

Mental health difficulties and academic attainment: Evidence for gender-specific developmental cascades in middle childhood

MARGARITA PANAYIOTOU AND NEIL HUMPHREY

University of Manchester

Abstract

We present a developmental cascade model of the longitudinal relationships between internalizing symptoms, externalizing problems, and academic performance in middle childhood, utilizing a large sample ($N = 1,771$) from the United Kingdom in a 3-year, cross-lag design. Three hypotheses were tested: adjustment erosion, academic incompetence, and (cumulative) shared risk. In addition, we sought to examine whether developmental cascade pathways varied across gender, while also statistically exploring indirect, mediation pathways. Structural equation models that accounted for within-time covariance, data nesting, and temporal stability provided evidence of gender-specific effects as follows: externalizing-attainment adjustment erosion pathways were found only in boys, while attainment-internalizing/externalizing academic incompetence pathways were found only in girls. Analysis of mediation pathways provided further support for these gender-specific longitudinal profiles. Protective longitudinal internalizing-externalizing and externalizing-internalizing pathways were found for both boys and girls. Finally, while it improved model fit for both genders, the influence of cumulative shared risk on the aforementioned pathways was relatively meager, substantively affecting only one (externalizing-attainment adjustment erosion pathway in boys). The implications of these findings are discussed, and study limitations noted.

Throughout the life span, success or failure within one domain of development may have a significant impact on future development in other domains (Masten, Burt, & Coatsworth, 2015). Researchers have begun to explore these bidirectional influences and progressive effects (Masten et al., 2005), with consequent implications for the advancement of developmental theory and prevention science (Masten et al., 2015). The consequences of interactions that spread across time and different domains of development through either direct or indirect pathways are known as “developmental cascades” (Masten & Cicchetti, 2010).

Developmental cascades research can explain how problems in one domain of functioning can spread into other areas of functioning, while also identifying the steps along that progression (Ehrlich, Miller, & Chen, 2016). In recent years, there has been increasing interest in the pathways between domains of mental health and academic attainment (Bornstein, Hahn, & Suwalsky, 2013; Burt & Roisman, 2010; Masten et al., 2005; Moilanen, Shaw, & Maxwell, 2010; Vaillancourt, Brittain, McDougall, & Duku, 2013). In the current study we examine these longitudinal associations during middle childhood, a critical period in which developmental and contextual changes take place that provide important foundations for adolescence and adult life (Huston & Ripke, 2006). Of particular note given

the focus of our paper is children’s increasingly complex and differentiated sense of identity and concept of self, “a sense of who they are and where they fit in the larger scheme of things . . . [and the formation of] concepts of what they do well and what they do not” (Huston & Ripke, p. 9), during middle childhood. This period also marks a significantly increased emphasis (from parents, teachers, and others) on the cognitive, social, and behavioral capabilities needed for positive adaptation in the school context, and the cultural capital attributed to them. Finally, middle childhood sees the continuation and intensification of gender socialization practices. Drawing these strands together, this period is one in which we would expect to see increasing interconnectivity between children’s academic performance and their mental health as their sense of self-worth becomes increasingly internalized from others’ feedback in the socializing environments of the school and home (Harter, 1999); however, there is also good reason to expect these interconnections to vary across gender given the differential expectations that are set in relation to emotional expression, academic motivation and achievement, and other pertinent domains of development.

There are generally three substantive theories that are used to explain how academic attainment and mental health interact with one another over time. We review each briefly below, before returning to the aim of the current study.

Adjustment Erosion Hypothesis

The adjustment erosion hypothesis posits that externalizing problems and/or internalizing symptoms lead to later academic

This research was supported by grants from the National Institute for Health Research.

Address correspondence and reprint requests to: Margarita Panayiotou, Manchester Institute of Education, University of Manchester, Manchester M13 9PL, UK; E-mail: margarita.panayiotou@manchester.ac.uk.

difficulties (Moilanen et al., 2010). For instance, externalizing problems (such as disruptive behaviors) can interfere with academic progress due to acceptance by more deviant peers (Schwartz, Gorman, Nakamoto, & McKay, 2006) and/or the disruption of interpersonal relationships in the classroom (Moilanen et al., 2010). In a similar way, internalizing symptoms (such as anxiety and withdrawal) may impede academic achievement by interfering with classroom participation (Roeser, van der Wolf, & Strobel, 2001). Despite some inconsistencies in the literature (e.g., Vaillancourt et al., 2013; Weeks et al., 2016; Zhang, 2013), there is a robust body of evidence indicating that early externalizing symptoms erode later academic performance (Ansary & Luthar, 2009; Burt & Roisman, 2010; Masten et al., 2005; Moilanen et al., 2010; Obradovic, Burt, & Masten, 2010; Salla et al., 2016; Stipek & Miles, 2008; Weeks et al., 2016; Wigelsworth, Qualter, & Humphrey, 2016). Although less well developed, a growing number of studies have also found support for the adjustment erosion hypothesis in terms of internalizing symptoms (Englund & Siebenbruner, 2012; Grover, Ginsburg, & Ialongo, 2007; Obradovic et al., 2010; Weeks et al., 2016).

The relationship between externalizing problems and internalizing symptoms has also been of interest to researchers. Although there is sufficient evidence to suggest that these domains of development are interrelated (Bornstein et al., 2013; Burt & Roisman, 2010; Englund & Siebenbruner, 2012; Moilanen et al., 2010; Weeks et al., 2016), findings from longitudinal data have been inconsistent. Some studies have found that initial externalizing problems exacerbate internalizing symptoms at later time points (Bornstein et al., 2013; Capaldi, 1992; Lahey, Loeber, Burke, Rathouz, & McBurnett, 2002; Moilanen et al., 2010; Vaillancourt et al., 2013). Consistent with the failure theory (Patterson & Capaldi, 1990) and related research (Kiesner, 2002), it is possible that problematic behaviors lead to failures in personal relationships and tasks, which in turn results in the development or maintenance of internalizing problems, such as depression.

By contrast, the influence of early internalizing symptoms on later externalizing problems has been found to be “protective” (i.e., high internalizing symptoms predict lower externalizing problems) in several studies (Bornstein, Hahn, & Haynes, 2010; Bornstein et al., 2013; Englund & Siebenbruner, 2012; Masten et al., 2005; Mesman, Bongers, & Koot, 2001; Rogosch, Oshri, & Cicchetti, 2010). The experience of internalizing symptoms may interrupt the trajectories from risk to disruptive behaviors, perhaps due to the increased self-isolation and withdrawal associated with the former (Masten et al., 2005). Children experiencing symptoms of anxiety and shyness are expected to exhibit fewer dangerous and disruptive behaviors (Mesman et al., 2001).

Academic Incompetence Hypothesis

In contrast to the adjustment erosion hypothesis, the academic incompetence model claims that initial problems relating to

academic performance can trigger or exacerbate internalizing symptoms and externalizing problems (Moilanen et al., 2010). Children and adolescents who struggle academically can experience considerable frustration and disaffection (Wigelsworth et al., 2016), while their placement in classes with other low-achieving children can expose them to social stigma and peers that may reinforce disruptive and aggressive behaviors (Masten et al., 2015).

Although findings vary across different developmental periods, there is emerging consensus that poor initial academic performance leads to elevated externalizing (Ansary & Luthar, 2009; Burt & Roisman, 2010; Englund & Siebenbruner, 2012; Moilanen et al., 2010; Morgan, Farkas, Tufis, & Sperling, 2008; Vaillancourt et al., 2013; Weeks et al., 2016; Wigelsworth et al., 2016) and internalizing (Burt & Roisman, 2010; Englund & Siebenbruner, 2012; Masten et al., 2005; Morgan et al., 2008; Obradovic et al., 2010; Weeks et al., 2016) symptomatology. Such findings are consistent with the notion that salient developmental domains (such as academic performance) are used to make judgments about others or oneself (Masten et al., 2005). Therefore, failure in such domains could potentially lead to maladaptive outcomes over time.

Shared Risk Hypothesis

The shared risk hypothesis suggests that the cascading effects central to the other two models may be the product of a “third variable” that affects multiple domains of development (Moilanen et al., 2010). Risk factors such as low socioeconomic status, poor parenting, and low intellectual ability have been found to account for some of the associations noted above (Masten et al., 2005; McLoyd, 1998). Such factors may place the individual on a path that promotes cascading pathways between academic competence and domains of child and adolescent functioning across time (Moilanen et al., 2010). Several developmental cascades studies have controlled for at least some of these confounders including, but not limited to, intellectual ability, parental influences, socioeconomic status, and stressful life events (Bornstein et al., 2013; Burt & Roisman, 2010; Englund & Siebenbruner, 2012; Masten et al., 2005; Moilanen et al., 2010; Vaillancourt et al., 2013; Weeks et al., 2016; Wigelsworth et al., 2016).

The Current Study

The aim of the current study was to examine the longitudinal relationships between academic attainment, externalizing problems, and internalizing symptoms during middle childhood. We set out to build upon and extend findings from existing research by addressing several important gaps, limitations, and inconsistencies in the existing developmental cascades literature. First, work examining the associations noted above has been conducted primarily with samples from North America. Cultural transferability of findings cannot be assumed (Deighton et al., *in press*). Second, studies in the field have generally utilized relatively small sample sizes to date (often

fewer than 300; e.g., see Masten et al., 2005; Moilanen et al., 2010). Aside from core issues of limited representativeness and generalizability, these samples are often insufficient to accommodate the modeling complexity required in developmental cascades research, reducing the accuracy with which tests are able to detect nonzero paths (Masten et al., 2005), and restricting researchers' ability to incorporate additional parameters (Bornstein et al., 2013), account for data nesting, or estimate correlated errors (Deighton et al., *in press*).

Third, exploration of mediation pathways between domains across time has been limited to date. To test mediation models within developmental cascades, it is optimal to have three or more constructs measured across three or more time points (Cole & Maxwell, 2003). Of the studies assessing academic attainment, externalizing problems, and internalizing symptoms across three waves, several have portrayed the pathways in terms of direct and indirect effects (Masten et al., 2005; Weeks et al., 2016), but only a couple have statistically assessed them (Moilanen et al., 2010; Vaillancourt et al., 2013).

Fourth, the potentially critical role of gender as a moderator of developmental cascades pathways remains underexplored. This is in spite of clear evidence of gender differences in both mental health difficulties (Green, McGinnity, Meltzer, Ford, & Goodman, 2005) and academic attainment (Department for Education, 2011). More important, from a theoretical perspective, gendered socialization practices relating to emotional expression (Chaplin, Cole, & Zahn-Waxler, 2005), academic motivation and achievement (Leaper & Friedman, 2007), and other pertinent domains could feasibly trigger gender-specific developmental cascade pathways, either directly or indirectly (e.g., via effects on stress response and coping mechanisms; Rosenfield & Mouzon, 2013). One might also expect to see differential expression of these pathways through the course of development, as gender intensification increases (Priess & Lindberg, 2011).

Some studies sidestep this issue by focusing only on one gender (e.g., Moilanen et al., 2010), while others have ignored it entirely (e.g., Wigelsworth et al., 2016). Of the limited number of studies that have explicitly investigated sex differences in developmental cascades, findings have been mixed. For example, while van Lier and Koot (2010) found minimal evidence that gender moderated the longitudinal relationships between peer relationships, externalizing problems, and internalizing symptoms from kindergarten to middle childhood, Deighton et al. (*in press*) found clear evidence of sex differences in their study exploring the links between externalizing problems, internalizing symptoms, and academic attainment in their middle childhood sample. Specifically, an adjustment erosion pathway from externalizing problems to later academic attainment was found for boys only, while academic incompetence pathways varied as a function of gender (e.g., poor initial academic attainment was predictive of later internalizing symptoms in girls, and externalizing symptoms in boys).

Fifth, while several individual risk factors have been accounted for in previous studies, our understanding of the

influence of *cumulative* risk exposure on developmental cascade pathways remains extremely limited (save for Obradovic et al., 2010). Cumulative risk theory has gained attention in developmental psychopathology due to the finding that exposure to multiple sources of risk is significantly more predictive of detrimental developmental consequences than any single risk factor in isolation (Evans, Li, & Whipple, 2013; Sameroff, 2006). This theory is based on the assumption that the accumulation of risk factors significantly raises the likelihood of negative outcomes through disruption of proximal developmental processes and increased allostatic load on the individual (Evans et al., 2013). Cumulative risk research typically makes use of an index calculated by summing exposure to each factor of interest. Unlike research on individual risk factors, it captures the *number* of risks experienced rather than their *intensity* (Evans et al., 2013). This can better conceptualize the effects of multiple adversities on development (Raviv, Taussig, Culhane, & Garrido, 2010). Although the cumulative risk approach means that information on risk factor intensity and interactive effects among factors is lost, this method has been found to be a more powerful predictor of developmental outcomes (Evans et al., 2013; Raviv et al., 2010), both cross-sectionally and longitudinally (e.g., Appleyard, Egeland, van Dulmen, & Sroufe, 2005).

In the current study, we present a developmental cascade model of the longitudinal relationships between internalizing symptoms, externalizing problems, and academic performance in middle childhood, utilizing a large sample from the United Kingdom. We draw upon cumulative risk theory in our conceptualization of shared risk, test for evidence of mediation effects, and examine the role of gender. The conceptual developmental cascades model shown in Figure 1 is our starting point. In this model, we examine cross-lag and mediation pathways to test the three hypotheses proposed by Moilanen et al. (2010), while accounting for within-time covariance, temporal stability, and data clustering, thereby increasing analytical rigor.

On the basis of the extant theory and research, the following predictions were made. First, consistent with the adjustment erosion model, we predicted that early externalizing problems (Hypothesis 1a) and internalizing symptoms (Hypothesis 1b) would be inversely related to later academic attainment. We also predicted that early externalizing problems would predict later internalizing symptoms (Hypothesis 1c); however, we anticipated that early internalizing symptoms would be *inversely* related to later externalizing problems (Hypothesis 1d). Second, in alignment with the academic incompetence model, we predicted that early academic problems would predict later internalizing symptoms (Hypothesis 2a) and externalizing problems (Hypothesis 2b). Third, given the generally inconsistent findings of studies vis-à-vis the shared risk model, we predicted that while accounting for cumulative risk exposure would significantly improve absolute model fit, it would not alter the magnitude or statistical significance of previously established cross-lag cascade pathways (Hypothesis 3). Fourth, we predicted that the developmental

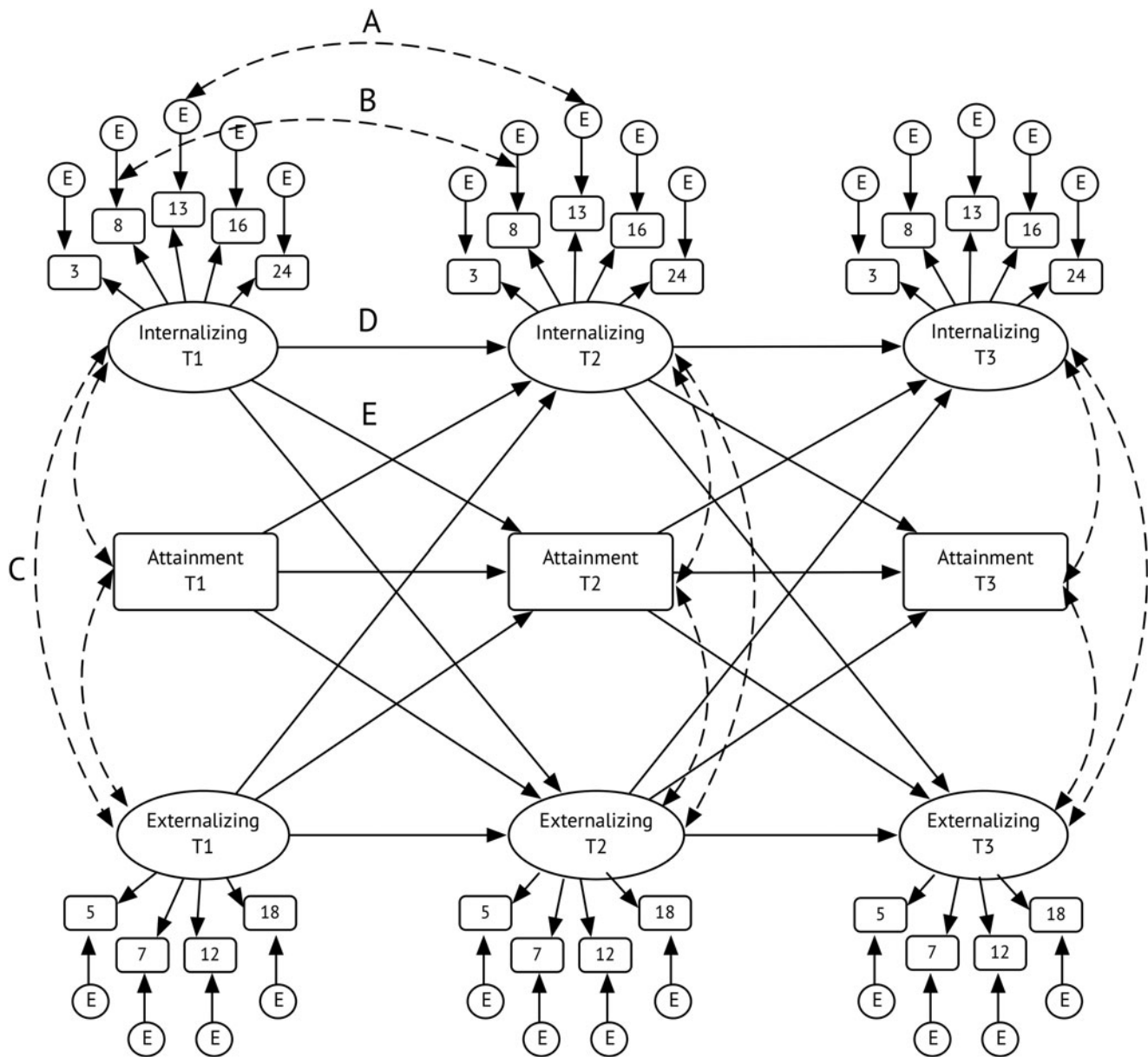


Figure 1. Conceptual developmental cascades model.

cascades pathways would vary across gender (Hypothesis 4). Fifth, we anticipated significant mediation effects to be evident alongside the basic cross-lag cascade pathways outlined above (Hypothesis 5).

Method

Design

A 3-year longitudinal cross-lag design was utilized, with annual (Times 1–3 [T1–T3], respectively) assessments of children's academic attainment, internalizing symptoms, and externalizing problems. Data were drawn from a major research project on school-based social and emotional learning funded

by the National Institute for Health Research (Project Ref. 10/3006/01; Humphrey et al., 2016).

Participants

The sample consisted of 1,771 children (908 male, 863 female) drawn from 16 mainstream primary schools in the northwest of England. Children's age ranged between 6.67 and 9.67 at T1 ($M = 8.14$, $SD = 0.86$), between 7.67 and 10.67 at T2 ($M = 9.14$, $SD = 0.86$), and between 8.67 and 11.67 at T3 ($M = 10.14$, $SD = 0.86$). In line with national averages (Department for Education, 2013), 17.5% ($N = 310$) were identified as having special educational needs (SEN). Deprivation levels in the current sample as measured by

eligibility for free school meals (FSM; 23%) were somewhat higher than national levels (Department for Communities and Local Government, 2015; Department for Education, 2012b). In terms of ethnicity, and similar to national picture of state-funded primary schools (Department for Education, 2012b) the majority of the participants were White ($N = 1,134$, 64%) followed by Asian ($N = 235$, 13.3%), Black ($N = 131$, 7.4%), mixed ($N = 127$, 7.2%), other/unclassified ($N = 55$, 3.1%), and Chinese ($N = 7$, 0.4%). The remaining 82 (4.6%) had incomplete data.

Measures

Academic attainment. The Interactive Computerised Assessment System (InCAS; Merrell & Tymms, 2007) standardized reading score was used to assess children’s academic attainment. InCAS reading captures a child’s ability to recognize words, to break them into sounds, and to choose the appropriate words to complete a passage (<http://www.cem.org>). Scores are age standardized to a mean of 100 and a standard deviation of 15. The Rasch person reliability of the different components of the InCAS reading test exceeds 0.92 in all cases, and scores are predictive of future, external assessment scores. For example, the correlation with performance indicators in primary schools reading test scores administered 12 months later is 0.72 (Centre for Evaluation and Monitoring, 2013).

Internalizing symptoms and externalizing problems. The five-item emotional symptoms and conduct problems subscales of the teacher informant-report version of the Strengths and Difficulties Questionnaire (SDQ; Goodman, 1997) were utilized. Respondents read a series of statements about a child (e.g., “often fights with other children” for conduct problems; “often unhappy, downhearted” for emotional symptoms) and endorse them using a 3-point scale (*not true, somewhat true, certainly true*). The two subscales have robust psychometric properties, with good internal consistency (conduct problems $\alpha = 0.70$; emotional symptoms $\alpha = 0.73$) and test–retest reliability (conduct problems $r = .77$; emotional symptoms $r = .72$), and satisfactory capacity to discriminate between children with and without psychiatric diagnoses (Stone, Otten, Engels, Vermulst, & Janssens, 2010).

Although the internal consistency of the two scales was found to be good in the current sample (see Table 1), preliminary analyses indicated that the removal of one item (“steals from home, school, or elsewhere”) would improve the Cronbach α coefficient of the conduct problems domain at all time points. Additional analyses assessing the measurement invariance of the two subscales across gender resulted in a warning about a correlation of 1 (i.e., not statistically distinguishable) between this item and others in the bivariate table in girls, thus indicating that it should be excluded from the main analyses. It is believed that this issue was caused by the low endorsement (0.1%–1%) of response 2 (*certainly true*) at all time points for girls.

Table 1. Study variable correlations and descriptive data

	1	2	3	4	5	6	7	8	9	10
1. Internalizing T1	—									
2. Internalizing T2	.352**	—								
3. Internalizing T3	.312**	.434**	—							
4. Attainment T1	-.149**	-.167**	-.152**	—						
5. Attainment T2	-.182**	-.166**	-.157**	.834**	—					
6. Attainment T3	-.147**	-.169**	-.176**	.780**	.864**	—				
7. Externalizing T1	.229**	.070**	.066*	-.202**	-.255**	-.261**	—			
8. Externalizing T2	.067**	.177**	.094**	-.191**	-.205**	-.224**	.573**	—		
9. Externalizing T3	.076**	.128**	.161**	-.200**	-.197**	-.211**	.521**	.635**	—	
10. Cumulative risk	.166**	.109**	.086**	-.313**	-.311**	-.270**	.320**	.259**	.237**	—
M	1.45	1.50	1.41	102.44	103.21	102.20	0.89	0.92	0.81	1.17
SD	2.04	2.04	1.94	14.94	14.39	13.21	1.61	1.51	1.42	1.12
Min–max	0–10	0–10	0–10	55–139	55–141.32	55–145	0–8	0–8	0–8	0–4
Cronbach α	0.78	0.79	0.78	—	—	—	0.80	0.76	0.77	—

Note: The cross-domain correlations within time are in bold. T1–T3, Times 1–3. * $p < .05$. ** $p < .01$.

Cumulative risk exposure. Exposure to five distinct risk factors known to be associated to both mental health difficulties and academic attainment was assessed. *Familial socioeconomic disadvantage* (Department for Education, 2012a; Green et al., 2005) was measured using FSM eligibility (0 = *no*, 1 = *yes*). Children in England are deemed eligible to receive FSM if they meet criteria set by the government relating to parental income. *Neighborhood deprivation* (Department for Education, 2014; Green et al., 2005) was ascertained using Income Deprivation Affecting Children Index scores, which are based on the proportion of children under 16 in a local area who live in low-income households. *SEN status* (Department for Education, 2014; Green et al., 2005) of children is reported by schools as part of their Pupil Level Annual School Census return and provides an index of whether a given child is considered to have SEN (0 = *no*, 1 = *yes*). Data pertaining to each of these first three risk factors were extracted from the National Pupil Database. This was supplemented by *lack of school connectedness* (Bond et al., 2007; Stewart, 2008) and *lack of peer and social support* (Bond et al., 2007; Stewart, 2008) derived from the corresponding subscales of the child self-report version of the KIDSCREEN-27 health-related quality of survey (Ravens-Sieberer et al., 2007). In each four-item subscale, children read questions (e.g., “In the last week, have you been able to rely on your friends?” [peer and social support]; “In the last week, have you enjoyed going to school?” [school environment]) and respond on a 5-point scale (*never, one day, some days, most days, every day*). The two subscales are psychometrically robust, with good internal consistency (peer and social support $\alpha = 0.81$, school connectedness $\alpha = 0.81$), test–retest reliability (peer and social support intraclass correlation coefficient = 0.61, school connectedness intraclass correlation coefficient = 0.74), and the ability to discriminate between children with good health and those with poorer physical or mental health status (as determined through external assessment; Ravens-Sieberer et al., 2007; Robitail et al., 2007).

In line with standard practice in cumulative risk research, an index was calculated by summing exposure to each factor of interest (coded as 1 = *exposure*, 0 = *no exposure*).¹ Thus, every child was initially allocated a score of 0–5, with a higher score representing greater exposure to cumulative risk. Subsequently, the groups representing four and five risk factors were merged to represent 4+ factors. This was guided by preliminary risk–outcome relationship analyses that indicated stability on the outcome variables from four to five factors and due to the fact that only a small minority of participants ($N = 10$, 0.6%) were classified as being exposed to all five risk factors. Thus, the final cumulative risk

index ranged between 0 and 4 ($M = 1.17$, $SD = 1.11$), with the largest proportion of participants not exposed to any risk factors ($N = 537$, 30.3%). A further 471 (26.6%) were exposed to one, 346 (19.5%) to two, and 142 (8%) to three, with only 58 (3.3%) reporting experience of four or more risk factors. The remaining 217 (12.3%) participants had incomplete data.

Statistical analysis

All models were tested by structural equation modeling (SEM) with weighted least squares with means and variance adjustment in MPlus 7.4. Weighted least squares with means and variance adjustment was used to account for the categorical nature of the indicators of the internalizing and externalizing latent factors, which can handle computationally demanding models with large sample sizes and many latent factors (Brown, 2015). The results from an exploratory factor analysis (not presented here) confirmed the use of the externalizing and internalizing constructs as latent factors. Latent variables estimate and remove item-level measurement error, providing more accurate estimates of autoregressive and cross-lag estimates (Newsom, 2015).

Absolute model fit was evaluated using comparative fit index (CFI), Tucker–Lewis index (TLI), and root mean square error of approximation (RMSEA). CFI and TLI values above 0.95 and RMSEA values below 0.05 were considered to indicate good model fit (Byrne, 1998; Hu & Bentler, 1999). The chi-square statistic is reported in the results but was not used to assess model fit, given its sensitivity to sample size. The mediation analyses were conducted using 10,000 resamples. For all the indirect effects that were statistically significant, the standardized beta coefficients along with the 95% bias-corrected confidence intervals are reported. Finally, following Keith (2015) we set a threshold of $\beta > 0.1$ (in addition to $p < .05$) for individual path coefficients as being indicative of practical significance.

Measurement invariance. Establishing longitudinal measurement invariance provides empirical evidence that the meaning of the constructs has not changed across the different waves of data collection (Little, Preacher, Selig, & Card, 2007). Given that full measurement invariance frequently does not hold, partial measurement invariance was considered acceptable for proceeding to the examination of the cross-lagged models (Byrne, Shavelson, & Muthén, 1989; Steenkamp & Baumgartner, 1998). To test for measurement invariance, nested models were estimated as described by Muthén and Muthén (1998–2012), and their model fit was assessed using the chi-square difference. Based on results not presented here (but available from the authors on request), full longitudinal and gender measurement invariance was established for both latent constructs.

Gender structural invariance. Having established measurement invariance across gender, structural invariance was

1. In the case of the two continuous variables derived from the KIDSCREEN-27, the bottom 25% of scores were recoded as 1 (*exposure*), and the remaining 75% as 0 (*no exposure*). Regarding the Income Deprivation Affecting Children Index scores, the top 25% (high deprivation) were recoded as 1 (*exposure*), and the remaining 75% as 0 (*no exposure*).

tested to determine whether the cross-lagged model parameters were equal across groups. To assess for gender differences, a multigroup SEM was applied. An unconstrained model (H1) in which all paths were allowed to be unequal (simultaneously freed) between boys and girls was compared to a more constrained nested model (H0) where gender equality constraints were imposed. In this case, a nonsignificant chi-square difference test would indicate invariance (i.e., the constraints did not worsen model fit).

Evaluation of developmental cascades models. Although all cross-lag paths were estimated as shown in Figure 1, only those that were statistically significant are presented in the following figures. In addition, for simplicity, cross-domain correlations were omitted from the figures but are reported in Table 2. Factor loadings for all latent constructs are reported in the Results section. The developmental cascades models were tested through SEM, accounting for several important modeling issues related to latent variable longitudinal cross-lagged models (see Cole & Maxwell, 2003; Little et al., 2007; Newsom, 2015). First, both internalizing and externalizing constructs were modeled as latent variables, which, as previously mentioned, enabled the estimation of autocorrelations among residual covariances (Path A in Figure 1), thus removing stable specific variance that would otherwise inflate the estimates of stability over time (Cole & Maxwell, 2003; Newsom, 2015). Second, given that longitudinal measurement invariance was established, the factor loadings of each latent factor were constrained to be equal across time (Path B), thereby increasing the precision of the model estimates (Newsom, 2015). Third, the cross-lagged SEM analyses allowed for the examination of the stability of all paths, while controlling for the cross-domain associations within time (Path C) and cross-time associations within domains (temporal stability; Path D). As illustrated in the Path E example, the current cascades model took into consideration all possible cross-lag paths from each domain to others across all waves.

Similar to gender invariance, the effects of cumulative risk were explored through the comparison of nested models. The cumulative risk index was included in the proposed model,

although all pathways from this variable were fixed to zero (model H0). Using the chi-square difference test, this model was compared against a less restrictive model (H1) where all pathways related to the cumulative risk index were allowed to be freely estimated. The effects of cumulative risk were assessed separately for boys and girls.

Results

Preliminary analysis

Missing data in the current sample (3.8%–18.4%) were attributed to school absences and error-based omission for individual survey items. Attainment data was higher (18.4%) due to testing difficulties in the administration of computer-based tasks in primary school settings. On this basis, data were assumed to be missing at random, thus enabling the use of all available information (i.e., full information). Means, standard deviations, internal consistency coefficients, and bivariate correlations among the study variables are shown in Table 1. As expected, externalizing problems and internalizing symptoms were shown to negatively correlate with academic attainment at all waves, and externalizing problems correlated positively with internalizing symptoms. Cumulative risk was found to positively correlate with internalizing symptoms and externalizing problems, while negatively correlating with academic attainment. Although it was found to have statistically significant associations with all study variables, cumulative risk had stronger correlations with constructs at T1, gradually decreasing at T2 and T3.

Gender structural invariance

The aim of the gender structural invariance analysis was to address existing gaps in the literature and assess the degree to which the proposed developmental cascades model varied across gender. An acceptable model fit for both groups should be established before proceeding to multigroup analysis; therefore, prior to the examination of gender differences,

Table 2. Within time correlations in boys and girls

Gender	Model 1A: Cumulative Risk Freely Estimated			Model 1B: Cumulative Risk Held at Zero		
	Attainment– Internalizing	Attainment– Externalizing	Internalizing– Externalizing	Attainment– Internalizing	Attainment– Externalizing	Internalizing– Externalizing
Boys (<i>N</i> = 908)						
T1	–.110*	–.114*	.299***	–.210***	–.255***	.382***
T2	–.043	.134	.592***	–.012	.140	.588***
T3	–.035	.128	.245***	.003	.141	.229***
Girls (<i>N</i> = 863)						
T1	–.132***	–.210*	.432***	–.164***	–.265***	.461***
T2	.018	.207	.473***	.021	.208	.470**
T3	–.068	–.026	.450***	–.056	–.028	.445***

Note: T1–T3, Times 1–3.
p* < .05. *p* < .01. ****p* < .001.

Confirmatory factor analysis was performed to assess the fit of the overall model (Figure 1), results of which indicated a good fit for both boys, χ^2 (389) = 529.234, $p < .001$; RMSEA = 0.020, 90% CI [0.015, 0.024], $p > .05$; CFI = 0.971, TLI = 0.966, and girls, χ^2 (389) = 519.423, $p < .001$; RMSEA = 0.020, 90% CI [0.015, 0.024], $p > .05$; CFI = 0.961, TLI = 0.954. Multigroup analysis indicated that the unconstrained model, in which all parameters were allowed to vary across gender, was found to fit the data well: χ^2 (806) = 1,091.841, $p < .001$; RMSEA = 0.020, 90% CI [0.017, 0.023], $p > .05$; CFI = 0.967, TLI = 0.962, as did the constrained model, in which all parameters were held equal across gender: χ^2 (848) = 1,125.626, $p < .001$; RMSEA = 0.019, 90% CI [0.016, 0.022], $p > .05$; CFI = 0.968, TLI = 0.965. In line with Hypothesis 4, a statistically significant difference in model fit, χ^2 diff. (42) = 70.567, $p < .01$, suggested that the cross-lag model parameters varied across gender. Taking these findings into consideration, subsequent analyses were, therefore, conducted separately for boys and girls.

Developmental cascades

The factor loadings of the externalizing and internalizing latent factors were examined. Latent factors with loadings above .40 were considered to represent a meaningful construct. The results were satisfactory for both latent constructs across all waves. Factor loadings ranged between .80 and .92 in boys, and from .73 to .87 in girls for the externalizing latent factor. Similarly, high factor loadings were observed in the internalizing latent factor (boys: $\lambda = .63-.87$; girls: $\lambda = .64-.87$).

The cross-sectional correlations between the externalizing and internalizing latent variables, and academic attainment are shown in Table 2. Externalizing problems and internalizing symptoms were found to be positively associated across all time points. Academic attainment was found to have negative cross-sectional correlations with externalizing and internalizing only at T1 in both groups.

While controlling for temporal stability, cross-domain associations within time, data clustering, and the effects of cumulative risk, a number of statistically significant cascade paths were identified, results of which are shown in Figure 2a and Figure 3a for boys and girls, respectively. Consistent with previous findings, academic attainment and externalizing problems were highly stable across time (see Table 1). An evident drop in stability was, however, observed in girls from externalizing problems at T2 to T3. Internalizing symptoms were found to be stable in both groups.

Hypothesis 1a was supported only for boys, as higher levels of externalizing problems at T1 and T2 significantly predicted poorer academic performance at T2 and T3, respectively. It is, however, important to note the small magnitude ($\beta = -0.056$; Keith, 2015) of the latter pathway (T2 \rightarrow T3). Contrary to our expectations, initial internalizing symptoms were *not* found to predict later academic attainment in either girls or boys (Hypothesis 1b). In addition, despite the

evident positive longitudinal relationship between externalizing problems and internalizing symptoms in the extant literature, Hypothesis 1c was not supported. Thus, higher initial levels of externalizing problems were found to lead to *lower* internalizing symptoms in both boys (T1 \rightarrow T2) and girls (T2 \rightarrow T3). However, as hypothesized (Hypothesis 1d), a significant path was observed between internalizing symptoms at T1 and externalizing problems at T2, such that higher initial internalizing symptoms lead to reduced externalizing problems later in time. This path was of greater magnitude in girls ($\beta = -0.419$) than in boys ($\beta = -0.121$). Finally, gender differences were observed regarding the academic incompetence hypothesis. It found support only in girls, such that greater academic incompetence at T1 was associated with increased internalizing symptoms at T2, while greater academic incompetence at T2 was associated with increased internalizing symptoms (Hypothesis 2a) and externalizing problems (Hypothesis 2b) at T3.

Cumulative risk analysis

The inclusion of cumulative risk in the model was shown to have a significant impact on the variables at T1 only (with the exception of T3 academic attainment in girls) in both groups. It was found to explain more variance in externalizing problems, followed by academic attainment and internalizing symptoms. In other words, there appears to be a more robust connection between multiple adversities and externalizing problems. This path (as well as that between cumulative risk and internalizing symptoms) was positive, such that increased risk exposure was associated with increased externalizing problems and internalizing symptoms. In contrast, higher risk exposure was associated with poorer academic performance.

Nested model comparisons examining the effect of cumulative risk are presented in Figures 2 and 3 for boys and girls, respectively. The full model (Figures 2a, 3a), where cumulative risk is freely estimated, was compared against a constrained model (Figures 2b, 3b), in which the paths from cumulative risk are held to zero. The full model fit the data very well for boys, χ^2 (389) = 529.234, $p < .001$; RMSEA = 0.020, 90% CI [0.015, 0.024], $p > .05$; CFI = 0.971, TLI = 0.966, as well as for girls, χ^2 (389) = 519.423, $p < .001$; RMSEA = 0.020, 90% CI [0.015, 0.024], $p > .05$; CFI = 0.961, TLI = 0.954. The constrained model achieved an acceptable but worse model fit than the full model in both boys, χ^2 (398) = 673.191, $p < .001$; RMSEA = 0.028, 90% CI [0.024, 0.031], $p > .05$; CFI = 0.945, TLI = 0.936, and girls, χ^2 (398) = 585.707, $p < .001$; RMSEA = 0.023, 90% CI [0.019, 0.027], $p > .05$; CFI = 0.945, TLI = 0.935. In alignment with Hypothesis 3, the statistically significant difference in model fit, boys χ^2 diff. (9) = 184.402, $p < .001$; girls χ^2 diff. (9) = 107.845, $p < .001$, indicated that the full model had a better fit and, accordingly, was chosen to be the focus of our discussion. However, the inclusion of cumulative risk did not have a substantive impact on the statistical

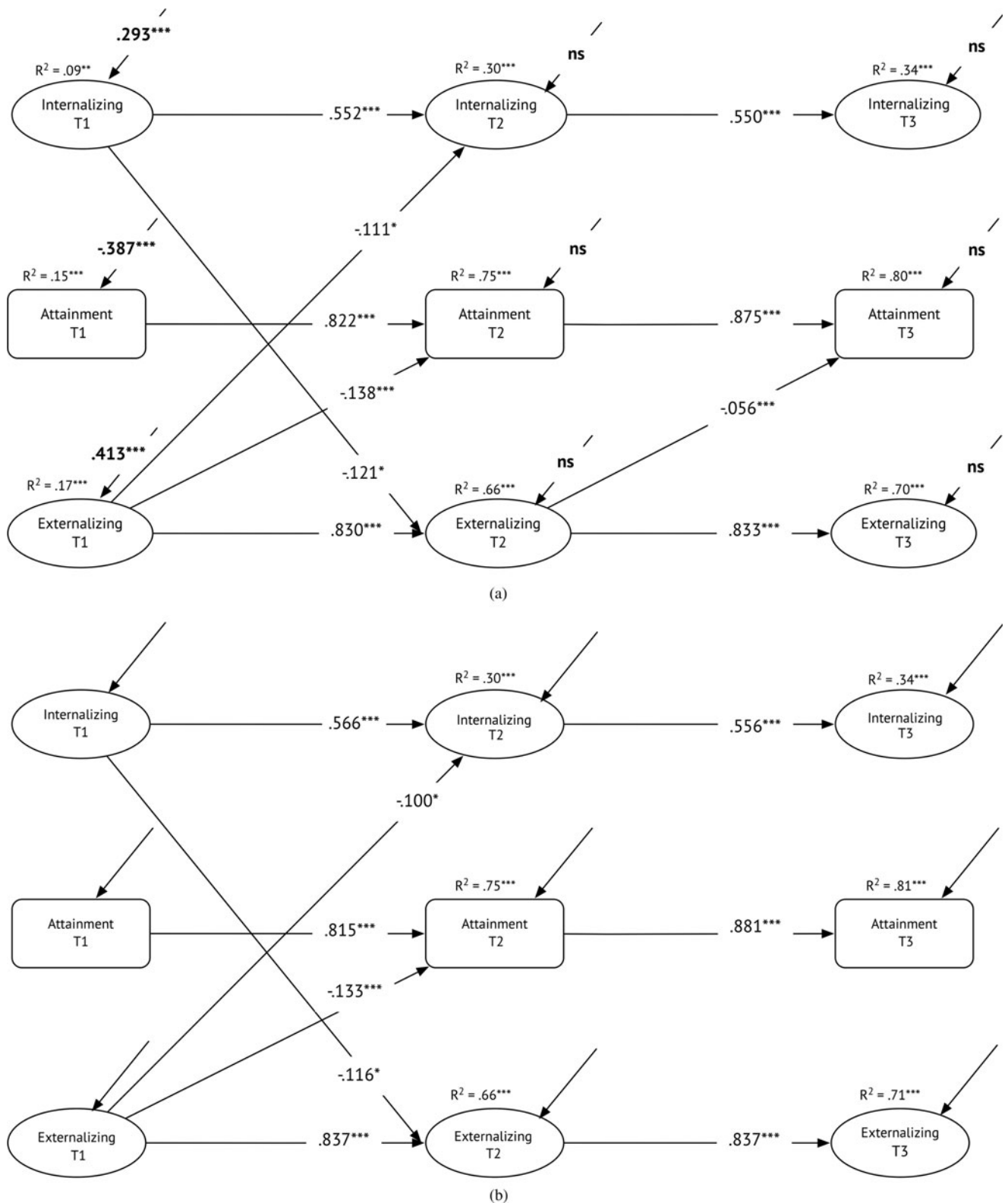


Figure 2. Cascade model for boys: (a) Cascade model with cumulative risk freed. The effects of cumulative risk are in bold, and beta values are standardized. $N = 908$; $\chi^2(389) = 529.234$, $p < .001$; root mean square error of approximation = 0.20, 90% confidence interval [0.015, 0.024], $p > .05$; comparative fit index = 0.971, Tucker–Lewis index = 0.966. * $p < .05$, ** $p < .01$, *** $p < .001$. χ^2 diff. (9) = 184.402, $p < .001$. (b) Cascade model with cumulative risk held at 0. Beta values are standardized. $N = 979$; $\chi^2(398) = 673.191$, $p < .001$; root mean square error of approximation = 0.28, 90% confidence interval [0.024, 0.031], $p > .05$; comparative fit index = 0.945, Tucker–Lewis index = 0.936. * $p < .05$, ** $p < .01$, *** $p < .001$.

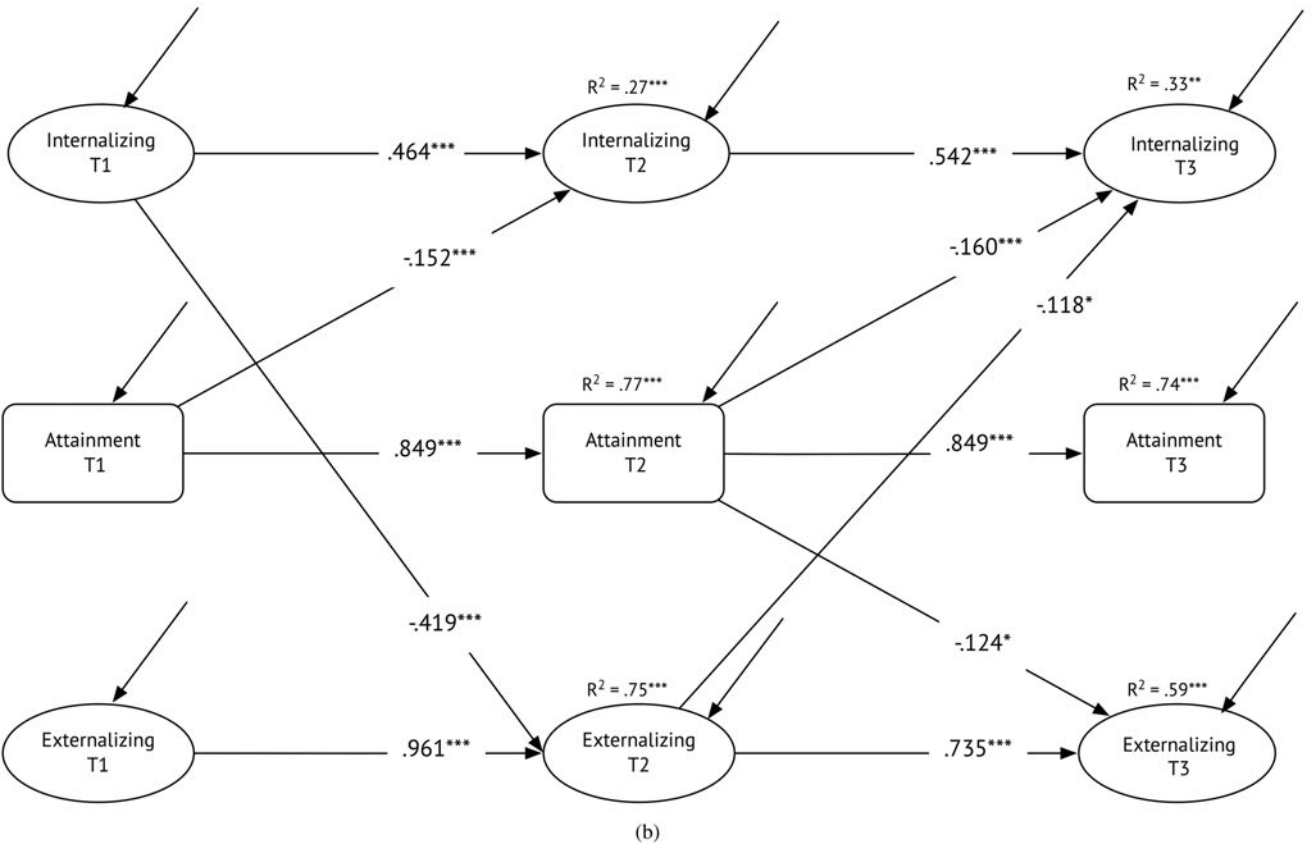
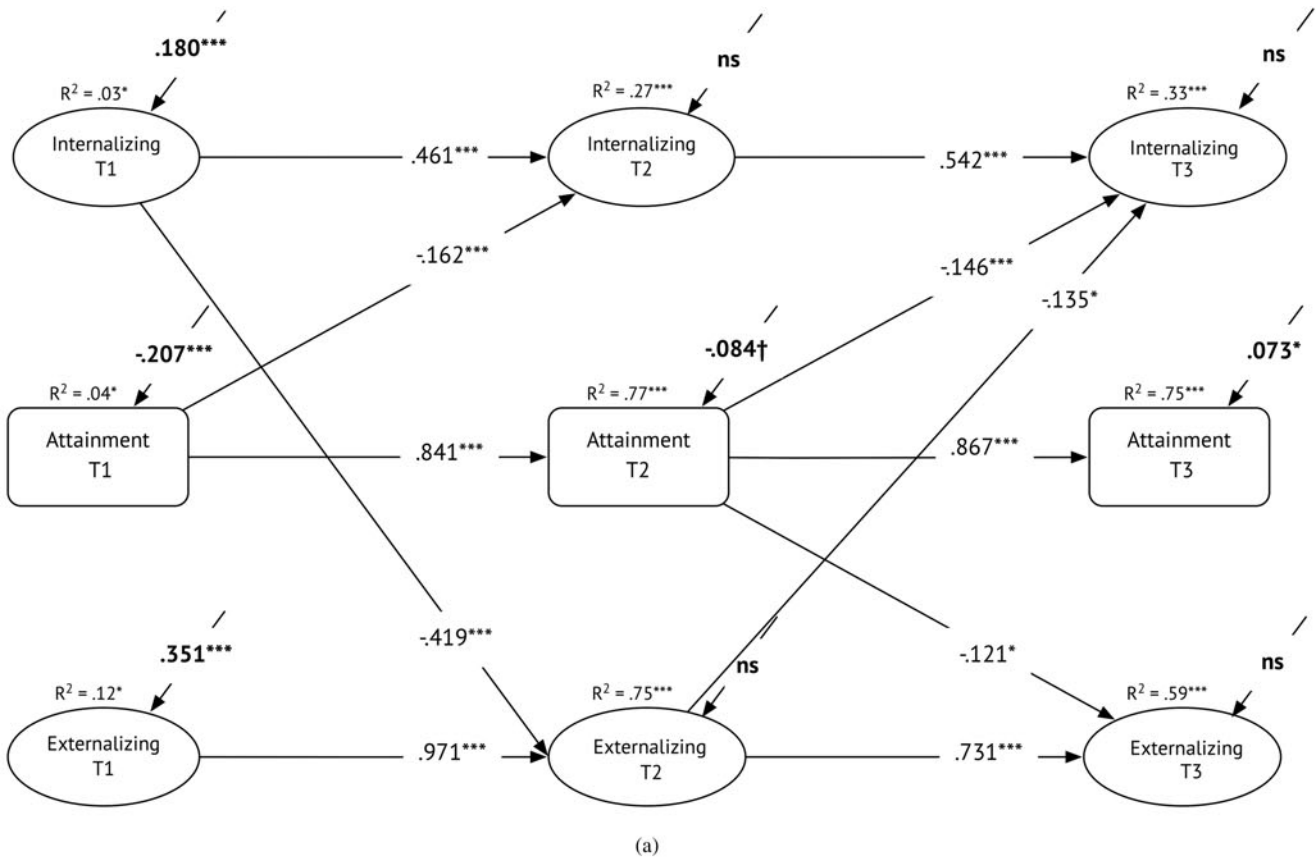


Figure 3. Cascade model for girls: (a) Cascade model with cumulative risk freed. The effects of cumulative risk are in bold, and beta values are standardized. $N = 863$; $\chi^2(389) = 519.423$, $p < .001$; root mean square error of approximation = 0.20, 90% confidence interval [0.015, 0.024], $p > .05$; comparative fit index = 0.961, Tucker–Lewis index = 0.954. † $p = .05$, * $p < .05$, ** $p < .01$, *** $p < .001$. χ^2 diff. (9) = 107.845, $p < .001$. (b) Cascade model with cumulative risk held at 0. Beta values are standardized. $N = 863$; $\chi^2(398) = 585.707$, $p < .001$; root mean square error of approximation = 0.23, 90% confidence interval [0.019, 0.027], $p > .05$; comparative fit index = 0.945, Tucker–Lewis index = 0.935. * $p < .05$, ** $p < .01$, *** $p < .001$.

significance, direction, or magnitude of the cascade paths. This was the case for both boys and girls, with a single exception: the path between externalizing problems at T2 and academic attainment at T3 for boys reached statistical significance only after controlling for the effects of cumulative risk.

Indirect pathways

Possible indirect pathways were identified and statistically assessed in the full models (Hypothesis 5, Figures 2a, 3a). Of the 27 mediation paths tested, only a few were found to be statistically significant (see Table 3). As would be expected, statistically significant mediation paths ($p < .001$) were observed within variables across time in both boys and girls. For instance, internalizing symptoms at T2 were found to account for the relationship between internalizing symptoms at T1 and T3 (boys $\beta = 0.31$; girls $\beta = 0.26$). This was also the case for externalizing problems (boys $\beta = 0.70$; girls $\beta = 0.69$) and academic attainment (boys $\beta = 0.73$; girls $\beta = 0.72$). Additional findings from mediation analysis provided further confirmation of gender differences in developmental cascades (Hypothesis 4). Two of the mediation paths unique to boys had externalizing problems as their starting point. More precisely, externalizing problems were found to predict more externalizing problems as well as poorer academic performance a year later, which in turn predicted further academic incompetence (thus supporting Hypothesis 1a). In contrast, three of the paths unique to girls had academic incompetence as their starting point. Poor initial academic performance in girls was found to predict further academic incompetence, and increased internalizing symptoms a year later, which then lead to further internalizing symptoms (Hypothesis 2a). In addition, girls whose academic attainment was poor at T1 were found to report more externalizing problems at T2, which was in turn shown to predict further such problems at T3 (Hypothesis 2b). Finally, both boys and girls with initial internalizing symptoms were found to report less externalizing problems a year later, which led to an increase in externalizing problems at T3 (Hypothesis 1d).

Discussion

In the current study we sought to explore the developmental and reciprocal relationships between academic attainment, externalizing problems, and internalizing symptoms in middle childhood. Much of the available work on developmental cascades is limited to studies with small sample sizes and long intervals between assessments (e.g., Masten et al., 2005; Obradovic et al., 2010), which preclude firm conclusions on the exact timing and processes by which such cascades occur (Masten et al., 2005). With a large sample and a rigorous methodology to draw on, the current study explored the longitudinal associations between these salient developmental domains, while also taking into consideration the role of gender differences and indirect effects, which to date remain largely understudied.

Expected positive cross-sectional associations between externalizing problems and internalizing symptoms were found at all time points. A negative association was found between academic attainment and both externalizing problems and internalizing symptoms, but only for T1 variables, even after accounting for cumulative risk. Adding to existing research (Englund & Siebenbruner, 2012; Vaillancourt et al., 2013), all variables exhibited high stability across time, particularly academic attainment and externalizing problems.

While accounting for within-time covariance, data nesting, and temporal stability, overall findings from the current study concur with the developmental cascades theory (Masten & Cicchetti, 2010), by which important domains of development spread into other domains across time both directly and indirectly. Consistent with the findings by Deighton et al. (in press), the proposed developmental cascades model was found to vary by gender. An adjustment erosion path from early externalizing problems to later academic performance was found exclusively for boys. This was not surprising as boys tend to be more prone to externalizing problems relative to girls, who are expected to experience more internalizing symptoms (Green et al., 2005; McCarty, 2008). Such differences are considered by some to be the influence

Table 3. Indirect pathways in boys and girls

Gender	Predictor T1	Indirect Paths		Coefficients		95% Bias-Corrected CI	
		Mediator T2	Outcome T3	β	SE	LL	UL
Boys ($N = 908$)							
	Externalizing →	Attainment →	Attainment	-0.106**	0.035	-0.177	-0.041
	Internalizing →	Externalizing →	Externalizing	-0.102*	0.049	-0.198	-0.008
	Externalizing →	Externalizing →	Attainment	-0.057*	0.024	-0.109	-0.014
Girls ($N = 863$)							
	Attainment →	Internalizing →	Internalizing	-0.093***	0.022	-0.143	-0.056
	Attainment →	Attainment →	Internalizing	-0.132**	0.037	-0.204	-0.056
	Internalizing →	Externalizing →	Externalizing	-0.273**	0.075	-0.445	-0.146
	Attainment →	Externalizing →	Externalizing	-0.079*	0.033	-0.139	-0.012

Note: T1–T3, Times 1–3; LL, lower level; UL, upper level.

* $p < .05$. ** $p < .01$. *** $p < .001$.

of socialization pressures (Chaplin et al., 2005). For instance, shyness and good prosocial behaviors, instead of deviant ones, are often encouraged in girls by their parents (Keenan & Shaw, 1997). These findings are of utmost importance as they provide evidence for gender-specific longitudinal profiles and strongly suggest the need for further systematic exploration.

In line with the findings of previous studies (Deighton et al., *in press*), the adjustment erosion pathway from internalizing symptoms to academic attainment was not supported. While this was unexpected, it might be due to the use of teacher-report measures for the assessment of externalizing and internalizing symptoms. It is possible that children's internalizing symptoms have been underestimated by their teachers because they are harder to observe, compared to externalizing problems (De Los Reyes et al., 2015). However, even in studies where youth self-report methods have been used to assess internalizing symptoms (such as Masten et al., 2005), robust longitudinal connections to later academic attainment are not found consistently. Thus, like Masten et al. (2005), we are left to speculate that negative cascades in the academic domain may only be evident for subgroups that fall in the clinical range for internalizing problems.

To our surprise and in contrast with the theory underlying adjustment erosion (Moilanen et al., 2010), initial externalizing problems were found to predict *less* internalizing symptoms at later times in both boys and girls. One potential explanation for this finding is that externalizing behaviors can be associated with higher levels of self-esteem (Baumeister, Smart, & Boden, 1996). Thus, students' engagement in externalizing behaviors with delinquent peers is hypothesized to increase their sense self-worth and belonging (Kaplan, 1975), thus alleviating their subsequent internalizing symptoms. One might also consider the possibility that the attention received from teachers in response to the display of externalizing behaviors may act as a protective factor.

From a methodological perspective, the aforementioned reliance on teacher-report ratings, along with the fact that these ratings came from different teachers at each time point, might have influenced the direction of the externalizing–internalizing pathways. One should also consider that the yearly time intervals of the current study might have been too small to capture the expected change in behavior (externalizing problems leading to more internalizing symptoms). High and directionally unclear within-time associations due to small time lags can obscure the direction of effects (Masten & Cicchetti, 2010). Some authors (e.g., Vaillancourt et al., 2013) suggest allowing more time to elapse between waves as it may provide a more accurate representation of cascade effects. To confirm this theory, we conducted a post hoc analysis to assess the longitudinal association between externalizing problems at T1 and internalizing symptoms at T3. This analysis revealed a positive cascade effect in both boys ($\beta = 0.418, p < .01$) and girls ($\beta = 0.228, p < .05$), suggesting that initial externalizing problems predict *more* internalizing

symptoms 2 years after. It is of note, however, that several other studies examining this particular adjustment-erosion pathway have found null results (Bornstein et al., 2010; Burt & Roisman, 2010; Englund & Siebenbruner, 2012; Masten et al., 2005; Obradovic et al., 2010; Rogosch et al., 2010). Clearly, there is a need for further investigation.

In line with our hypotheses and existing findings (Bornstein et al., 2010, 2013; Englund & Siebenbruner, 2012; Masten et al., 2005; Mesman et al., 2001; Rogosch et al., 2010) the adjustment erosion cascade from internalizing symptoms to externalizing behaviors followed the expected direction. Internalizing symptoms thus appeared to act as a protective factor guarding against later externalizing problems, possibly due to the increased self-isolation and withdrawal associated with the former (Masten et al., 2005). It is of interest that the magnitude of this path was stronger in girls, who tend to report more internalizing symptoms (McCarty, 2008), including withdrawal, and thus are less likely to engage in deviant behaviors at later points in time (Moilanen et al., 2010).

Consistent with the extant literature relating to the academic incompetence model (Moilanen et al., 2010; Weeks et al., 2016), poor academic performance was found to exacerbate externalizing and internalizing behaviors, though only in girls. Such group differences are consistent with theories suggesting that girls tend to evaluate themselves more negatively than boys (Pomerantz, Altermatt, & Saxon, 2002) and are more concerned than boys with pleasing their parents and teachers (Hoffman, 1972). This may in turn lead to internalizing symptoms such as distress over possible failure, which can be intensified when failure is actually experienced (Pomerantz et al., 2002). It is here where a connection between our findings and the education policy context in the United Kingdom may be made. Through successive governments, the last several years have witnessed the proliferation of “high stakes” academic testing in schools, driven by an accountability agenda (Hutchings, 2015). The likely impact on children's mental health has been noted (House of Commons Education and Health Committees, 2017). Given the distinct, gender-specific pathways identified between academic performance and later internalizing symptoms in the current study, we might reasonably anticipate differential changes in the prevalence of such difficulties over time. Scrutiny of several recent longitudinal analyses supports this assertion, with significant increases in internalizing symptoms noted among early adolescent females in the last decade, in contrast to relative stability elsewhere (Bor, Dean, Najman, & Hayatbakhsh, 2014; Fink et al., 2015; Lessof, Ross, Brind, Bell, & Newton, 2016). Our findings suggest that these trends may, at least in part, be driven by increased academic pressure in schools.

Despite the clear improvement in model fit after controlling for cumulative risk, we found limited support for the shared risk hypothesis, a finding that is consistent with existing research (Deighton et al., *in press*; Masten et al., 2005; Moilanen et al., 2010; Wigelsworth et al., 2016). Only the pathway from externalizing problems at T2 to academic

attainment at T3 in boys became statistically significant after accounting for cumulative risk. Our findings collectively suggest that the influence of risk factors on cascade effects, even when their cumulative influence is assessed, may be less powerful than initially hypothesized (Deighton et al., *in press*). However, the limited variability in cumulative risk groups may have also influenced the results.

The current study is one of few (Moilanen et al., 2010; Vaillancourt et al., 2013) to statistically explore indirect cascade pathways while accounting for cumulative risk. Irrespective of gender, internalizing symptoms at T1 were indirectly related to internalizing symptoms at T3 via prior externalizing symptoms. The same pattern was observed for academic attainment and externalizing problems, which confirms the progressive change in the developmental domains over time. Initial internalizing problems in both boys and girls were inversely related to externalizing problems at T2, which in turn predicted increased externalizing problems at T3. It is, however, worth mentioning that the magnitude of the path from T2 externalizing problems to T3 externalizing problems was found to be equal to or smaller than that of previous time points (T1 to T2), indicating the positive long-term impact of internalizing symptoms.

The findings from mediation analysis lend further support for the existence of gender differences in developmental cascades. In boys, the indirect effects observed appear to have externalizing symptoms at their starting point. More precisely, externalizing problems at T1 were found to exacerbate externalizing problems and academic problems at T2, which in turn predicted further academic incompetence at T3. All indirect pathways unique to girls had academic attainment as their starting point. Academic attainment at T1 was indirectly associated to internalizing symptoms at T3 through academic attainment and internalizing symptoms at T2. Moreover, T1 academic attainment was indirectly related to externalizing problems at T3 via prior externalizing problems.

Taken in sum, our findings are strongly suggestive of gender-specific developmental cascade profiles operating both directly and indirectly over time. We theorize that these are the result of gendered socialization practices relating to emotional expression (Chaplin et al., 2005) and academic motivation and achievement (Leaper & Friedman, 2007). As our analysis was limited to a single developmental phase (middle childhood), we were unable to explore whether, in line with gender intensification theory (Priess & Lindberg, 2011), there is differential expression of these pathways through the course of development. However, when contrasting our findings with, for example, those of van Lier and Koot's (2010) study focusing on pathways from early to middle childhood (in which minimal gender differences were found), there is support for the notion of increasingly gendered socialization practices in key developmental contexts such as the home and school beginning to show their effects over time. Clearly, a significant line of inquiry for future research is to explore this possibility.

A number of limitations should be taken into consideration when interpreting the findings of the current study. First, al-

though the sample was sociodemographically diverse, it was not fully representative of the population from which it was drawn (e.g., with higher levels of socioeconomic disadvantage than are seen nationally), thus limiting the generalizability of findings. Second, as previously noted, the current study relied on teacher reports for the assessment of mental health domains. Despite sufficient evidence for the robust psychometric properties of the teacher version of the SDQ (Stone et al., 2010), the degree to which it accurately captured children's internalizing symptoms is unclear. This could also explain the lower stability observed in this domain. Teachers are likely not only to report more externalizing than internalizing problems but also to do so more for boys than for girls (e.g., Winsler & Wallace, 2002), which might explain the low endorsement of the SDQ item "steals from home, school, or elsewhere" in girls. Researchers are, therefore, encouraged to consider using clinical interviews and self-report measures, which are regarded to be more adequate for the assessment of internalizing symptoms in children and adolescents (Whitcomb & Merrell, 2013). Third, although accounted for, the cluster sample size ($N = 16$) was slightly lower than the recommended cutoff of 20 for multilevel analysis (Snijders & Bosker, 2012). Fourth, although the sample of the current study was sufficient for the number of parameters examined ($>5:q$ ratio; Kline, 2016), future studies should consider conducting a Monte Carlo analysis prior to the investigation of complex longitudinal models. Fifth, our cumulative risk index was not comprehensive, as it did not capture more extreme risks such as child maltreatment and parental mental illness. However, we note the ubiquity of this issue in secondary analyses of existing data sets: one can only analyze what was measured in the first place.

Sixth and finally, we should also note recent challenges to the use of cross-lagged panel models that are used in much developmental cascades research (including this paper) and contemporary developmental science more generally (see, e.g., Berry & Willoughby, 2016; Hamaker, Kuiper, & Grasman, 2015). The principal concern raised is that such models fail to adequately disaggregate within- and between-person effects, thus failing to align with the developmental theories they are intended to test, and yielding parameter estimates are difficult to interpret in a meaningful way. These critiques suggest that future research in this area will need to adapt models to include growth factors and random intercepts.

Conclusion

The current study contributes to the growing field of developmental cascades research, providing important insights into the longitudinal associations between academic attainment, externalizing problems, and internalizing symptoms in a large sample of English children. The methodological advancements reported herein, including the use of longitudinal- and gender-invariant latent variables; the consideration of data nesting, cumulative risk, and gender differences; and the statistical exploration of indirect pathways, provide a more

robust and comprehensive analysis of these important developmental processes than has been evident in much previous research. Our findings highlight that developmental cascades vary across gender in middle childhood, with consequent implications for the explanatory models outlined in our introduction. For example, externalizing-attainment adjustment

erosion pathways were found only in boys, whereas attainment-internalizing (and externalizing) academic incompetence pathways were found only in girls. Further research is needed, however, to explore whether and how these gender-specific pathways intensify or attenuate through the course of development.

References

- Ansary, N. S., & Luthar, S. S. (2009). Distress and academic achievement among adolescents of affluence: A study of externalizing and internalizing problem behaviors and school performance. *Development and Psychopathology, 21*, 319–341. doi:10.1017/S0954579409000182
- Appleyard, K., Egeland, B., van Dulmen, M. H., & Sroufe, L. A. (2005). When more is not better: The role of cumulative risk in child behavior outcomes. *Journal of Child Psychology and Psychiatry, 46*, 235–245. doi:10.1111/j.1469-7610.2004.00351.x
- Baumeister, R. F., Smart, L., & Boden, J. M. (1996). Relation of threatened egotism to violence and aggression: The dark side of high self-esteem. *Psychological Review, 103*, 5–33. doi:10.1037/0033-295X.103.1.5
- Berry, D., & Willoughby, M. T. (2016). On the practical interpretability of cross-lagged panel models: Rethinking a developmental workhorse. *Child Development*. Advance online publication. doi:10.1111/cdev.12660
- Bond, L., Butler, H., Thomas, L., Carlin, J., Glover, S., Bowes, G., & Patton, G. (2007). Social and school connectedness in early secondary school as predictors of late teenage substance use, mental health, and academic outcomes. *Journal of Adolescent Health, 40*, 357. doi:10.1016/j.jadohealth.2006.10.013
- Bor, W., Dean, A. J., Najman, J., & Hayatbakhsh, R. (2014). Are child and adolescent mental health problems increasing in the 21st century? A systematic review. *Australian and New Zealand Journal of Psychiatry, 48*, 606–616. doi:10.1177/0004867414533834
- Bornstein, M. H., Hahn, C. S., & Haynes, O. M. (2010). Social competence, externalizing, and internalizing behavioral adjustment from early childhood through early adolescence: Developmental cascades. *Development and Psychopathology, 22*, 717–735. doi:10.1017/S0954579410000416
- Bornstein, M. H., Hahn, C. S., & Suwalsky, J. T. (2013). Developmental pathways among adaptive functioning and externalizing and internalizing behavioral problems: Cascades from childhood into adolescence. *Applied Developmental Science, 17*, 76–87. doi:10.1080/10888691.2013.774875
- Brown, T. A. (2015). *Confirmatory factor analysis for applied research* (2nd ed.). New York: Guilford Press.
- Burt, K. B., & Roisman, G. I. (2010). Competence and psychopathology: Cascade effects in the NICHD study of early child care and youth development. *Development and Psychopathology, 22*, 557–567. doi:10.1017/S0954579410000271
- Byrne, B. M. (1998). *Structural equation modeling with LISREL, PRELIS, and SIMPLIS: Basic concepts, applications, and programming*. Mahwah, NJ: Erlbaum.
- Byrne, B. M., Shavelson, R. J., & Muthén, B. O. (1989). Testing for the equivalence of factor covariance and mean structures: The issue of partial measurement invariance. *Psychological Bulletin, 105*, 456–466. doi:10.1037/0033-2909.105.3.456
- Capaldi, D. M. (1992). Co-occurrence of conduct problems and depressive symptoms in early adolescent boys: II. A 2-year follow-up at Grade 8. *Development and Psychopathology, 4*, 125–144. doi:10.1017/s0954579400005605
- Centre for Evaluation and Monitoring. (2013). *InCAS technical report*. Durham: Author.
- Chaplin, T. M., Cole, P. M., & Zahn-Waxler, C. (2005). Parental socialization of emotion expression: Gender differences and relations to child adjustment. *Emotion, 5*, 80–88. doi:10.1037/1528-3542.5.1.80
- Cole, D. A., & Maxwell, S. E. (2003). Testing mediational models with longitudinal data: Questions and tips in the use of structural equation modeling. *Journal of Abnormal Psychology, 112*, 558–577. doi:10.1037/0021-843X.112.4.558
- Deighton, J., Humphrey, N., Patalay, P., Belsky, J., Wolpert, M., & Vostanis, P. (in press). Longitudinal pathways between mental health difficulties and academic performance during middle childhood and early adolescence. *British Journal of Developmental Psychology*.
- De Los Reyes, A., Augenstein, T. M., Wang, M., Thomas, S. A., Drabick, D. A., Burgers, D. E., & Rabinowitz, J. (2015). The validity of the multi-informant approach to assessing child and adolescent mental health. *Psychological Science, 141*, 858–900. doi:10.1037/a0038498
- Department for Communities and Local Government. (2015). *The English Indices of Deprivation*. Retrieved from https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/465791/English_Indices_of_Deprivation_2015_-_Statistical_Release.pdf
- Department for Education. (2011). *Interim results for key stage 2 and 3 national curriculum assessments in England, 2010/11* (SFR 18/2011). Retrieved from https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/219052/sfr18-2011v2.pdf
- Department for Education. (2012a). *National curriculum assessments at key stage 2 in England, 2011/2012* (rev.; SFR33/2012). Retrieved from https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/219151/sfr33-2012v2.pdf
- Department for Education. (2012b). *Schools, pupils, and their characteristics, January 2012* (SFR10/2012). Retrieved from https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/219260/sfr10-2012.pdf
- Department for Education. (2013). *Special educational needs in England, January 2013* (SFR 30/2013). Retrieved from https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/225699/SFR_30-2013_Text.pdf
- Department for Education. (2014). *National curriculum assessments at key stage 2 in England, 2014* (rev.; SFR 50/2014). Retrieved from https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/428838/SFR50_2014_Text.pdf
- Ehrlich, K. B., Miller, G. E., & Chen, E. (2016). Childhood adversity and adult physical health. In D. Cicchetti (Ed.), *Developmental psychopathology: Risk, resilience, and intervention* (3rd ed., Vol. 4, pp. 1–42). Hoboken, NJ: Wiley.
- Englund, M. M., & Siebenbruner, J. (2012). Developmental pathways linking externalizing symptoms, internalizing symptoms, and academic competence to adolescent substance use. *Journal of Adolescence, 35*, 1123–1140. doi:10.1016/j.adolescence.2012.03.004
- Evans, G. W., Li, D., & Whipple, S. S. (2013). Cumulative risk and child development. *Psychological Bulletin, 139*, 1342–1396. doi:10.1037/a0031808
- Fink, E., Patalay, P., Sharpe, H., Holley, S., Deighton, J., & Wolpert, M. (2015). Mental health difficulties in early adolescence: A comparison of two cross-sectional studies in England from 2009 to 2014. *Journal of Adolescent Health, 56*, 502–507. doi:10.1016/j.jadohealth.2015.01.023
- Goodman, R. (1997). The Strengths and Difficulties Questionnaire: A research note. *Journal of Child Psychology and Psychiatry, 38*, 581–586. doi:10.1111/j.1469-7610.1997.tb01545.x
- Green, H., McGinnity, A., Meltzer, H., Ford, T., & Goodman, R. (2005). *Mental health of children and young people in Great Britain*. Newport: Office for National Statistics.
- Grover, R. L., Ginsburg, G. S., & Ialongo, N. (2007). Psychosocial outcomes of anxious first graders: A seven-year follow-up. *Depression and Anxiety, 24*, 410–420. doi:10.1002/da.20241
- Hamaker, E. L., Kuiper, R. M., & Grasman, R. P. P. (2015). A critique of the cross-lagged panel model. *Psychological Methods, 20*, 102–116. doi:10.1037/a0038889
- Harter, S. (1999). *The construction of self: A developmental perspective*. New York: Guilford Press.
- Hoffman, L. W. (1972). Early childhood experiences and women's achievement motives. *Journal of Social Issues, 28*, 129–155. doi:10.1111/j.1540-4560.1972.tb00022.x

- House of Commons Education and Health Committees. (2017). *Children and young people's mental health—The role of education*. London: House of Commons.
- Hu, L. T., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling*, 6, 1–55. doi:10.1080/10705519909540118
- Humphrey, N., Barlow, A., Wigelsworth, M., Lendrum, A., Pert, K., Joyce, C., . . . Turner, A. (2016). A cluster randomized controlled trial of the promoting alternative thinking strategies (PATHS) curriculum. *Journal of School Psychology*, 58, 73–89. doi:10.1016/j.jsp.2016.07.002
- Huston, A. C., & Ripke, M. N. (2006). Middle childhood: Contexts of development. In A. C. Huston & M. N. Ripke (Eds.), *Developmental contexts in middle childhood* (pp. 1–22). Cambridge: Cambridge University Press.
- Hutchings, M. (2015). *Exam factories? The impact of accountability measures on children and young people*. London: National Union of Teachers.
- Kaplan, H. B. (1975). Sequelae of self-derogation predicting from a general theory of deviant behavior. *Youth & Society*, 7, 171–197.
- Keenan, K., & Shaw, D. (1997). Developmental and social influences on young girls' early problem behavior. *Psychological Bulletin*, 121, 95–113.
- Keith, T. Z. (2015). *Multiple regression and beyond: An introduction to multiple regression and structural equation modeling* (2nd ed.). New York: Taylor & Francis.
- Kiesner, J. (2002). Depressive symptoms in early adolescence: Their relations with classroom problem behavior and peer status. *Journal of Research on Adolescence*, 12, 463–478. doi:10.1111/1532-7795.00042
- Kline, R. B. (2016). *Principles and practice of structural equation modeling* (4th ed.). New York: Guilford Press.
- Lahey, B. B., Loeber, R., Burke, J., Rathouz, P. J., & McBurnett, K. (2002). Waxing and waning in concert: Dynamic comorbidity of conduct disorder with other disruptive and emotional problems over 17 years among clinic-referred boys. *Journal of Abnormal Psychology*, 111, 556–567. doi:10.1037/0021-843X.111.4.556
- Leaper, C., & Friedman, C. K. (2007). The socialization of gender. In J. E. Grusec & P. D. Hastings (Eds.), *Handbook of socialization: Theory and research* (pp. 561–587). New York: Guilford Press.
- Lessof, C., Ross, A., Brind, R., Bell, E., & Newton, S. (2016). *Longitudinal Study of Young People in England Cohort 2: Health and wellbeing at wave 2 research report*. London: Department for Education.
- Little, T. D., Preacher, K. J., Selig, J. P., & Card, N. A. (2007). New developments in latent variable panel analyses of longitudinal data. *International Journal of Behavioral Development*, 31, 357–365. doi:10.1177/0165025407077757
- Masten, A. S., Burt, K. B., & Coatsworth, D. J. (2015). Competence and psychopathology in development. In D. Cicchetti & D. J. Cohen (Eds.), *Developmental psychology* (2nd ed., pp. 696–738). Hoboken, NJ: Wiley.
- Masten, A. S., & Cicchetti, D. (2010). Developmental cascades. *Development and Psychopathology*, 22, 491–495. doi:10.1017/S0954579410000222
- Masten, A. S., Roisman, G. I., Long, J. D., Burt, K. B., Obradovic, J., Riley, J. R., . . . Tellegen, A. (2005). Developmental cascades: Linking academic achievement and externalizing and internalizing symptoms over 20 years. *Developmental Psychology*, 41, 733–746. doi:10.1037/0012-1649.41.5.733
- McCarty, C. A. (2008). Adolescent school failure predicts depression among girls. *Journal of Adolescent Health*, 43, 180–187. doi:10.1016/j.jadohealth.2008.01.023
- McLoyd, V. C. (1998). Socioeconomic disadvantage and child development. *American Psychologist*, 53, 185–204. doi:10.1037/0003-066X.53.2.185
- Merrell, C., & Tymms, P. (2007). Identifying reading problems with computer-adaptive assessments. *Journal of Computer Assisted Learning*, 23, 27–35. doi:10.1111/j.1365-2729.2007.00196.x
- Mesman, J., Bongers, I. L., & Koot, H. M. (2001). Preschool developmental pathways to preadolescent internalizing and externalizing problems. *Journal of Child Psychology and Psychiatry*, 42, 679–689. doi:10.1111/1469-7610.00763
- Moilanen, K. L., Shaw, D. S., & Maxwell, K. L. (2010). Developmental cascades: Externalizing, internalizing, and academic competence from middle childhood to early adolescence. *Development and Psychopathology*, 22, 635–653. doi:10.1017/S0954579410000337
- Morgan, P. L., Farkas, G., Tufis, P. A., & Sperling, R. A. (2008). Are reading and behavior problems risk factors for each other? *Journal of Learning Disabilities*, 41, 417–436. doi:10.1177/0022219408321123
- Muthén, L. K., & Muthén, B. O. (1998–2012). *Mplus user's guide* (7th ed.). Los Angeles: Author.
- Newsom, J. T. (2015). *Longitudinal structural equation modeling: A comprehensive introduction*. New York: Routledge.
- Obradovic, J., Burt, K. B., & Masten, A. S. (2010). Testing a dual cascade model linking competence and symptoms over 20 years from childhood to adulthood. *Journal of Clinical Child and Adolescent Psychology*, 39, 90–102. doi:10.1080/15374410903401120
- Patterson, G. R., & Capaldi, D. M. (1990). A mediational model for boys' depressed mood. In J. Rolf, A. S. Masten, D. Cicchetti, K. H. Nuechterlein, & S. Weintraub (Eds.), *Risk and protective factors in the development of psychopathology* (pp. 141–163). New York: Cambridge University Press.
- Pomerantz, E. M., Altermatt, E. R., & Saxon, J. L. (2002). Making the grade but feeling distressed: Gender differences in academic performance and internal distress. *Journal of Educational Psychology*, 94, 396–404. doi:10.1037/0022-0663.94.2.396
- Priess, H. A., & Lindberg, S. M. (2011). Gender intensification. In R. J. R. Levesque (Ed.), *Encyclopedia of adolescence*. New York: Springer.
- Ravens-Sieberer, U., Auquier, P., Erhart, M., Gosch, A., Rajmil, L., Bruil, J., . . . The European KIDSCREEN Group. (2007). The KIDSCREEN-27 quality of life measure for children and adolescents: Psychometric results from a cross-cultural survey in 13 European countries. *Quality of Life Research*, 16, 1347–1356. doi:10.1007/s11136-007-9240-2
- Raviv, T., Taussig, H. N., Culhane, S. E., & Garrido, E. F. (2010). Cumulative risk exposure and mental health symptoms among maltreated youth placed in out-of-home care. *Child Abuse and Neglect*, 34, 742–751. doi:10.1016/j.chiabu.2010.02.011
- Robitail, S., Ravens-Sieberer, U., Simeoni, M. C., Rajmil, L., Bruil, J., Power, M., . . . The KIDSCREEN Group. (2007). Testing the structural and cross-cultural validity of the KIDSCREEN-27 quality of life questionnaire. *Quality of Life Research*, 16, 1335–1345. doi:10.1007/s11136-007-9241-1
- Roeser, R. W., van der Wolf, K., & Strobel, K. R. (2001). On the relation between social-emotional and school functioning during early adolescence preliminary findings from dutch and american samples. *Journal of School Psychology*, 39, 111–139. doi:10.1016/S0022-4405(01)00060-7
- Rogosch, F. A., Oshri, A., & Cicchetti, D. (2010). From child maltreatment to adolescent cannabis abuse and dependence: A developmental cascade model. *Development and Psychopathology*, 22, 883–897. doi:10.1017/S0954579410000520
- Rosenfeld, S., & Mouzon, D. (2013). Gender and mental health. In C. S. Aneshensel, J. C. Phelan, & A. Bierman (Eds.), *Handbook of the sociology of mental health* (pp. 277–296). New York: Springer.
- Salla, J., Michel, G., Pingault, J. B., Lacourse, E., Paquin, S., Galera, C., . . . Cote, S. M. (2016). Childhood trajectories of inattention-hyperactivity and academic achievement at 12 years. *European Child & Adolescent Psychiatry*, 25, 1195–1206. doi:10.1007/s00787-016-0843-4
- Sameroff, A. (2006). Identifying risk and protective factors for healthy child development. In A. Clarke-Stewart & J. Dunn (Eds.), *Families count: Effects on child and adolescent development* (pp. 53–76). Cambridge: Cambridge University Press.
- Schwartz, D., Gorman, A. H., Nakamoto, J., & McKay, T. (2006). Popularity, social acceptance, and aggression in adolescent peer groups: Links with academic performance and school attendance. *Developmental Psychology*, 42, 1116–1127. doi:10.1037/0012-1649.42.6.1116
- Snijders, T. A. B., & Bosker, R. J. (2012). *Multilevel analysis: An introduction to basic and advanced multilevel modeling*. London: Sage.
- Steenkamp, J.-B. E. M., & Baumgartner, H. (1998). Assessing measurement invariance in cross-national consumer research. *Journal of Consumer Research*, 25, 78–90. doi:10.1086/209528
- Stewart, E. B. (2008). School structural characteristics, student effort, peer associations, and parental involvement: The influence of school- and individual-level factors on academic achievement. *Education and Urban Society*, 49, 179–204. doi:10.1177/0013124507304167
- Stipek, D., & Miles, S. (2008). Effects of aggression on achievement: Does conflict with the teacher make it worse? *Child Development*, 79, 1721–1735. doi:10.1111/j.1467-8624.2008.01221.x
- Stone, L. L., Otten, R., Engels, R. C. M. E., Vermulst, A. A., & Janssens, J. M. A. M. (2010). Psychometric properties of the parent and teacher versions of the Strengths and Difficulties Questionnaire for 4- to 12-year-olds: A review. *Clinical Child and Family Psychology Review*, 13, 254–274. doi:10.1007/s10567-010-0071-2
- Vaillancourt, T., Brittain, H. L., McDougall, P., & Duku, E. (2013). Longitudinal links between childhood peer victimization, internalizing and

- externalizing problems, and academic functioning: Developmental cascades. *Journal of Abnormal Child Psychology*, *41*, 1203–1215. doi:10.1007/s10802-013-9781-5
- van Lier, P. A., & Koot, H. M. (2010). Developmental cascades of peer relations and symptoms of externalizing and internalizing problems from kindergarten to fourth-grade elementary school. *Development and Psychopathology*, *22*, 569–582. doi:10.1017/S0954579410000283
- Weeks, M., Ploubidis, G. B., Cairney, J., Wild, T. C., Naicker, K., & Colman, I. (2016). Developmental pathways linking childhood and adolescent internalizing, externalizing, academic competence, and adolescent depression. *Journal of Adolescence*, *51*, 30–40. doi:10.1016/j.adolescence.2016.05.009
- Whitcomb, S. A., & Merrell, K. W. (2013). *Behavioral, social, and emotional assessment of children and adolescents* (4th ed.). New York: Routledge.
- Wigelsworth, M., Qualter, P., & Humphrey, N. (2016). Emotional self-efficacy, conduct problems, and academic attainment: Developmental cascade effects in early adolescence. *European Journal of Developmental Psychology*. Advance online publication. doi:10.1080/17405629.2016.1180971
- Winsler, A., & Wallace, G. L. (2002). Behavior problems and social skills in preschool children: Parent-teacher agreement and relations with classroom observations. *Early Education & Development*, *13*, 41–58. doi:10.1207/s15566935eed1301_3
- Zhang, X. (2013). The longitudinal interplay of psychopathology and social competence during Chinese children's transition to preschool. *Infant and Child Development*, *22*, 198–215. doi:10.1002/icd.1780