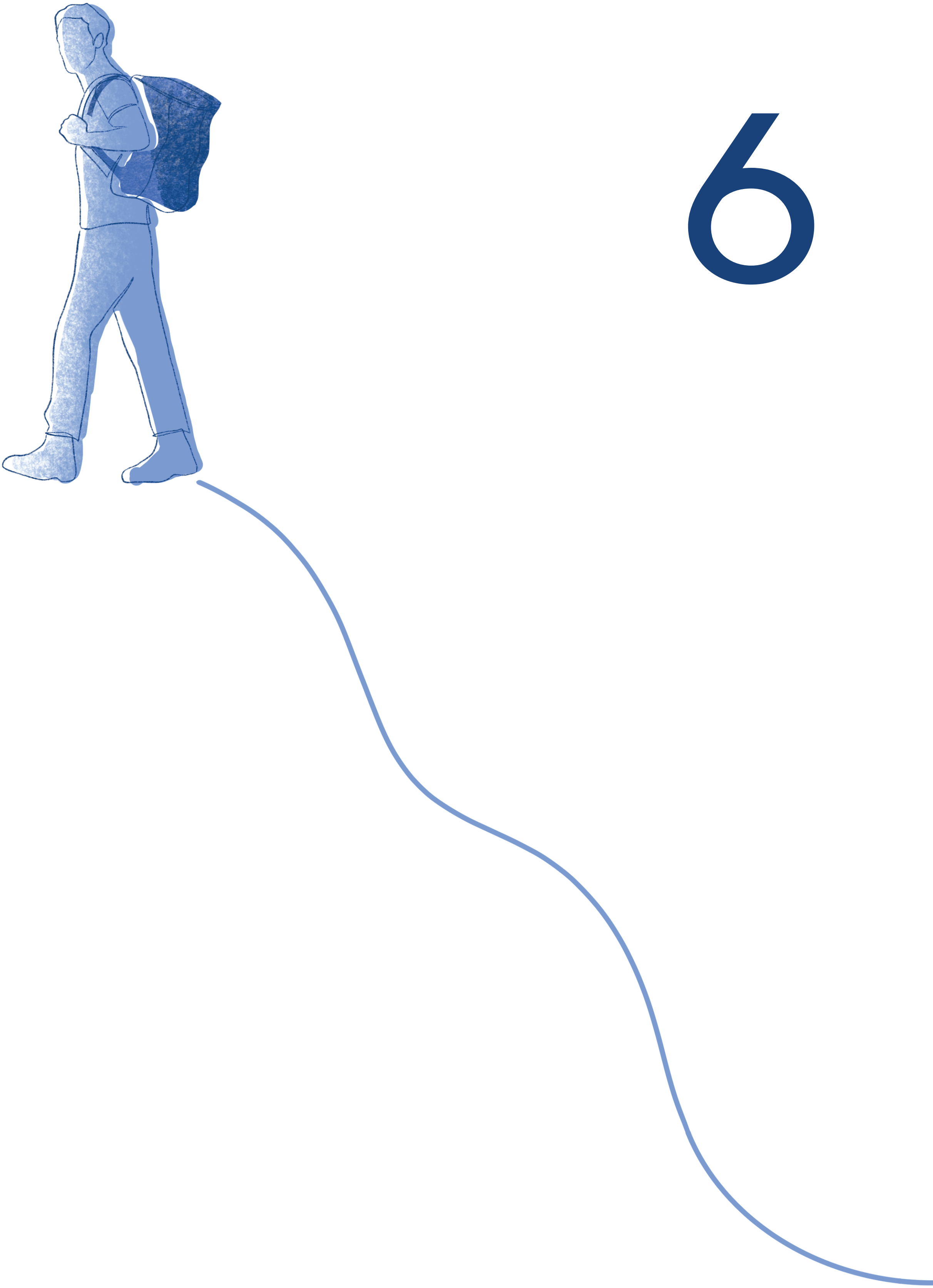
Family matters: a multi-perspective approach to the link between parenting and offspring mental health problems

Kullberg, M.L.J.

GROWING UP TOGETHER



6

Harsh parental discipline and offspring emotional and behavioural problems: Comparing findings from the random-intercept cross-lagged panel model with the monozygotic twin difference cross-lagged panel model

Under review: Kullberg, M. L.,
van Schie, C., Allegrini, A., Ahmadzadeh, Y. I., Wechsler, D.,
Elzinga, B., & McAdams, T. (2021). Harsh parental discipline
and offspring emotional and behavioural problems: Comparing
findings from the random- intercept cross-lagged panel model with the
monozygotic twin difference cross- intercept cross-lagged panel model
with the monozygotic twin difference cross- lagged panel model.

Abstract

Objective. This study examines associations between parental harsh discipline and child emotional and behavioural problems in identical twins aged 9, 12 and 16. We compare results obtained when applying the monozygotic twin difference version of the cross-lagged panel model (MZD-CLPM) and a random intercept cross-lagged panel model (RI-CLPM) to the same data with the goal to illustrate similarities and differences in findings and the conclusions that can be drawn using each method.

Method. Child reports of 5,698 identical twins from the Twins Early Development Study (TEDS) were analysed. We ran a regular CLPM to anchor are finding in the current literature. Next, we applied the MZD-CLPM and the RI-CLPM to the data.

Results. Given the study aim to infer causation, interpretation of models focussed primarily on the magnitude and significance of cross-lagged associations. In both the MZD-CLPM and the RI-CLPM behavioural problems at age 9 resulted in harsher parental discipline at age 12. Other effects were not consistently significant across the 2 models, although the majority of estimates pointed in the same direction.

Conclusion. Findings can be interpreted as corroborating (but not definite) evidence in favour of a causal effect of child behavioural problems at age 9 on experienced harsh parental discipline at age 12. Yet, in light of the triangulated methods, differences in the results between causal inference methods underline the importance of careful consideration of what sources of unmeasured confounding different models control for and that nuance is required when interpreting findings using such models. We provide an overview of what the CLPM, RI-CLPM and MZD-CLPM can and cannot control for in this respect and which conclusions can be drawn from each model.

Keywords: child, parental discipline, causal inference, emotional problems, behavioural problems

Background

Extensive research has shown that harsh parental discipline, including physical and verbal punishment and parenting strategies involving unclear and inconsistent communication with the child, is associated with the development of emotional and behavioural problems in children (see e.g. Baumrind, 1991; Crosnoe & Cavanagh, 2010; Fletcher, Steinberg, & Williams-Wheeler, 2004). Equally, some studies have also demonstrated reversed effects. That is, child emotional and behavioural problems might elicit certain parenting practices (Pinquart, 2017b, 2017a). Understanding the nature and direction of associations between parenting behaviour and child problems is likely to prove important for prevention and intervention strategiesto support families in fostering child mental well-being. In order to understand the direction of associations between harsh parental discipline and emotional and behavioural problems, researchers have used longitudinal data. The direction of effects between parenting and child emotional and behavioural problems have been mainly tested using the cross-lagged panel model (CLPM; Hipwell et al., 2008; Lansford et al., 2011; Serbin, Kingdon, Ruttle, & Stack, 2015; Wang & Kenny, 2014). The CLPM estimates the effect of a predictor on an outcome while controlling for prior differences in the outcome. As such the CLPM can help to determine whether one variable predicts the other and/or vice versa. In these models, longitudinal influences of a variable on the other over time are referred to as cross-lagged effects. Nevertheless, CLPMs have several shortcomings when it comes to determining the direction of effects between associated variables (Hamaker, Kuiper, & Grasman, 2015). For example, CLPMs do not distinguish within-person changes from between-person differences across repeated measures in the trait-like, time-invariant stability of many psychological constructs. As such, it is not possible to draw any conclusions on within-person changes using CLPMs. This is a limitation given that developmental researchers are often most interested in processes within a person. Also, CLPMs cannot account for unmeasured sources of confounding, and this may distort estimates of cross-lagged effects.

Two commonly used alternatives to the CLPM that are designed to strengthen the capacity for causal inference are the monozygotic twin difference CLPM (MZD-CLPM; Moscati, Verhulst, McKee, Silberg, & Eaves, 2018; Ritchie, Bates, & Plomin, 2015) and the random intercept CLPM (RI-CLPM; Hamaker et al., 2015). The MZD-CLPM and RI-CLPM account for stable between-subject differences across development and have each been designed with the intention of controlling for potential sources of unobserved confounding. However, these models differ conceptually and in their underlying assumptions, and thus, conclusions that can be drawn from these methods differ as well. An overview of features from each model can be found in Table 1.

MZ twin difference method

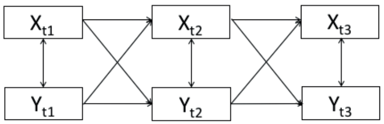
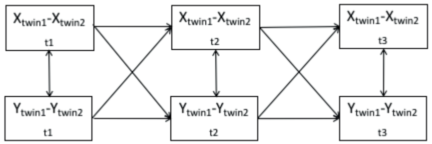
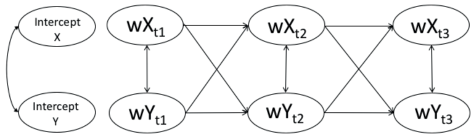
MZ twins who grow up together share 100% of their genes and many aspects of their rearing environment. As such, when twins differ on an exposure of interest, they can each act as an almost-perfect matched control for the other (McAdams, Rijsdijk, Zavos, & Pingault, 2021; Pingault et al., 2018). In the MZ twin difference method, difference scores between twins in a pair are calculated for each variable of interest (e.g. differences between twins' perceived parental discipline; and emotional and behavioural problems). These difference scores can then be used as predictors of one another (e.g. we can test whether twin differences in exposure to harsh parental discipline predict twin differences in behavioural problems). Longitudinal studies of MZ twins have demonstrated that negative parenting experiences are associated with increased child behavioural problems over time, even after accounting for potential genetic and shared environmental confounds (Burt, McGue, Iacono, & Krueger, 2006; Cecil, Barker, Jaffee, & Viding, 2012; Larsson, Viding, Rijsdijk, & Plomin, 2008; Lynch et al., 2006; Oliver, 2015; Viding, Fontaine, Oliver, & Plomin, 2009).

Twin differences in one variable prospectively predicting differences in another (e.g. as in an MZD-CLPM), indicates that any population-level prediction is not attributable to genetic confounding or environmental confounding shared by MZ twins. A limitation of the MZ twin difference method is that environmental influences not-shared between twins (ie. 'unique' or 'child-specific' environmental effects) are not accounted for and can still serve to confound associations.

The random intercept CLPM (RI-CLPM)

The MZ twin difference method can only be used on MZ twin data and so questions remain about the generalisability of results to non-twin populations. Another approach to strengthening researchers' ability to draw causal inferences that can be applied to non-twin data is the random intercept CLPM (RI-CLPM; see Table 1 for details). The random intercepts in this model accounts for stable between-person differences in each variable, thus capturing the effects of any unobserved confounding that is stable across time (Rohrer & Murayama, 2021; Usami, Murayama, & Hamaker, 2019). The residual within-person variance on each variable (which indicates variation around a person's mean) is used to evaluate whether an increase or reduction of one variable predicts an increase or reduction in another variable within-person. Cross-lagged relationships between the residual within-person variances of each variable therefore indexes within-person effects after controlling for sources of between-person variance (Hamaker et al., 2015). Controlling for stable (time-invariant) between-person differences, therefore strengthens the ability to draw causal inference from significant cross-lags in the RI-CLPM.

Table 1. Overview of model features of three CLPM models to test direction of effects in longitudinal data

	CLPM	MZD-CLPM	RI-CLPM
Model type	Cross-lagged panel model	MZ twin difference cross-lagged panel model	Random intercept cross-lagged panel model
Illustration of the model			
Model features	Cross-lagged paths are intended to inform on direction of effects between variables over time. Cross-lagged effects indicate to what extent e.g. trait X_t prospectively predicts trait Y_{t+1} controlling for autoregressive prediction from Y_t, and vice versa.	Cross-lagged paths are intended to inform on direction of effects between variables over time. Accounts for genetic and environmental effects shared by monozygotic twins by modeling relationships between twin differences in traits X and Y. Cross-lagged effects indicate to what extent within-twin pair differences in X_t prospectively predict the same in Y_{t+1}, controlling for prior differences in Y_t, and vice versa. Differences between twins can be attributed to environmental influences unique to each twin.	Cross-lagged paths are intended to inform on direction of effects between variables over time. Accounts for stable between-subject differences across development by inclusion of random intercepts. Separation of between- subject (time-invariant) and within-subject (time-variant) variation. Cross-lagged effects indicate to what extent within-person deviation from the trait level of X_t prospectively predicts change in the within-person deviation from the trait level of Y_{t+1}, controlling for prior deviation at Y_{t1}, and vice versa.
Controls for	Autoregressive effects and within-time covariance.	Autoregressive effects and within-time covariance. Any unmeasured confounders indexed by the genetic and environmental influences shared by MZ twins.	Autoregressive effects and within-time covariance. Stability within a trait (i.e. trait-like, *time-invariant* stability) and associated unobserved time-invariant confounding.

Example RQ	Are children who report higher levels of harsh parental discipline (compared to others in the sample) at age 9 also likely to show more behavioural problems (compared to others in the sample) at age 12?	Are children who report higher levels of harsh parental discipline (compared to their co-twin) at age 9 also likely to show more behavioural problems (compared to their co-twin) at age 12?	Are children who report higher deviations from their own average of harsh parental discipline at age 9 also likely to show higher deviations from their own average of behavioural problems at age 12?
Limitations	Does not control for unmeasured confounders. Does not distinguish between-person effects from within-person processes. Cannot be used to investigate whether the change in one variable between two measurement occasions is associated with change in another.	Cannot be used to investigate whether the change in one variable between two measurement occasions is associated with change in another.	Does not control for time-varying sources of confounding within individuals.
Example of conclusion that can be drawn on behavioural problems and parental discipline	"When an individual reports more behavioural problems at Tx (relative to others), (s)he will be more likely to experience a subsequent rank-order increase in harsh parental discipline at Tx+1."	"When an individual reports more behavioural problems than his/her MZ cotwin at Tx, (s)he will be more likely to experience a subsequent increase in parental discipline as compared to their cotwin at Tx +1. Moreover, any association will be explained by environmental factors unique to that child" *	"When an individual reports more behavioural problems than expected based on that individual's average (intercept), (s)he will be more likely to experience a subsequent increase in parental discipline."

Note. X=observed phenotype 'X', Y=observed phenotype 'Y', t1=first measurement occasion, t2 = second measurement occasion, t3 = third measurement occasion, Xtwin1-Xtwin2 = difference scores of X between twins in a pair, Ytwin1-Ytwin2= difference scores of Y between twins in a pair, wX= within-person deviation from the trait level of X, wY= within-person deviation from the trait level of Y, *irrespective of their rank position in the total distribution

A growing number of studies has investigated within-person associations between parenting and various developmental outcomes controlling for the between-person variance (Aunola, Tolvanen, Viljaranta, & Nurmi, 2013; Janssen, Elzinga, Verkuil, Hillegers, & Keijsers, 2021; Nelemans et al., 2020). Altough findings of between-person studies indicated longitudinal links between harsh parental discipline and emotional and behavioural problems of the child (e.g. Serbin et al., 2015), there was hardly any evidence of time-lagged within-person effects of parenting and child problems (Boele, Denissen, Moopen, & Keijsers, 2020). Studying within-person time-lagged effects, while controlling for between-person associations, may lead to a more nuanced understanding of the within-family processes of harsh parental discipline and child emotional and behavioural problems. Consequently, it could strenghten our ability to make causal inferences and prove useful for families and mental healthcare professionals (Dietvorst, Hiemstra, Hillegers, & Keijsers, 2018).

Comparing the MZD-CLPM and the RI-CLPM

The MZD-CLPM and the RI-CLPM have been proposed as ways of improving the capacity of the CLPM for drawing causal inference. Each model allows researchers to control for unobserved sources of confounding. It is well-known that identifying, measuring, and controlling for potential confounders helps to delineate cause and effect in data analysis, but that it is challenging to capture all confounding effects in this manner. As such, any method that allows researchers to control for unobserved confounding strengthens our ability to draw causal inferences. Besides being designed for different types of data (MZ twins vs singleton offspring), the MZD-CLPM and RI-CLPM also differ in the types of unobserved confounding they control for (see Table 1). Since both models control for overlapping but non-identical sources of unmeasured confounding, comparing results from these two approaches will be informative and may aid researchers in understanding how and why findings converge/diverge when using distinct methods to control for unobserved confounding. Moreover, this study will help identify which results are robust and could inform clinical practice.

Present study

The goal of the present study is to gain a better understanding of any potentially causal links between harsh parental discipline and child emotional and behavioural problems at age 9, 12 and 16 in a sample of MZ twins. To do so, we tested cross-lagged associations using the MZD-CLPM and RI-CLPM and compared findings across models. To the best of our knowledge, this is the first study to compare the RI-CLPM to MZD-CLPM within the same sample. Converging evidence from these models would serve to reinforce confidence in any putative causal relationships between harsh parental discipline and child emotional and behavioural problems.

Methods

Sample

We examined data from 5,698 monozygotic (MZ) twin pairs from the Twins Early Development Study (TEDS), an ongoing longitudinal study of twin pairs born between 1994 and 1996 in England and Wales. Participants were identified through birth records and approached for recruitment to the study (involving 16,810 families). The first TEDS data collection was conducted when twins were around 18 months of age. The sample is reasonably representative of the England and Wales population in terms of ethnic and socio-economic diversity, as well as sex and zygosity of twins (Rimfeld et al., 2019). Overall, 55% of the MZ twin sample is female, and 93% of parents identified their twins as 'white' (using a single item with response options 'Asian', 'Black', 'Mixed', 'White' and 'Other'). Participant retention in the complete TEDS sample is significantly associated with female sex, monozygosity and identifying racially as 'white'. More detailed information on TEDS can be found elsewhere (Rimfeld et al., 2019). Here we focus on child rated measures, administered when the twins were aged 9, 12 and 16.

Measures

Child problems

Emotional and behavioural problems were measured via child self-report using the strengths and difficulties questionaire (SDQ, Goodman, 1997). The SDQ is designed to assess psychological adjustment in youth aged 3–16 years. The emotional problems subscale included five statements such as 'I have many fears, I am easily scared' and 'I am nervous in new situations, I easily lose confidence'. The behavioural problems scale included five items such as 'I am often accused of lying and cheating', and 'I fight a lot, I can make other people do what I want' over the past 3 months. Ratings were on a three-point scale (with response options 'not true'=0, 'somewhat true'=1 and 'certainly true'=2). Sum scores on child self-reported emotional problems and behavioural problems of the twins at age 9, 12, and 16 years were used seperately in the analyses. Cronbach alpha's were as follows: EP9=.686, BP9 = .592, EP12 =.682, BP12 =.597, EP16=.694, BP16=.536.

Harsh parental discipline

Harsh parental discipline was assessed by four items derived from the parenting domain of a semi-structured interview (see Deater-Deckard et al., 1998). Children reported how often parents used various disciplinary strategies to deal with instances of child misbehaviour (i.e 'When I misbehave I am smacked or slapped'; 'When I misbehave Mum/Dad makes a joke out of it'). All items were answered on a 3-point scale, with the options: 'Not true'=0; 'Quite true'=1; and 'Very true'=2. The two positive items, 'When I misbehave Mum/Dad is firm and calm with me' and 'When I misbehave Mum/Dad explains why what I have done is wrong' were reversecoded, to ensure that higher scores reflect harsher parenting. Sum scores on child-reported harsh parental discipline at age 9, 12 and 16 were used for the analyses. Although Cronbach's alphas were low (9=.420, 12=.455, 16=.575), this is likely because the scale was made of only 4 items (alpha strongly depends on the number of items in a scale). Spearman's Rho correlations across years suggest that test-retest reliability is sufficient for this measure (9 -12 years: r_s=.32, 9-16 years: r_s=.15, and 12-16 years: r_s=.28; see Table 2). Also, face validity of the measure is adequate and appropriate for the aim of the study to compare findings across two statistical models.

Statistical analyses

A Cross-Lagged Panel Model (CLPM) was specified to model the relationship between harsh parental discipline and child emotional and behavioural problems at 9, 12 and 16 years. Results from this model were intended to anchor our findings within the literature, which has predominantly used a CLPM to model bi/multivariate relationships over time. We then fitted a CLPM with MZ twin differences scores (MZD-CLPM) and a RI-CLPM (Hamaker et al. 2015). Twin age and sex were regressed out of all variables and unstandardized residual scores were used in all models.

To inform the degree of correspondence between the RI-CLPM and the MZD-CLPM Spearman's Rho and R-squared between beta coefficients were calculated. The full sample of MZ twins was used for all analyses. To deal with non-independence of twins' data we added 'family ID' to control for family clustering in the CLPM and RI-CLPM. Model fitting was carried out in R using the Lavaan-package (Rosseel, 2012). We estimated all models using maximum likelihood estimation with robust standard errors for nonnormality of the data (MLR estimator) and full information maximum likelihood to deal with missing data. To evaluate model fit, we inspected Chi squared test statistics, comparative fit index (CFI), standardized root mean square residuals (SRMR) and root mean square error of approximation (RMSEA). A CFI >.90 and a RMSEA <.08 are considered good (Hu & Bentler, 1999). The analytic plan for this study was uploaded to Open Science Framework prior to analyses (https://osf.io/29s53/).

Moderation analyses

Part of our pre-registered analytic plan was to explore whether associations between harsh parental discipline and child emotional and behavioural problems differed between twins from families with high versus low levels of home chaos and socioeconomic status (SES). We ran two multiple group models, one to test home chaos and one to test SES as moderators. A more detailed description of the home chaos and SES constructs and the analytical models can be found in online supplementary materials (appendix 1).

Results

Descriptive statistics

Means, standard deviations and bivariate correlations among study variables are reported in Table 2. Spearman's Rho correlations indicated that harsh parental discipline, behavioural problems and emotional problems were significantly correlated with one another at all ages (range r_s=.12 - .33).

Main analysis

Next, we tested assocations between harsh parental discipline, behavioural problems and emotional problems using the CLPM, MZD-CLPM and RI-CLPM. Based on a comparison of model fit between constrained and unconstrained models, we specified that correlations and cross-lagged effects were allowed to vary across time in all models. Model fit indices are presented in Table 3 and show all models fitted the data adequatly. Estimated cross-lagged paths and autoregressive effects of the MZD-CLPM and RI-CLPM can be found in Table 4. Results from the CLPM are presented in appendix 1 of the online supplementary materials as baseline information. Harsh parental discipline at age 9 prospectively predicted behavioural problems at age 12 and vice versa (β=0.10, *SE*=0.02, p<.001); behavioural problems at age 9 prospectively predicted harsh parental discipline at age 12 (β=0.14, *SE*=0.02, p<.001). Also at the next occasions, age 12 and 16, harsh parental discipline (β=0.06, *SE*=0.02, p=.045) and behavioural problems (β=0.07, *SE*=0.04, p=.045) predicted each other prospectively. No cross-lagged associations between emotional problems and harsh parental discipline were found (β range -0.02 - 0.04, p>.095).

Table 2. Means, SDs and phenotypic Spearman's Rho correlations at age 9, 12 and 16

		M	*SD*	1	2	3	4	5	6	7	8	9
1	Parental discipline at 9	3.20	1.63	1								
2	Emotional problems at 9	3.22	2.40	.157**	1							
3	Behavioural problems at 9	2.21	1.85	.327**	.330**	1						
4	Parental discipline at 12	3.13	1.49	.317**	.077**	.201**	1					
5	Emotional problems at 12	2.18	2.07	.108**	.381**	.148**	.124**	1				
6	Behavioural problems at 12	1.87	1.65	.248**	.206**	.395**	.293**	.319**	1			
7	Parental discipline at 16	2.72	1.70	.153**	.055*	.130**	.276**	.035	.113**	1		
8	Emotional problems at 16	2.74	2.27	.063*	.237**	.037	.029	.378**	.083**	.120**	1	
9	Behavioural problems at 16	1.56	1.40	.125*	.069**	.210**	.143**	.124**	.329**	.180**	.200**	1

Note. ** Correlation is significant at the 0.01 level (2-tailed). * Correlation is significant at the 0.05 level (2-tailed).

Table 3. Fit indices

Model	Chi-square	df	AIC	BIC	RMSEA	CFI	TLI
CLPM	36.7	9	93221	93511	0.026	0.988	0.956
RI-CLPM	1.72	3	93193	93521	0.000	1.000	1.006
MZD-CLPM	14.13	9	96606	96896	0.011	0.983	0.932

MZD-CLPM

The MZD-CLPM is displayed in Figure 1 (estimates and 95% CI can be found in Table 4). Twin differences in behavioural problems at age 9 predicted twin differences in harsh parental discipline at age 12 (β=0.08, SE=0.03, p=0.015) but not vice versa. We also found that twin differences in harsh parental discipline at age 9 predicted twin difference in emotional problems at age 12 (β=0.08, SE=0.05, p=0.024), but not vice versa (β=-0.02, SE=0.02, p=0.526). All other cross-lagged associations were non-significant (β range [-0.03,0.04]). Autoregressions and contemporaneous associations can be found in Table 4.

RI-CLPM

The RI-CLPM is displayed in Figure 2 (estimates and 95% CI can be found in Table 4). At the between-person level, significant positive associations were found among harsh parental discipline, behavioural problems and emotional problems (PD-EP; β=0.72, SE=0.12, p=.001, PD-BP; β=0.79, SE=0.08, p<.001, EP-BP; β=0.42, SE=0.11, p=.005). This indicates that children reporting higher levels of harsh parental discipline across all ages reported more behavioural problems and emotional problems as well. At the within-person level, behavioural problems at age 9 were associated with harsh parental discipline at age 12 (β=0.11, SE=0.03, p=0.004) but not vice versa (β=0.07, SE=0.03, p=0.062). This means that, when children report more behavioural problems at age 9 than their average, they prospectively report more harsh parental discipline at age 12. We did not find significant within-person cross-lagged paths between harsh parental discipline at age 9 and emotional problems at age 12 (β= -0.04, SE=0.05, p=0.410) nor vice versa (β= -0.06, SE=0.03, p=0.123). From age 12 to 16, we found that emotional problems were predictive of decreased harsh parental discipline at age 16 (β= -0.10, SE=0.04, p=0.021). In other words, when children report more emotional problems at age 12 than expected (based on that individual's average), they report less harsh parental discipline at age 16 than expected. All the other cross-lagged associations between harsh parental discipline and child problems were non-significant (β range [-0.10, 0.07]. Autoregressions and contemporaneous correlations can be found in Table 4.

Comparing the MZD-CLPM and RI-CLPM

We calculated correlations between the beta coefficients estimated in the MZD-CLPM and RI-CLPM. Spearman's Rho and R-squared (r_s=.594, p=.009, R2 =0.298) indicated that the coefficients from each model were significantly correlated (see Figure 4). However, focusing only on the cross-lagged estimates this was not the case, r_s=.273, p=.391, R2 = 0.031.

Table 4. Parameter Estimates

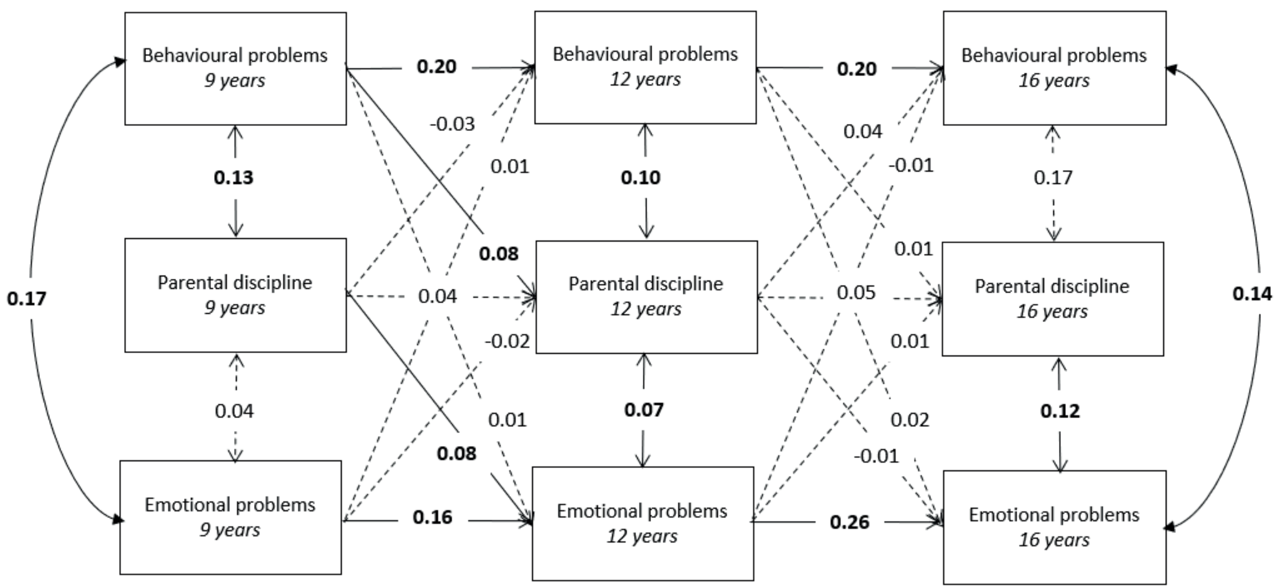
	MZD-CLPM				RI-CLPM			
Parameter	Estimate	SE	p-value	Std beta (95% C)	Estimate	SE	p-value	Std beta (95% CI)
Between-person effects								
PD <-> EM					0.40	0.12	**0.001**	0.72 (0.15, 1.30)
PD <-> BE					0.30	0.08	**<.001**	0.80 (0.39, 1.20)
EM <-> BE					0.32	0.11	**0.005**	0.42 (0.18, 066)
Cross-lagged effects								
PD9 -> EM12	0.11	0.05	**0.024**	0.08 (0.01, 0.14)	-0.04	0.05	0.410	-0.04 (-0.12, 0.05)
PD9 -> BE12	-0.03	0.04	0.435	-0.03 (-0.09, 0.04)	0.06	0.03	0.062	0.07 (0.14, -0.03)
EM9 -> PD12	-0.01	0.02	0.526	-0.02 (-0.09, 0.04)	-0.04	0.03	0.123	-0.06 (-0.14, 0.02)
EM9 -> BE12	0.01	0.02	0.690	0.01 (-0.05, 0.08)	0.09	0.03	**0.001**	0.12 (0.05, 0.20)
BE9 -> PD12	0.07	0.03	**0.015**	0.08 (0.02, 0.14)	0.09	0.03	**0.004**	0.11 (0.03, 0.19)
BE9 -> EM12	0.01	0.04	0.715	0.01 (-0.05, 0.08)	0.12	0.05	**0.009**	0.11 (0.03, 0.20)
PD12 -> EM16	-0.02	0.06	0.797	-0.01 (0.09, 0.07)	-0.13	0.07	0.052	-0.10 (-0.19, 0.00)
PD12 -> BE16	0.04	0.04	0.257	0.04 (-0.03, 0.11)	-0.03	0.05	0.484	-0.04 (-0.15, 0.07)
EM12 -> PD16	0.01	0.03	0.735	0.01 (-0.07, 0.10)	-0.09	0.04	**0.021**	-0.10 (-0.19, -0.02)
EM12 -> BE16	-0.01	0.03	0.850	-0.01 (-0.09, 0.07)	-0.04	0.04	0.287	-0.07 (-0.19, 0.05)
BE12 -> PD16	0.01	0.05	0.808	0.01 (-0.08, 0.10)	0.03	0.06	0.515	0.03 (-0.06, 012)
BE12 -> EM16	0.02	0.06	0.697	0.02 (-0.07, 0.10)	0.06	0.06	0.953	0.05 (-0.05, 0.14)

Table continues on the next page

Contemporaneous associations								
EM9 <-> BE9	0.80	0.16	**<.001**	0.17 (0.10, 0.24)	1.15	0.13	**<.001**	0.32 (0.26, 0.38)
EM9 <-> PD9	0.16	0.13	0.207	0.04 (-0.02, 0.10)	0.26	0.13	0.065	0.08 (-0.01, 0.16)
BE9 <-> PD9	0.40	0.10	**<.001**	0.13 (0.07, 0.20)	0.66	0.10	**<.001**	0.26 (0.19, 0.32)
EM12 <-> BE12	0.83	0.10	**<.001**	0.27 (0.17, 0.28)	0.70	0.09	**<.001**	0.29 (0.22, 0.36)
EM12 <-> PD12	0.22	0.09	**0.010**	0.07 (0.02, 0.12)	0.03	0.09	0.774	0.01 (-0.07, 0.09)
BE12 <-> PD12	0.25	0.06	**<.001**	0.10 (0.05, 0.15)	0.31	0.06	**<.001**	0.17 (0.11, 0.23)
EM 16 <-> PD16	0.48	0.18	**0.007**	0.12 (0.03, 0.20)	0.17	0.11	0.142	0.06 (-0.02, 0.13)
EM16 <-> BE16	0.46	0.14	**0.001**	0.14 (0.06, 0.21)	0.28	0.10	**0.004**	0.14 (0.05, 0.24)
BE16 <-> PD16	0.15	0.16	0.122	0.17 (-0.02, 0.13)	0.15	0.08	**0.049**	0.09 (0.01, 0.17)
Autoregressions								
PD9 -> PD12	0.04	0.03	0.271	0.04 (-0.03, 0.11)	0.21	0.05	**<.001**	0.23 (0.13, 0.34)
PD12-->PD16	0.06	0.05	0.254	0.05 (-0.04, 0.13)	0.24	0.07	**<.001**	0.20 (0.09, 0.32)
EM9 ->EM12	0.14	0.03	**<.001**	0.16 (0.09, 0.23)	0.16	0.05	**<.001**	0.19 (0.09, 0.30)
EM12-> EM16	0.27	0.05	**<.001**	0.26 (0.17, 0.34)	0.18	0.07	**0.010**	0.17 (0.04, 0.30)
BE9 -> BE12	0.18	0.03	**<.001**	0.20 (0.12, 0.27)	0.20	0.04	**<.001**	0.23 (0.15, 0.32)
BE12 -> BE16	0.18	0.04	**<.001**	0.20 (0.12, 0.28)	0.11	0.05	**0.020**	0.16 (0.03, 0.28)

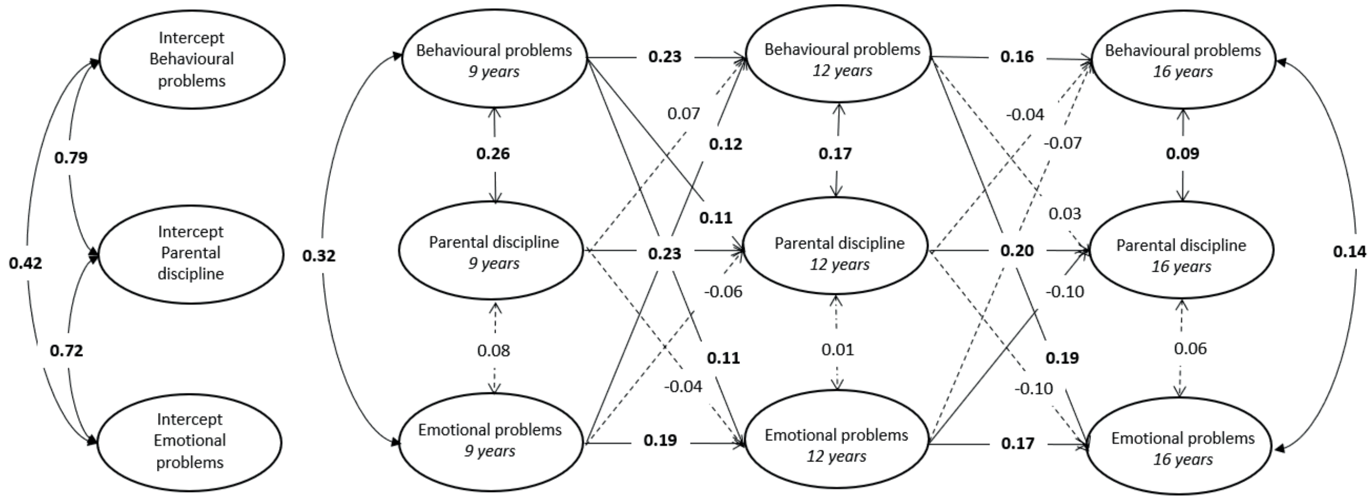
PD = harsh parental discipline, EP = emotional problems, BP = behavioural problems p-values in bold are significant (p<.05). Confidence intervals of the standardized estimates were simple symmetric 95% CIs.

Figure 1. Representation of MZD-CLPM with standardized effects (β).



Note. Covariance between behavioural problems and emotional problems at age 12 is missing in the depiction = 0.27. Estimates in bold /solid lines are significant ($p<.05$).

Figure 2. Simplified representation of the unconstrained RI-CLPM with standardized effects (β).



Note. Covariance between behavioural problems and emotional problems at age 12 is missing in the depiction = 0.29. Estimates in bold /solid lines are significant ($p<.05$).

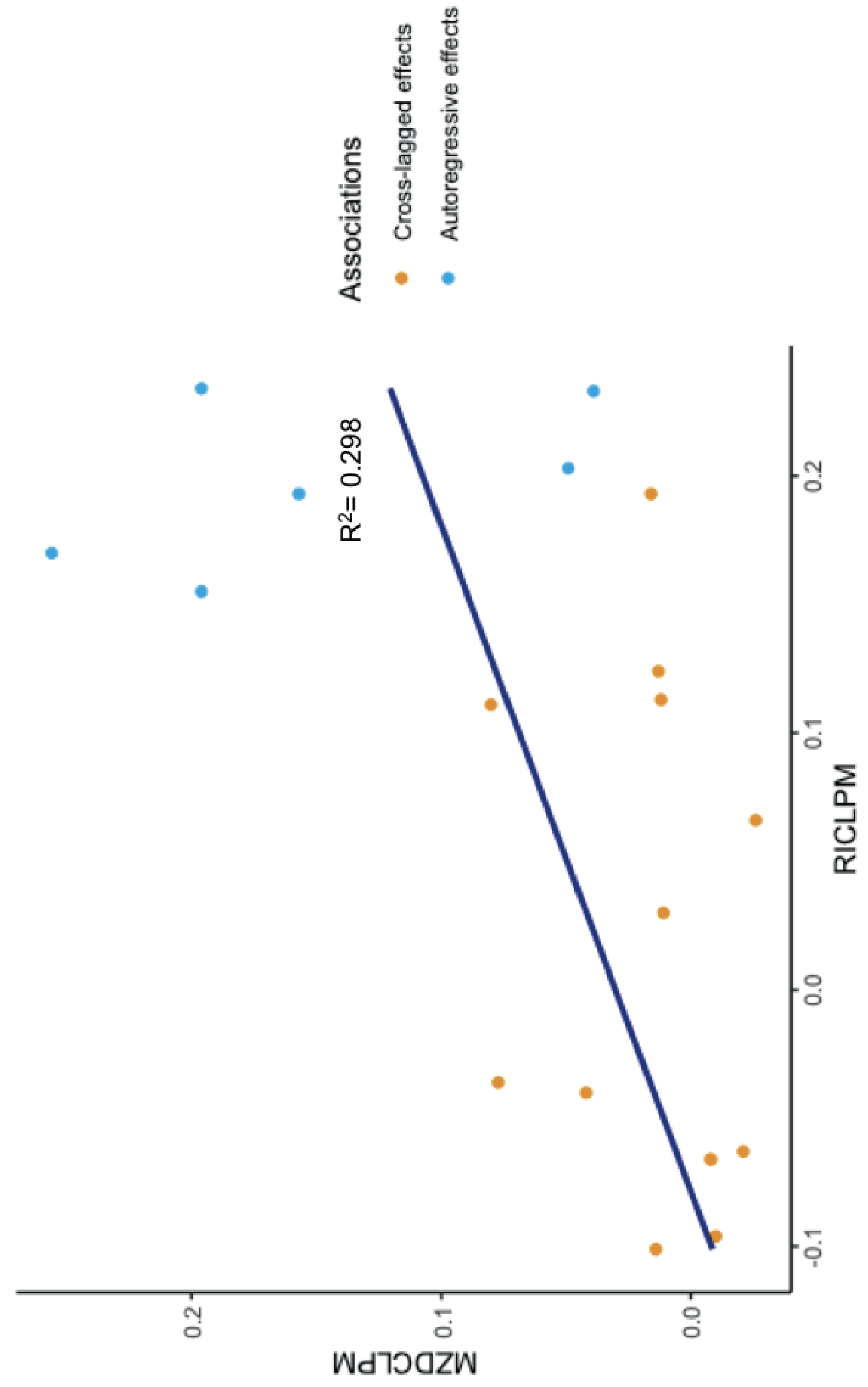


Figure 3. Beta coefficients of the regression effects from the MZD-CLPM and RI-CLPM.

Exploratory post-hoc analyses

We explored whether the associations between harsh parental discipline and child emotional problems and behavioural problems differed between twins from families with high versus low levels of home chaos and high versus low levels of SES. This was not the case in our study. Results from these analyses can be found in online supplementary materials (appendix 2).

Discussion

In this study we investigated longitudinal relationships between harsh parental discipline, and child emotional and behavioral problems, triangulating evidence across two extensions of the cross-lagged panel model (CLPM): the monozygotic twin difference cross-lagged panel model (MZD-CLPM) and the random intercept CLPM (RI-CLPM). One longitudinal cross-lagged association was significant in all three models: children's behavioural problems at age 9 predicted increased harsh parental discipline at age 12. The converging evidence from the MZD-CLPM and RI-CLPM strengthens our confidence in a potentially causal relationship between child behavioural problems at age 9 and consequent experienced harsh parental discipline at age 12. For all other associations there was a lack of convergence across models, as also illustrated by the small correlations between cross-lagged parameters.

Why do findings from the MZD-CLPM and RI-CLPM not converge?

A comparison of beta coefficients from the RI-CLPM and MZD-CLPM (Figure 4) indicated a similar pattern of results from both models: beta coefficients were positively correlated, suggesting that, on average, associations were in the same direction. Nonetheless, substantial differences in the magnitude of associations estimated in each model were also evident. One noteworthy difference between models was that the link from harsh parental discipline at age 9 to emotional problems at age 12 was non-significant in the RI-CLPM, whilst results from the MZD-CLPM indicated that perceiving more harsh parental discipline at age 9 than one's sibling was related to having more emotional problems at age 12. Differences between RI-CLPM and MZD-CLPM likely exist because the two approaches control for overlapping, but not identical, confounding influences. The MZD-CLPM controls for all genetic factors and all environmental influences shared between MZ twins, these can be both time-invariant (stable) and time-varying (unstable). The random intercept of the RI-CLPM, however, controls only for time-invariant (stable) effects, which could entail genetic and environmental influences, but does not account for time-variant (unstable) within-person confounding effects, see also Table 1 for an overview. Although genetic influences on traits are generally considered stable, evidence shows some change over time, certainly across childhood and adolescence (Hannigan et al., 2017). Broadly speaking, the MZD-CLPM design can therefore be seen as a more stringent approach compared to the RI-CLPM, given that the MZD-

CLPM controls for both time-varying and time-invariant shared effects, while the RI-CLPM accounts for time-invariant stability only. Although it should be noted that the RI-CLPM does account for all stable environmental influences, whereas the MZD-CLPM only accounts for environmental influences shared between twins. Our results underline the important nuances in conclusions that can be drawn from each of these these distinct and complementary approaches (Lüdtke & Robitzsch, 2021, and see also Table 1).

Associations between traits that change over time, or traits that are only in part genetically linked (e.g. parental discipline, child emotional and behavioural problems) may substantially vary across the MZD-CLPM and RI-CLPM. In the RI-CLPM, we are unable to capture developmental changes of the child and their environment, as the model controls for time-invariant confounding only. It could also be expected that the models might converge more when modelling associations among traits in adulthood, when we might expect genetic and environmental effects to be more stable. As a result, findings from the MZD-CLPM and RI-CLPM may be more similar when modelling traits that are stable over time and under stable genetic influence (e.g. impulsivity; Niv, Tuvblad, Raine, Wang, & Baker, 2012).

Strengths, limitations and future studies

Some limitations to our study should be noted. First, we like to underline that we did find evidence for contemporaneous associations between harsh parental discipline and child emotional and behavioural problems. So, we do not know whether non-significant findings regarding the time-lagged effects indicate absence of longitudinal associations or that, for instance, the effects of harsh parental discipline and child problems unfold on a shorter time scale (Bolger & Laurenceau, 2013). The long intervals, i.e. three and four years between measurement occasions, may have limited our ability to capture processes between parents and children that might unfold on monthly, daily, weekly or even hourly basis (Granic & Patterson, 2006). Future studies are needed to address the dynamic processes between harsh parental discipline and child emotional and behavioural problems on different timescales. Moreover, it should be noted that within-person effects in the RI-CLPM are based on scores that capture temporary fluctuations around a person’s mean. As such, the RI-CLPM may be a more appropriate method to answering questions about short-term within-person effects using time series data and/or intensive longitudinal designs than for understanding long-term changes in longitudinal data with less measurements and long intervals (Lüdtke & Robitzch; 2021; Orth, Clark, Donnellan, & Robins, 2021). Second, it should be noted that in comparing the MZD-CLPM cross-lagged path estimates with those of the RI-CLPM we assessed the correlation among 12 beta estimates. Low statistical power due to the low number of data points may have contributed to the non-significance of the correlation. As such, the comparison of the models by assessing correlations among estimates, as depicted in Figure 3, has an illustrative purpose. Lastly, our measure of parental discipline scored low on Cronbach's alpha, indicating low reliabilities. The low number of items might be a reason for the low alpha. However, test-retest reliability seemed sufficient and the measure correlated with emotional and behavioural prbolems as predicted, so it can be assumed that the harsh parental discipline scale measured a stable trait-like phenotype. The analytical models could be replicated with higher number of parental discipline items or with a latent factor model. It is expected that our findings are still robust and reliable, as face validity is reasonable and the low number of heterogeneous parenting items are appropriate for the aim of the current study.

Conclusions and recommendations

Taken together, our study described the types of unmeasured confounding different longitudinal designs can account for, and underlined how slight differences in the sort of confounding being controlled for can lead to quite different conclusions, even when using the exact same data (see Table 1). Findings from all models, i.e. the CLPM, RI-CLPM and MZD-CLPM, indicated that child behavioural problems at age 9 predicted increased harsh parental discipline at age 12. These results can be interpreted as corroborating (albeit not conclusive) evidence in favour of a causal relationship.

Importantly however, results also illustrate divergence in the MZD-CLPM and RI-CLPM outcomes. While both methods are intended to improve the ability of researchers to draw causal inference, they do not lead to the same conclusions. The substantial differences in results underline that nuance in conclusions is required when interpreting findings using such models and that triangulating results across multiple (longitudinal) methods strengthens the ability to draw conclusions on causality.

Conflict of interest. All authors declare that they have no conflicts of interest.

References

Aunola, K., Tolvanen, A., Viljaranta, J., & Nurmi, J.-E. (2013). Psychological control in daily parent–child interactions increases children's negative emotions. *Journal of Family Psychology*, *27*(3), 453–462. https://doi.org/10.1037/a0032891

Baumrind, D. (1991). The Influence of Parenting Style on Adolescent Competence and Substance Use. *The Journal of Early Adolescence*, *11*(1), 56–95. https://doi.org/10.1177/0272431691111004

Boele, S., Denissen, J., Moopen, N., & Keijsers, L. (2020). Over-time Fluctuations in Parenting and Adolescent Adaptation Within Families: A Systematic Review. *Adolescent Research Review*, *5*(3), 317–339. https://doi.org/10.1007/s40894-019-00127-9

Bolger, N., & Laurenceau, J.-P. (2013). *Intensive longitudinal methods: An introduction to diary and experience sampling research. - PsycNET*. Retrieved from https://psycnet.apa.org/record/2012-17340-000

Burt, S. A., McGue, M., Iacono, W. G., & Krueger, R. F. (2006). Differential parent-child relationships and adolescent externalizing symptoms: Cross-lagged analyses within a monozygotic twin differences design. *Developmental Psychology*, *42*(6), 1289–1298. https://doi.org/10.1037/0012-1649.42.6.1289

Cecil, C. A. M., Barker, E. D., Jaffee, S. R., & Viding, E. (2012). Association between maladaptive parenting and child self-control over time: Cross-lagged study using a monozygotic twin difference design. *British Journal of Psychiatry*, *201*(4), 291–297. https://doi.org/10.1192/bjp.bp.111.107581

Deater-Deckard, K., Dodge, K. A., Bates, J. E., & Pettit, G. S. (1998). Multiple risk factors in the development of externalizing behavior problems: Group and individual differences. *Development and Psychopathology, 10*(3), 469–493.

Dietvorst, E., Hiemstra, M., Hillegers, M. H. J., & Keijsers, L. (2018). Adolescent Perceptions of Parental Privacy Invasion and Adolescent Secrecy: An Illustration of Simpson's Paradox. *Child Development*, *89*(6), 2081–2090. https://doi.org/10.1111/cdev.13002

Fletcher, A. C., Steinberg, L., & Williams-Wheeler, M. (2004). Parental Influences on Adolescent Problem Behavior: Revisiting Stattin and Kerr. *Child Development*, *75*(3), 781–796. https://doi.org/10.1111/j.1467-8624.2004.00706.x

Goodman, R. (1997). The Strengths and Difficulties Questionnaire: A Research Note. *Journal of Child Psychology and Psychiatry*, *38*(5), 581–586. https://doi.org/10.1111/j.1469-7610.1997.tb01545.x

Granic, I., & Patterson, G. R. (2006). Toward a comprehensive model of antisocial development: A dynamic systems approach. *Psychological Review*, *113*(1), 101–131. https://doi.org/10.1037/0033-295X.113.1.101

Hamaker, E. L., Kuiper, R. M., & Grasman, R. P. P. P. (2015). A critique of the cross-lagged panel model. *Psychological Methods*, *20*(1), 102–116.

https://doi.org/10.1037/a0038889
Hannigan, L. J., Walaker, N., Waszczuk, M. A., McAdams, T. A., & Eley, T. C. (2017). Aetiological Influences on Stability and Change in Emotional and Behavioural Problems across Development: A Systematic Review. *Psychopathology Review*, *a4*(1), 52–108. https://doi.org/10.5127/pr.038315
Hipwell, A., Keenan, K., Kasza, K., Loeber, R., Stouthamer-Loeber, M., & Bean, T. (2008). Reciprocal Influences Between Girls' Conduct Problems and Depression, and Parental Punishment and Warmth: A Six year Prospective Analysis. *Journal of Abnormal Child Psychology*, *36*(5), 663–677. https://doi.org/10.1007/s10802-007-9206-4
Hu, L., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling: A Multidisciplinary Journal*, *6*(1), 1–55. https://doi.org/10.1080/10705519909540118
Janssen, L. H. C., Elzinga, B. M., Verkuil, B., Hillegers, M. H. J., & Keijsers, L. (2021). The Link between Parental Support and Adolescent Negative Mood in Daily Life: between-Person Heterogeneity in within-Person Processes. *Journal of Youth and Adolescence*, *50*(2), 271–285. https://doi.org/10.1007/s10964-020-01323-w
Lansford, J. E., Criss, M. M., Laird, R. D., Shaw, D. S., Pettit, G. S., Bates, J. E., & Dodge, K. A. (2011). Reciprocal relations between parents' physical discipline and children's externalizing behavior during middle childhood and adolescence. *Development and Psychopathology*, *23*(1), 225–238. https://doi.org/10.1017/S0954579410000751
Larsson, H., Viding, E., Rijsdijk, F. V., & Plomin, R. (2008). Relationships between parental negativity and childhood antisocial behavior over time: A bidirectional effects model in a longitudinal genetically informative design. *Journal of Abnormal Child Psychology*, *36*(5), 633–645. https://doi.org/10.1007/s10802-007-9151-2
Lüdtke, O., & Robitzsch, A. (n.d.). *A Critique of the Random Intercept Cross-Lagged Panel Model*. https://doi.org/10.31234/OSF.IO/6F85C
Lynch, S. K., Turkheimer, E., D'onofrio, B. M., Mendle, J., Emery, R. E., Slutske, W. S., & Martin, N. G. (2006). *A Genetically Informed Study of the Association Between Harsh Punishment and Offspring Behavioral Problems*. https://doi.org/10.1037/0893-3200.20.2.190
McAdams, T. A., Rijsdijk, F. V, Zavos, H. M. S., & Pingault, J.-B. (2021). Twins and Causal Inference: Leveraging Nature's Experiment. *Cold Spring Harbor Perspectives in Medicine*, *11*(6), a039552. https://doi.org/10.1101/cshperspect.a039552
McKee, L., Colletti, C., Rakow, A., Jones, D. J., & Forehand, R. (2008). Parenting and child externalizing behaviors: Are the associations specific or diffuse?

Aggression and Violent Behavior, *13*(3), 201–215. https://doi.org/10.1016/J.AVB.2008.03.005
Moscati, A., Verhulst, B., McKee, K., Silberg, J., & Eaves, L. (2018). Cross-Lagged Analysis of Interplay Between Differential Traits in Sibling Pairs: Validation and Application to Parenting Behavior and ADHD Symptomatology. *Behavior Genetics*, *48*(1), 22–33. https://doi.org/10.1007/s10519-017-9882-y
Nelemans, S. A., Keijsers, L., Colpin, H., Leeuwen, K., Bijttebier, P., Verschueren, K., & Goossens, L. (2020). Transactional Links Between Social Anxiety Symptoms and Parenting Across Adolescence: Between- and Within-Person Associations. *Child Development*, *91*(3), 814–828. https://doi.org/10.1111/cdev.13236
Niv, S., Tuvblad, C., Raine, A., Wang, P., & Baker, L. A. (2012). Heritability and Longitudinal Stability of Impulsivity in Adolescence. *Behavior Genetics*, *42*(3), 378–392. https://doi.org/10.1007/s10519-011-9518-6
Nosek, B. A., Ebersole, C. R., Dehaven, A. C., Mellor, D. T., & Shiffrin, R. M. (2018). *The preregistration revolution*. *115*(11), 2600–2606. https://doi.org/10.1073/pnas.1708274114
Oliver, B. R. (2015). Unpacking externalising problems: negative parenting associations for conduct problems and irritability. *BJPsych Open*, *1*(1), 42–47. https://doi.org/10.1192/bjpo.bp.115.000125
Orth, U., Clark, D. A., Donnellan, M. B., & Robins, R. W. (2021). Testing prospective effects in longitudinal research: Comparing seven competing cross-lagged models. *Journal of Personality and Social Psychology*, *120*(4), 1013–1034. https://doi.org/10.1037/pspp0000358
Pingault, J.-B., O'Reilly, P. F., Schoeler, T., Ploubidis, G. B., Rijsdijk, F., & Dudbridge, F. (2018). Using genetic data to strengthen causal inference in observational research. *Nature Reviews. Genetics*, *19*(9), 566–580. https://doi.org/10.1038/s41576-018-0020-3
Pinquart, M. (2017a). Associations of parenting dimensions and styles with externalizing problems of children and adolescents: An updated meta-analysis. *Developmental Psychology*, *53*(5), 873–932. https://doi.org/10.1037/dev0000295
Pinquart, M. (2017b). Associations of Parenting Dimensions and Styles with Internalizing Symptoms in Children and Adolescents: A Meta-Analysis. *Marriage & Family Review*, *53*(7), 613–640. https://doi.org/10.1080/01494929.2016.1247761
Rimfeld, K., Malanchini, M., Spargo, T., Spickernell, G., Selzam, S., McMillan, A., … Plomin, R. (2019). Twins Early Development Study: A Genetically Sensitive Investigation into Behavioral and Cognitive Development from Infancy to Emerging Adulthood. *Twin Research and Human Genetics*, *22*(6), 508–513. https://doi.org/10.1017/thg.2019.56
Ritchie, S. J., Bates, T. C., & Plomin, R. (2015). Does Learning to Read Improve

Intelligence? A Longitudinal Multivariate Analysis in Identical Twins From Age 7 to 16. *Child Development*, *86*(1), 23–36. https://doi.org/10.1111/cdev.12272
Rohrer, J., & Murayama, K. (2021). *These are not the effects you are looking for: Causality and the within-/between-person distinction in longitudinal data analysis*. Retrieved from https://psyarxiv.com/tg4vj/download?format=pdf
Serbin, L. A., Kingdon, D., Ruttle, P. L., & Stack, D. M. (2015). The impact of children's internalizing and externalizing problems on parenting: Transactional processes and reciprocal change over time. *Development and Psychopathology*, *27*(4), 969–986. https://doi.org/10.1017/S0954579415000632
Usami, S., Murayama, K., & Hamaker, E. L. (2019). A unified framework of longitudinal models to examine reciprocal relations. *Psychological Methods*, *24*(5), 637–657. https://doi.org/10.1037/met0000210
Viding, E., Fontaine, N. M. G., Oliver, B. R., & Plomin, R. (2009). Negative parental discipline, conduct problems and callous–unemotional traits: monozygotic twin differences study. *British Journal of Psychiatry*, *195*(5), 414–419. https://doi.org/10.1192/bjp.bp.108.061192
Wang, M.-T., & Kenny, S. (2014). Longitudinal Links Between Fathers' and Mothers' Harsh Verbal Discipline and Adolescents' Conduct Problems and Depressive Symptoms. *Child Development*, *85*(3), 908–923. https://doi.org/10.1111/cdev.12143
Yves Rosseel (2012). lavaan: An R Package for Structural Equation Modeling. *Journal of Statistical Software*, 48(2), 1-36.