to as "states" and between-level variables as "traits". Associations are within and between-level correlation (R psych). For lagged variables (T-1) n_{ti} = 608. Descriptives at the within-level were based on raw data, and at the between-level based on aggregated raw data. Accuracies are presented as proportions, and reaction-time in seconds. [b] Time of day was centered at noon (-1 = morning, 09:00-10:59, 0 = noon, 11:00 - 12:59; 1 = afternoon, 13:00-15:00). [c] Day of week was centered mid-week (-2 = Monday, 0 = Wednesday, 2 = Friday).

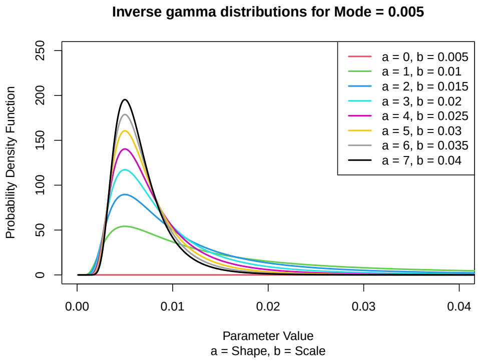
.

Table: Priors for model parameters

Parameter	Magnitude	Variance	Reference
Beta: acc_T on acc_{T-1}	0.29	1.00	Trevillion et al (2022)
Beta: acc_T on $resp_{CWCT}$	0.14	1.00	raw data ($r = .37$)^2
Mean: acc_B	0.92	1.00	raw data
Mean: ϵ	-3.00	1.00	Exp(0.05)
Mean: $resp_W$	1.08	1.00	raw data
Var: $resp_W$	4.00	0.20	IG for Mode = 0.04[a]
Var: acc_B	3.00	0.02	IG for Mode = 0.005[a]
Var: acc_W	4.00	0.20	IG for Mode = 0.04[a]
Var: ϵ_B	3.00	4.00	IG for Mode = 1.000[a]

Note: [a] We specified weakly informative admissible-range priors priors following (McNeish, 2019) for variances.

Following McNeish (2019) we use descriptive data to calculate the proportion of variance at the within (≈ 0.05) and between-levels (≈ 0.01). Then we find boundaries for conservative estimates (0.04 and 0.005) of an inverse gamma distributions.



Example data for fictive case

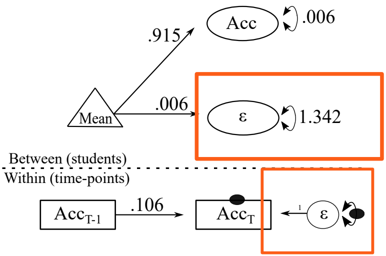
studid	obs	day	time	Acc_T	Acc_{T-1}	RT_T	RT_{T-1}
001	1	1	10:36	0.94	*	0.995	*
001	2	1	14:44	0.95	0.94	1.100	0.995
001	3	2	09:55	0.98	0.95	1.050	1.100
001	4	2	14:55	0.92	0.98	1.002	1.050
...	...	...	...	...	...	...	
001	18	9	14:00	-9	0.93	-9	1.120
001	19	10	10:10	0.92	-9	1.200	-9
001	20	10	14:55	0.95	0.92	1.102	1.200

Note: obs = rank-order observation (discrete time), time = chronological time, Acc = Accuracy (proportion, 0.40-1.00), RT = Response time (seconds), T = concurrent time-point, T-1 = lagged time-point, -9 = missing data (non-response), * = missing by design

Adjust for non-equidistant time-lags using the "TINTERVAL" function, in 3-hour chunks. Sample-size in simulation increases from n_{ti} = 651 to n_{ti} = 3,492.

In the MLM residuals are independent and identically distributed in a regression model, $\epsilon \sim$ i.i.d. $N(0, \sigma^2)$. We want to estimate individual residuals - these increase precision in other parameters

A. Stability of accuracy



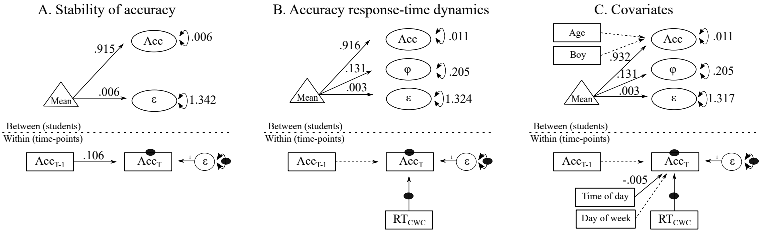


Figure: Dynamic structural equation models

Note: Diagrams are simplified, not all parameters reported. Rectangles indicate observed measures at the within-level, and ellipses latent constructs at the between-level. ϵ = individually varying residuals we report exp(ϵ). ϕ = random term for accuracy-on-response-time-slopes. Models were estimated in Mplus, using the time-interval function to adjust for non-equidistant time-points.

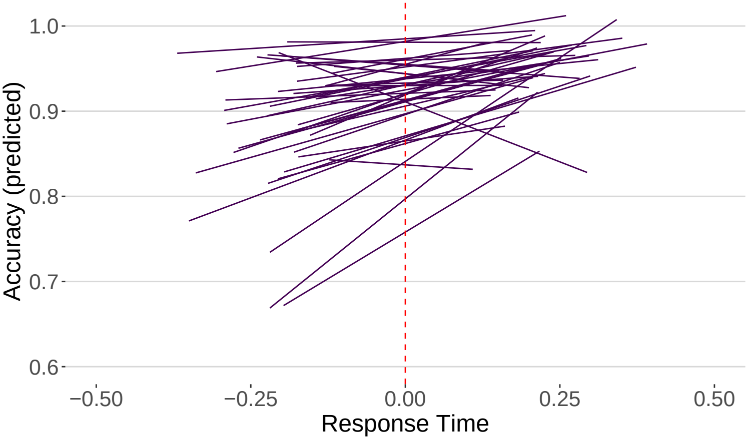


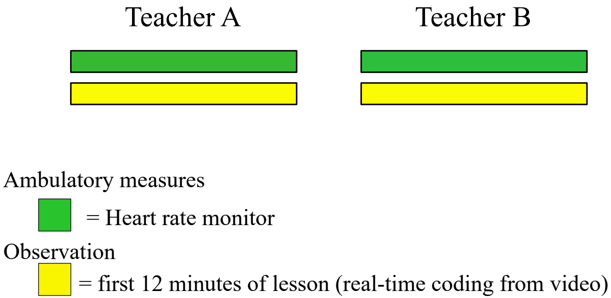
Figure: Individual difference in accuracy-on-response-time-slopes

When children who are less accurate slow down they perform better

DSEM with biophysiological measures (HR) and coded videos of interaction

Biophysiological measures

Linking observation and biophysiological data (in 5 sec segments). Observed agency (e.g., leadership) and communion (e.g., friendliness)



(Donker, Van Gog, & Mainhard, 2018)

UNIVERSITY OF OXFORD

"Objective stress" heart-rate (HR) ambulatory physiological measurement (5 sec segments), controlling for physical movement

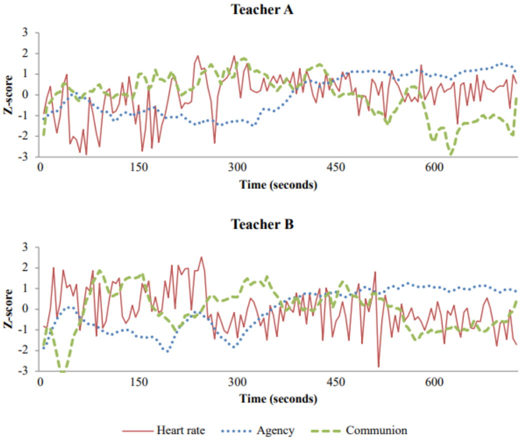


Figure 3. Time series for heart rate, Agency, and Communion.

Time-series modelling of HR $\rightarrow$ Agency $\rightarrow$ Communion

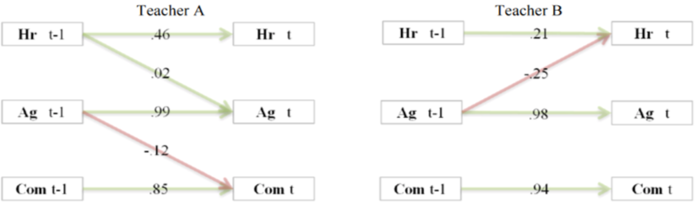


Figure 5. Dynamic Structural Equation Model (DSEM; Asparouhov et al., 2017). The subscript *t* refers to a time point and the subscript *t-1* refers to the previous time point of the time series. The reported autocorrelations are slightly different from section 3.2.2 due to the addition of the cross-lagged associations. Only statically significant cross-lagged effects are presented (*N*=144 time points per teacher). Hr = heart rate; Ag = Agency; Com = Communion.

Summary Part III

DSEM is a powerful tool

- Our own adventures do not take advantage of the full DSEM
- Many elements of the DSEM can be incorporated into models
- Are we used thinking in terms of time-series?
- "Small n" studies (with informative priors)
- Ratio between n and t (Schultzberg & Muthén, 2018)

Conclusions

Where do we go from here?

Future designs

There is growing body of intraindividual educational research

- Focus on processes
- Many complementary models for processes
- Real-time objective and subjective measures
- Modelling individualized learning

How is the software side?

- MPlus versus R brms, R Stan

Future designs

Can we turn process-studies into nudge/prompt/just-in-time designs?

- For medical students (Breitwieser, Neubauer, Schmiedek, & Brod, 2021), exam-results could be improved ($d = 0.21$), while short-term effects wore out fast
- Children's vocabulary learning using an app with or without prompts, or with or without planning (Breitweiser et al., 2023)
- Investigate effects of individualized support in class (Street et al 2023-27)

Acknowledgements

ISEM
Cross-classified
DSEM
Conclusions
Acknowledge
References

Collaborators:

- Andrew Martin, Cassandra Woolgar, Ted Walls, Todd Little, Timothy Lim, Hazel Hagger, Sophie Webster, Asko Tolvanen, Jari-Erik Nurmi, Eija Pakarinen, Kati Vasalampi, Joona Muotka, Noona Kiuru, Jake Argent, Kate Mee, Bethan Thomson, Chris Heemskerk, Karin Street, Charlotte Trevillion, Ekaterina Michou, Pei-Hsin Li, Kyle Davison, Dominik Bulla, Mirna Šumatić, Henry Lo, Abdul Karim Ismael, Patrick Esser, Helen Dawes, Claudia Roebers, Matilda Sorkkila

Funders

- Research Councils UK (RCUK), John Fell Fund, ESRC, Johan Jacob Foundation

Thank you!

UNIVERSITY OF OXFORD

References I

Baker, D. H., Vilidaite, G., Lygo, F. A., Smith, A. K., Flack, T. R., Gouws, A. D., & Andrews, T. J. (2021). Power contours: Optimising sample size and precision in experimental psychology and human neuroscience. *Psychological Methods*, *26*(3), 295-314. doi: 10.1037/met0000337

Breitwieser, J., Neubauer, A. B., Schmiedek, F., & Brod, G. (2021). Self-regulation prompts promote the achievement of learning goals – but only briefly: Uncovering hidden dynamics in the effects of a psychological intervention. *Learning and Instruction*, August, 101560. doi: 10.1016/j.learninstruc.2021.101560

Davidson, M. C., Amso, D., Anderson, L. C., & Diamond, A. (2006). Development of cognitive control and executive functions from 4 to 13 years: Evidence from manipulations of memory, inhibition, and task switching. *Neuropsychologia*, *44*(11), 2037-2078. doi: 10.1016/j.neuropsychologia.2006.02.006

Donker, M. H., Van Gog, T., & Mainhard, M. T. (2018). A quantitative exploration of two teachers with contrasting emotions: Intra-individual process analyses of physiology and interpersonal behavior. *Frontline Learning Research*, *6*, 162–184. doi: 10.14786/flr.v6i3.372

Friedman, N. P., & Miyake, A. (2017). Unity and diversity of executive functions: Individual differences as a window on cognitive structure. *Cortex*, *86*, 186-204. (Is a "single" brain model sufficient?) doi: https://doi.org/10.1016/j.cortex.2016.04.023

Gelman, A., Carlin, J. B., Stern, H. S., Dunson, D. B., Vehtari, A., & Rubin, D. B. (2013). *Bayesian Data Analysis, (3rd Ed.)*. Boca Raton: CRC Press.

Hamaker, E. L., Asparouhov, T., Brose, A., Schmiedek, F., & Muthén, B. O. (2018). At the frontiers of modeling intensive longitudinal data: Dynamic structural equation models for the affective measurements from the COGITO study. *Multivariate Behavioral Research*, *53*, 820–841. doi: 10.1080/00273171.2018.1463194

Hox, J., Van de Schoot, R., & Matthijsse, S. (2012). How few countries will do? Comparative survey analysis from a Bayesian perspective. *Survey Research Methods*, *6*, 87–93. doi: 10.18148/srm/2012.v6i2.5033

Lafit, G., Adolf, J. K., Dejonckheere, E., Myin-Germeys, I., Viechtbauer, W., & Ceulemans, E. (2021). Selection of the number of participants in intensive longitudinal studies: A user-friendly shiny app and tutorial for performing power analysis in multilevel regression models that account for temporal dependencies. *Advances in Methods and Practices in Psychological Science*, *4*(1). doi: 10.1177/2515245920978738

UNIVERSITY OF OXFORD

Li, P.-H., Mayer, D., & Malmberg, L.-E. (2024). Student engagement and teacher emotions in student-teacher dyads: The role of teacher involvement. *Learning and Instruction*, *91*, 101876. doi: 10.1016/j.learninstruc.2024.101876

Malmberg, L.-E. (2020). Intraindividual structural equation models for learning experiences. *International Journal of Research & Methods in Education*, *43*(4), 413-430. doi: 10.1080/1743727X.2020.1793939

McNeish, D. (2019). Two-level dynamic structural equation models with small samples. *Structural Equation Modeling: A Multidisciplinary Journal*, *26*(6), 948-966. (PMID: 32863699) doi: 10.1080/10705511.2019.1578657

Muthén, B., & Asparouhov, T. (2012). Bayesian structural equation modeling: A more flexible representation of substantive theory. *Psychological Methods*, *17*, 313–335. doi: 10.1037/a0026802

Ryu, E., & West, S. G. (2009). Level-specific evaluation of model fit in multilevel structural equation modeling. *Structural Equation Modeling: A Multidisciplinary Journal*, *16*, 583–601. doi: 10.1080/10705510903203466

Schmitz, B. (2006). Advantages of studying processes in educational research. *Learning and Instruction*, *16*, 433–449. doi: 10.1016/j.learninstruc.2006.09.004

Schmitz, B., & Skinner, E. (1993). Perceived control, effort, and academic performance: Interindividual, intraindividual, and multivariate time-series analyses. *Journal of Personality & Social Psychology*, *64*, 1010–1028. doi: 10.1037//0022-3514.64.6.1010

Schultzberg, M., & Muthén, B. O. (2018). Number of subjects and time points needed for multilevel time-series analysis: A simulation study of dynamic structural equation modeling. *Structural Equation Modeling: A Multidisciplinary Journal*, *25*(4), 495–515. doi: 10.1080/10705511.2017.1392862

Tetzlaff, L., Schmiedek, F., & Brod, G. (2021). Developing personalized education: a dynamic framework. *Educational Psychology Review*. doi: 10.1007/s10648-020-09570-w

Zitzmann, S., Lüdtke, O., Robitzsch, A., & Marsh, H. W. (2016). A Bayesian approach for estimating multilevel latent contextual models. *Structural Equation Modeling*, *23*(5), 661–679. doi: 10.1080/10705511.2016.1207179