Developmental pathways of physical aggression from infancy to early school age

Ane Nærde [1*], Harald Tage Janson[1], Mike Stoolmiller[1,2]

[1]The Norwegian Center for Child Behavioral Development, Oslo, Norway

[2]Oregon Research Institute, Eugene, Oregon, USA

*Corresponding author

E-mail: ane.narde@nubu.no (AN)

Abstract

This study identified latent trajectories of physical aggression from infancy to preschool age and tested for (a) early parent, parenting and child predictor effects on trajectory membership and (b) trajectory-effects net of parent, parenting, and child predictor effects on Grade 2 social, behavioral and academic functioning. We used data from the Behavior Outlook Norwegian Developmental Study (BONDS), comprising 1,159 children (559 girls). Parents reported on risk and protective factors, and on physical aggression from 1 to 5 years; teachers reported on Grade 2 outcomes. We employed latent class growth curve analyses to identify nine aggression trajectories. In fully adjusted models testing simultaneously all associations among predictors, trajectories, and outcomes, maternal and paternal harsh parenting, child gender, and sibling presence predicted trajectory membership, which significantly predicted Grade 2 externalizing. Child gender had a pervasive influence on all outcomes as well as on trajectory membership. This is the first trajectory study that attempts to sort out which predictors are most proximal, more distal, or just confounded, with their relative direct effect sizes, and to link early paternal as well as maternal harsh parenting with children's development of physical aggression from infancy to preschool age. Our findings underscore the need to include fathers in developmental research and early prevention and intervention efforts.

Introduction

In this study, we take a deep dive into the developmental patterns of physical aggression during early childhood and explore how these longitudinal patterns influence children's social, behavioral and academic functioning in school age when simultaneously considering early individual and contextual risk and protective factors. During the past few decades, it has become increasingly evident that children's use of physical aggression is predictable, normative, and close to ubiquitous during certain developmental periods. Indeed, behaviors such as kicking, hitting, and pushing have their origins in early childhood and serve developmentally salient and adaptive functions. At the same time, physical aggression may be highly problematic and is in fact considered as the most serious and socially unacceptable form of aggression, being ontogenetically antecedent to less "serious" forms such as verbal aggression or indirect aggression [1].

The systematic study of how physical aggression develops already from early childhood onwards is still recent [2], as most existing research focuses on aggression beginning in toddlerhood or preschool age [but see e.g., 3, 4]. Still, theoretical and empirical research efforts together with advances in quantitative theory and practice have moved the field forward. In short, there is compelling evidence that: (a) children start to physically aggress at an early age [5-7]; (b) such behavior typically peaks around 2-3 years of age followed by a steep decline [3, 8-10]; and (c) a small proportion of children, more often boys, maintain a high and stable frequency of physical aggression during middle childhood that is associated with a range of negative outcomes [11-17]. Yet, existing longitudinal research is limited by mainly starting out from the second year of life, comprising relatively few observations of aggressive behavior, relying solely on assessments from mothers, and

primarily *about* mothers, and not testing simultaneously all associations among predictors, trajectories, and outcomes. We add to the existing literature by identifying trajectories of physical aggression from infancy to preschool age (i.e., 1 to 5 years) in a large population-based sample using data from multiple informants and examining social, behavioral and academic adjustment in second grade as a function of these trajectories as well as of empirically driven early predictors. Following previous research [7, 18], we use the term physical aggression to denote physical acts directed against others in the form of observable child behaviors such as kicking and hitting, without any consideration of intentionality.

Developmental trajectories of physical aggression across early childhood

In a seminal paper at the turn of the century, Tremblay [1] reviewed the state of knowledge regarding the origins of aggressive behavior and highlighted the need to focus on the first few years of development if we are to understand physical aggression and prevent chronic cases. He argued that infants' physical aggression typically had not been deemed developmentally significant and stated that "the exciting work that needs to be done is to study the developmental trajectories of these aggressive behaviors at the intra-individual level during infancy and toddlerhood, and their relationship with later development" [1, p. 134]. Some twenty years later, this task seems well underway, with published results from several longitudinal studies mapping the development of aggression utilizing state-of-the-art analytic techniques. Among other things, empirical evidence shows that there is heterogeneity in the early development of physical aggression [2, 13, 14, 17, 19-21]. In the following, we provide an overview of research applying a developmental trajectory approach, that is, the identification of prototypic subgroup curves that represent different

developmental patterns of aggression [e.g., 16]. We focus on summarizing key results from large-scale longitudinal studies specifically exploring developmental pathways of physical aggression starting in toddlerhood, which is in accord with the focus of the current study.

Research addressing the course of aggression has found varying numbers of developmental trajectories. This is likely related to sample variations like age range, size, and risk factor composition, the number of repeated assessments and variation in the targeted constructs that measurements are able to capture, as well as model specification and model selection criteria [8, 22]. Several studies focusing on the course of physical aggression through early childhood have identified either 3 [8, 10, 17, 20], 4 [19, 23-27], or 5 trajectories of aggression [13, 21]. Some have also reported differing numbers of trajectory groups for girls and boys (i.e., 3 for girls and 5 for boys [2]. Results from two studies [13, 21] utilizing a sample of over 1,100 children identified 5 distinct aggression trajectories from age 2 through 9 years including a "very low" group characterized by low frequency in early childhood followed by a rapid decline (~45% of the sample), a “low-aggression” group (~25%), also quite low in aggression throughout, although higher than the lowest group, a "moderate-decline" group, showing a marked decline over time (~12%), a "moderate" group being moderately stable, albeit slightly declining (~15%), and a "high" group showing the highest level of aggression (~3%).

There are several similarities across findings. Groups of children who display high physical aggression at young ages tend to continue to display high aggression over time. Moreover, studies have typically identified at least one trajectory group showing low levels of aggression across time, as well as one characterized by a high and stable level. Whereas several, but not all, studies have identified trajectories characterized by declining physical aggression [21, 24, 27], no late onset high physical aggression groups tend to be identified.

Given empirical evidence that children start to physically aggress already at 6 to 8 months of age [5, 6], it is noteworthy that only a small minority of longitudinal studies addressing developmental trajectories of aggression have started out before the second year of life. Urges for researchers conducting longitudinal developmental research to start up in early childhood and conduct multiple observations across a host of data sources have been expressed, including by Jennings and Reingle [22] in their comprehensive meta-review of latent group-based trajectories of violence, aggression, and delinquency across the life span. Examples of longitudinal studies starting out this early include the Quebec Longitudinal Study of Child Development [QLSCD; see e.g., 2, 20, 28], the National Longitudinal Survey of Children and Youth in Canada [NLSCY; see e.g., 19], the NICHD Study of Early Child Care and Youth Development in the US [SECCYD; see e.g., 14, 21], and studies by Lorber et al., [4, 7] and Tremblay et al. [10]. Finally, it should be mentioned that despite the call from Tremblay [1] for longitudinal studies to include repeated measurements of aggressive behavior at least every two to three months from birth to 36 months, and ideally every 6 months from 36 months onward, existing evidence is typically based on data collected about yearly [2, 3, 27], every other year [8, 19], or a mix (17, 21). [But see e.g., 4, 7, 29].

Predictors of physical aggression trajectories

One implicit assumption when identifying population heterogeneity is that groups of children exhibiting atypical developmental patterns have specific risk characteristics [2, 8, 21]. Considerable efforts have thus been devoted to identifying early predictors of trajectory group membership. Results from twin research [30, 31] suggest that individual differences in the frequency of physical aggression may be substantially driven by genetically based

124 factors, and recent evidence [32, 33] indicate that epigenetic mechanisms could be involved
125 in the development of chronic aggression, highlighting the need to identify early bio-psycho-
126 social mechanisms [33, 34]. Longitudinal patterns of aggression have also been linked to
127 several environmental predictors (including gene-environment correlations), comprising *risk*
128 as well as *protective* factors. Protective factors contribute to positive outcomes especially in
129 the context of increased risk, but also regardless of risk level (the latter is sometimes
130 distinguished as *promotive* factors; a distinction which is meaningful but empirically not so
131 clear-cut [see e.g., 21, 35]. These risk and protective factors fit within Bronfenbrenner's [36]
132 socio-ecological model addressing the ecology of human development. In line with this,
133 Labella and Mason [37] recently discussed the most salient family-based risk and protective
134 factors for the development of aggression and violence including those relating to children's
135 proximal developmental contexts, which can be grouped into *parent/family factors* (i.e.,
136 sociodemographic risk, parental psychopathology), *parenting factors* (i.e., parenting style),
137 and *child factors* (i.e., gender, temperament).

138 ***Parent/family factors***

139 Sociodemographic risk and parental psychopathology represent two of the strongest
140 and most persistent predictors of multiple types of child maladjustment, including
141 physical aggression [37- 40]. For instance, results from the SECCYD NICHD ECCRN [21]
142 indicated that family poverty, low maternal education, and single parenthood predicted
143 higher and more stable maternal reported physical aggression trajectories among more
144 than 1,100 children from 2 years to third grade. Relatedly, Côté et al. explored maternal
145 reported trajectories of physical aggression from 2 to 8 years [19] and from 2 to 11 years [8]
146 among 1,183 children participating in the NLSCY and found that those in the high aggression
147 group were more likely than the ones following lower aggression groups to come from low-

148 income families, with mothers who had not completed high school. Similar findings were
149 also reported by Tremblay et al. [10] in that low income, early maternal childbearing,
150 smoking during pregnancy, the presence of similarly aged siblings, maternal antisocial
151 behavior in high school, and family dysfunction predicted the high aggression group at ages
152 17, 30, and 42 months among 572 Canadian families. Finally, Teymoori et al. [2] addressed
153 predictors of maternal reported physical aggression trajectories separately for boys and girls
154 among 2,223 Canadian children participating in the QLSCD from 1.5 to 13 years of age. The
155 girls in the high trajectory group were significantly different from those in the low on 9 out
156 of 12 family risk factors, including parent's education, depression, and antisocial behavior in
157 adolescence, household income, socioeconomic status, and number of siblings at birth and
158 at 17 months. The boys in the highest group furthermore had parents with lower education
159 and socioeconomic status, higher depression, and mothers who were younger at first
160 childbirth. According to the much-studied family stress model [41-44], poverty and financial
161 instability undermine parents' psychological well-being and lead to persistently elevated
162 levels of stress, which in turn predict their parenting practices and subsequent child
163 functioning.

164 Having siblings (younger or older) furthermore constitutes a major risk factor for
165 aggressive behavior [45-50]. Siblings represent easily accessible targets for and sources of
166 aggression, and physical aggression amongst siblings is the most prevalent, albeit often
167 overlooked form of family violence, for which the lack of a comprehensive theory has long
168 been noted [51, 52]. Relevant theoretical perspectives mainly include family systems theory,
169 evolutionary psychology, and social learning theories (49-52]. For instance, Patterson [49,
170 50] suggested a sibling-training hypothesis wherein coercive sibling interactions set the
171 stage for the development and escalation of aggressive behavior through direct practice,

172 modeling, and reinforcement. In accord with Patterson's coercion theory [53], parental
173 reactions and handling of their children's conflicts and aggressive behaviors further
174 contribute to exacerbate or diminish such sibling behavior [48].

175 ***Parenting factors***

176 There is much evidence to support that parenting style is predictive of children's
177 externalizing behavior. For example, results from the SECCYD NICHD ECCRN [21] indicated
178 that less sensitive and involved maternal parenting predicted higher and more stable
179 aggression trajectories from 2 years to third grade. Moreover, Tremblay et al. [10] found
180 that maternal coercive parenting behavior predicted membership in the high aggression
181 group from 17 to 42 months, which was also reported by Côté et al. [19] across 2 to 8 years.
182 Among the most well-established theories of development of aggression and externalizing
183 problem behavior is Patterson's [53] coercion theory, wherein child externalizing develops
184 from transactional reinforcement processes between the child and important others in the
185 early developmental period. Parenting behaviors characterized by lack of skills, warmth and
186 positive affect in a discipline encounter can lead the child to react with (even more) negative
187 and challenging behaviors, which may make the parents feel overwhelmed and frustrated
188 and possibly giving in to the child just to restore peace. This parental negative
189 reinforcement for child acting out behavior is what perpetuates an adverse parent-child
190 cycle [53, 54]. Such suboptimal parenting behaviors are typically associated with parental
191 psychopathology and depression (in research mainly focusing on mothers; [55-56] but see
192 also Cheung and Theule [57]), which illustrates how these risk factors tend to appear and
193 operate in tandem.

194 The recurring failure to integrate fathers into empirical research addressing child
195 development in general, and development of physical aggression in particular (although see

196 a recent exception by Yang et al. [58]), is serious and unfortunate as it might bring about
197 erroneous conclusions, such as misattributing influences of fathers to mothers or leaving
198 important sources of variation unexplained. Cabrera et al. [59] and Cabrera [60] recently
199 discussed the puzzling fact that fathers are still mostly absent from parenting research,
200 despite strong suggestions that traditional mother-focused models of developmental
201 influences are outdated. Father involvement is particularly high in societies where this is
202 facilitated by social policies [61]. In Norway, both parents have the right to a lengthy paid
203 leave in connection with childbirth, of which a particular quota (6 weeks in 2006, when the
204 current sample was recruited) is reserved for the father. Indeed, Norwegian fathers started
205 spending significantly more time with their children following the introduction of this quota
206 [62], and time-use data show that they tend to be highly involved in caregiving [63],
207 providing a unique opportunity to assess the importance of factors relating to mothers and
208 fathers alike (e.g., parenting behavior and psychopathology) for the development of
209 aggression. Finally, the dearth of research examining paternal sensitivity in conjunction with
210 maternal within developmental psychopathology more generally was highlighted in a recent
211 meta-analytic review [64] addressing (observed) paternal sensitivity and children's cognitive
212 and socioemotional outcomes.

213 ***Child factors***

214 It is well established that boys use direct forms of aggression (i.e., physical or verbal)
215 more often than girls. In line with this, Côté et al. [19] reported that children in the high
216 aggression trajectory group from 2 to 8 years (amongst other things) were more likely to be
217 boys. Researchers has frequently addressed various aspects of these gender differences,
218 including *when* they emerge, the *types* of aggressive behaviors boys and girls generally
219 engage in, and *alternative hypotheses* for such differences [5, 19, 23, 65-67]. Empirical

220 evidence suggests that boys and girls use physical force against others (e.g., pulling peoples
221 hair, hitting, biting) at similar rates in infancy and early toddlerhood [5, 66, 68], whereas
222 boys show significantly more physical aggression in the next few years [3, 5, 25, 69].
223 According to Hay [5], no single factor explains the widening gender gap, but rather a
224 developmental cascade of biological and social factors. One important contribution of
225 trajectory- or group-based approaches is the finding that emerging gender differences
226 primarily may be accounted for by more boys than girls making up the small group of
227 children who continue to deploy aggression at high rates [11].

228 Finally, the contribution of child temperament to the early development of
229 externalizing problem behaviors is well documented, including high negative emotionality,
230 activity level and reactivity, along with low effortful control, inadaptability, persistence, and
231 negative mood - sometimes subsumed under the broad construct difficult temperament [5,
232 25, 29, 70, 71]. Several theoretical models (see Slagt et al. [72] for overview) postulate
233 individual differences among children in their general susceptibility to parenting and
234 environmental influences based on their temperamental characteristics. A meta-analysis
235 [72] on parenting-by-temperament interactions reported partly support for Belsky's [73]
236 differential susceptibility hypothesis wherein young children with difficult temperament are
237 more likely to develop externalizing problems in the presence of negative parenting
238 compared with those having an easier temperament [47, 70, 74-76].

239 Outcomes of physical aggression trajectories

240 A key question when mapping the developmental course of aggression across early
241 childhood is the degree to which the trajectories predict subsequent behavior and
242 adjustment, and whether moderate versus elevated levels of aggression predict different

243 patterns of outcomes [12, 13, 65, 76, 77]. Much research exploring outcomes of aggression
244 trajectories has focused on more serious delinquency and antisocial behavior in adolescence
245 and emerging adulthood [12, 15, 16, 78, 79], as well as in adulthood [80-82], whereas less
246 has addressed a broader range of social and academic outcomes that may also confer risk
247 [but see e.g., 13, 17, 21, 24, 83]. Results from two studies using data from the SECCYD [13,
248 21] show that both level and stability of aggression from age 2 to 9 predict middle childhood
249 academic achievement and social functioning at home, school, and in the peer group. In the
250 first study [21], the five identified trajectories were compared on age 9 measures of
251 academic competence and cognitive functioning, teacher reported behavior problems and
252 social competence, self-reported peer relationships, friendships, and hostile attribution
253 biases, observed behavior in school, and mother-reported discipline and mother-child
254 conflict. Relatedly, in a second study, Campbell et al. [13] explored age 9 to 12 adjustment
255 (i.e., teacher reported behavior problems, social competence, and academic achievement,
256 child self-reported peer relations, depression, and risk taking, observed classroom behavior
257 and friendship interaction) as a function of the same 5 trajectories. Overall, the results
258 indicated that outcomes varied by trajectory group. As expected, children in the high- and
259 moderate-stable groups showed continuing behavioral and social problems relative to those
260 very low in aggression; the high-stable group evidenced signs of significant antisocial
261 behavior, whereas the moderate-stable showed symptoms of ADHD and more minor peer
262 problems [13, 21]. The researchers suggested that the two groups seemed to be on
263 different developmental pathways. Moreover, children in the moderate decreasing group
264 displaying transient and age-related aggression that disappeared before school age, showed
265 good adjustment both at age 9 [21], and across the transition to adolescence [13]. A more
266 surprising finding was the differences between the two low aggression groups (i.e., low-

stable, and very-low), given that the aggression in the low-stable group initially seemed trivial. Still, relative to the very-low group, children in the low-stable group had elevated scores on several measures of internalizing and externalizing problems and poor social skills in elementary school [13]. Consequently, moderate, or low-level physical aggression that persists from toddlerhood through school entry may signal that adjustment problems will continue or emerge by middle childhood, albeit not at the level of the high-stable trajectory group.

The present study

In the present study, we extend previous research by scrutinizing developmental pathways of physical aggression spanning the age range during which it emerges among most children, when it is particularly prevalent and viewed as normative, as well as the period when such behavior usually declines. The study follows up on previous work [9] utilizing the same sample, which addressed the developmental course of physical aggression from infancy to age 2 applying growth curve modeling rather than a developmental trajectory approach and comprising early predictors but no outcomes. We include a comprehensive set of 20 early predictors and several school-age outcomes, enabling us to disentangle the relative importance of children's longitudinal patterns of physical aggression for their subsequent social, behavioral, and academic functioning, while simultaneously considering the direct and indirect impact of early risk and protective factors. As far as we know, this is the first study to examine trajectories of physical aggression across the developmental period from 1 year until pre-school age including frequent parental reports of child aggression (up to 21 reports with 90% of the sample having at least ten) and to test in the same model the effects of a comprehensive set of early risk and protective factors on

trajectory membership, as well as those of trajectory membership on teacher-reported school-age outcomes net of early predictors. In contrast, most prior research has either focused on effects of precursors on trajectories, or of trajectories on outcomes, or has included a somewhat restricted set of covariates. Finally, we take advantage of very recent advances in statistical methods for trajectory analysis, so called 3-step procedures [84], that were not available for many prior studies, when incorporating multiple predictors and multiple outcomes in the trajectory model to obtain the best possible estimates of both effect sizes and their associated standard errors.

Research issues

We employed state-of-the-art analytic techniques to:

1. Model trajectories of physical aggression from infancy (i.e., 1 year) to preschool age (i.e., 5 years);
2. Predict aggression-trajectory membership from empirically driven parent, parenting, and child predictors measured at child age 0.5 or 1 year both unadjusted and adjusted for all other predictors; and
3. Predict children's teacher-rated Grade 2 adjustment (externalizing and internalizing, social skills, and academic competence) as a function of aggression-trajectory membership as well as the early predictors, both unadjusted and adjusted for all other predictors and for trajectory membership.

Materials and methods

Sample and procedures

We utilized data from the Behavior Outlook Norwegian Developmental Study (BONDS), a population-based longitudinal study of 1,159 children (559 girls and 600 boys)

313 from five municipalities in southeast Norway conducted by the Norwegian Center for Child
314 Behavioral Development. The BONDS is approved by the Regional Committee for Medical
315 and Health Research Ethics and the Norwegian Social Sciences Data Services (approval
316 numbers S-06067; 2009/224a), and all parents provided informed written consent.
317 Participants were recruited from September 13, 2006, until December 31, 2008, by nurses
318 at the 5-month visit at child health clinics, which are attended almost universally. Inclusion
319 criteria were the child being of the appropriate age and one parent being able to participate
320 without an interpreter. Families of 1,931 eligible children were informed about the project.
321 Of the 1,465 (76%) who agreed to be contacted, 1,159 (79%, or 60% of the eligible group)
322 opted to participate. The sample was fairly representative of the general population,
323 however, somewhat biased toward mothers with higher education, fewer immigrant
324 parents, more firstborns, and fewer single mothers [85]. Personal interviews, which
325 included a computerized questionnaire, were conducted when children were aged 6
326 months, 1, 2, 3, and 4 years, and in the child's first school year. Whereas both parents were
327 invited to participate in the first interview, fathers were primarily targeted in the 1- and 3-
328 year waves, and mothers in the 2- and 4-year waves. At 1, 2, and 3 years, the assessment
329 included videotaped structured parent-child interaction tasks, in which more fathers thus
330 participated at ages 1 and 3 years, and more mothers at age 2. The parental interviews
331 (including the interaction tasks) were conducted by trained assistants in local offices (or in
332 the home setting if parents preferred). At two occasions (i.e., 5 years as well as in the child's
333 second school year), lengthier telephone interviews including many of the study's variables
334 were conducted instead of a personal interview. In addition to the yearly main data
335 collections, brief telephone interviews comprising a small selection of items relevant to the
336 study's key outcomes were conducted at about 8 months, 10 months, 1 year and 3 months,

337 1 year and 6 months, 1 year and 9 months, 2 years and 4 months, 2 years and 8 months, and
338 3½ years. Brief telephone interviews were also added 2 and 1 months prior to and 1-, 2-,
339 and 3-months following entry into center-based childcare when the child's date of entry was
340 known to the project and these ages were not covered by another interview. At age four,
341 which was the endpoint for the original participation consent, 1,121 children (96.7%)
342 remained in the study. Parents of 1,027 children (88.6%) renewed their informed written
343 consent for participation in a continuation of the study until second grade. At the Grade 2
344 data collection, 996 children (85.9% of the original sample) remained in the study. Teachers
345 of 901 children contributed reports on the child in Grade 2.

346 Child physical aggression measure

347 Our measurement design for physical aggression included repeated measurements
348 per child across time and as many items on physical aggression as thought to be relevant
349 and feasible at each wave of data collection. Thus, from 10 months onward, each
350 questionnaire (i.e., at 1, 2, 3, and 4 years in the current dataset) and each telephone
351 interview (i.e., in the current dataset including 7 target ages from 1 year and 3 months up to
352 5 years, and in addition prior to and after entry into center-based childcare as described
353 above) included 2 to 9 items on child physical aggression. We used an original instrument
354 developed for the BONDS, with idiomatic wordings and a selection of items widely
355 acceptable to parents of small children, including exclusively physical behavior directed at
356 others covering also the lower range of intensity (Supporting information A). In telephone
357 interviews, respondents were asked if the behavior was true now or in the past two weeks,
358 with a dichotomous (*yes/no*) response format. In the questionnaire format, respondents
359 were asked how often the child engaged in the behavior, with responses given on a seven-

point frequency scale ranging from 1 (*never/not in the past year*) to 7 (*three times daily or more*). The present set of analyses used physical aggression data starting with the 1-year interview and including up to the 5-year telephone interview. The number of regularly scheduled interviews from 1 to 5 years was eleven; a smaller number of interviews occurred when families discontinued participation or declined to participate in specific interviews; a greater number of interviews occurred when additional telephone interviews were added prior to and after entry in center-based day care, and when both father and mother contributed with an interview at the same age. Reports on 1,141 children were available. Table 1 shows item inclusion and frequencies for telephone interviews from around 1 year 3 months to around 5 years, as well as for the 1- to 4-year questionnaires. The number of reports per child ranged from one (7 cases) to 21 (1 case) with a mean of 12.0 (*SD*=2.5), with 90% having at least ten reports and 43% having at least thirteen. In total, 26% of the reports were by fathers.

374 **Table 1. Item inclusion and frequencies for telephone interviews from around 15 to around 60 months and 1- to 4-year questionnaires.**

	Telephone-interview responses (% Yes for past two weeks)							Questionnaire responses (*M* [*SD*] for frequency)[a]			
Target age (months)	**15**	**18**	**21**	**28**	**32**	**42**	**60**	**12**	**24**	**36**	**48**
Age range (months)	**14-16[b]**	**17-19[b]**	**20-22[b]**	**27-29[b]**	**31-33[b]**	**41-44[b]**	**59-61[b]**	**10-15[c]**	**22-29[c]**	**34-39[c]**	**46-51[c]**
Hits you	44 %	59 %	60 %	63 %	56 %	41 %	27 %	1.8 (0.8)	2.2 (0.6)	2.0 (0.5)	1.8 (0.6)
Hits siblings[d]	37 %	57 %	62 %	66 %	61 %	50 %	53 %	1.7 (0.7)	2.3 (0.7)	2.1 (0.7)	2.0 (0.6)
Hits other adults	--	--	--	--	--	--	--	--	1.7 (0.6)	--	--
Pushes someone to get his/her will	38 %	48 %	54 %	59 %	54 %	44 %	32 %	1.6 (0.7)	2.1 (0.7)	2.0 (0.6)	1.9 (0.6)
Pulls someone's hair	--	--[e]	--[e]	19 %	17 %	7 %	4 %	--	1.6 (0.7)	1.4 (0.6)	1.3 (0.5)
Pinches someone	--	--	--	--	--	--	14 %	--	1.6 (0.7)	1.5 (0.6)	1.4 (0.5)
Throws things at others	--[e]	31 %	37 %	35 %	34 %	23 %	18 %	--	1.9 (0.7)	1.8 (0.5)	1.6 (0.5)
Bites someone[f]	--[f]	29 %	25 %	19 %	15 %	6 %	--	--[f]	1.6 (0.6)	1.4 (0.5)	1.2 (0.4)
Kicks someone	--[e]	8 %	10 %	17 %	18 %	18 %	18 %	--	1.4 (0.5)	1.5 (0.5)	1.5 (0.6)
N responses[g]	1603	1391	1211	1127	1098	1034	919	1215	1141	1091	1127
Father responses	23 %	21 %	22 %	23 %	19 %	18 %	7 %	69 %	8 %	70 %	9 %

	Telephone-interview responses (% Yes for past two weeks)							Questionnaire responses (*M* [*SD*] for frequency)[a]			
Target age (months)	**15**	**18**	**21**	**28**	**32**	**42**	**60**	**12**	**24**	**36**	**48**
Age range (months)	**14-16[b]**	**17-19[b]**	**20-22[b]**	**27-29[b]**	**31-33[b]**	**41-44[b]**	**59-61[b]**	**10-15[c]**	**22-29[c]**	**34-39[c]**	**46-51[c]**
N children	1096	1086	1068	1057	1058	1015	880	1101	1075	1024	1059

375 [a] Means and standard deviations are based on the reduced 4-point frequency scale: 1 (never/not in the past year), 2 (on single occasions/1-3 times per
376 month/once per week), 3 (2-3 times per week/1-2 times per day), and 4 (3 times per day or more). [b]Children's age in whole months, i.e., from -1.0 to +2.0
377 months from target age. Mostly due to telephone interviews around entry in center-based childcare, there were 716 (7.9%) telephone-interview responses
378 from 526 different children outside of these age ranges; while those responses were included in the data set, percentages tabulated here include only
379 responses within the indicated age ranges. [c]Age range in whole months for all questionnaire responses. [d]Administered only to children with a similar-aged
380 sibling in the family (who ranged from 39% at age 1 to 82% at age 5). [e]According to study design, this question should not have been included at this age.
381 However, due to interviewer choice, the data set included 10 or fewer responses for this age; a percentage out of this small number is not tabulated here.
382 [f]Excluding during breastfeeding. Data prior to 18 months excluded. [g]Multiple telephone-interviews responses per child up to 42 months occurred as a result
383 of frequent interviews around center-based day care entry. Two questionnaire responses per child, or two 60-month telephone interviews, occurred when
384 both parents participated in the yearly parental interview.

We employed Rasch scaling [86] to equate response formats, link response patterns with varying sets of completed items, and address measurement invariance. The following is a summary; a fuller account is given in Supporting information B. We first performed Rasch scaling in two calibration subsamples – one of questionnaire responses, and one of telephone-interview responses – each consisting of responses covering the age span sampled such that each child contributed only one response. Rasch rating scale analyses suggested that a reduction of the scale from seven to four categories (1 [*never/not in the past year*], 2 [*on single occasions/1-3 times per month/once per week*], 3 [*2-3 times per week/1-2 times per day*], and 4 [*three times daily or more*]) gave a more optimally functioning rating scale. There was almost perfect correlation (r=.986) between item scale value estimates from questionnaires and the corresponding estimates from telephone interviews. In both calibration subsamples, we tested and found reasonable measurement invariance of the Rasch model with respect to item scale value estimates for respondent (mother vs. father) and child gender (boy vs. girl) by means of Rasch-model differential item functioning (DIF) analyses. For child age, six items' scale values were allowed to vary as a function of age based on the DIF analyses in the calibration subsamples and further analyses of item displacement in eleven age-specific subsamples. The arbitrary average difference in person locations that was a result of the two data types (questionnaire vs. telephone interview) was addressed by adding a constant of about 2 to questionnaire data. We obtained person location estimates for all measurements in the sample according to the estimation procedure presented by Linacre [87].

In the final data set after recoding of extreme values and equivalating for forms, the 13,673 person location estimates from the Rasch analyses that we used as our physical aggression measurements ranged from -5.30 to 5.02 with a mean of -1.86 and standard

deviation of 1.87 (variance 3.49; skewness 0.28; kurtosis -0.23). The estimates of the measures' standard error averaged 1.22, and the average squared standard error estimate was 1.67. Thus, Rasch-model based person reliability estimated in the entire sample by means of these terms equaled .52, suggesting modest overall reliability of an individual measurement. On the other hand, from a standard (i.e., 1-class) growth model parameterized the same way as our latent class growth models (4th order polynomial trajectories, see below), the average reliability of the child level aggression was estimated at .88 at age 2.2 years, the growth curve intercept (.82 at 1.2 years; .81 at 4.4 years), and the growth curves accounted for about 55 percent of the total outcome variance. Thus, although an individual aggression measurement at one point in time has only modest reliability, sets of repeated aggression measurements over time are substantially more reliable and suitable for growth curve or latent class growth curve analysis.

The average time-to-time correlation of physical aggression measurements for 2- to 12-month intervals between study design target ages was .46. A simple, non-parametric regression (Loess) plot of the repeated measures of physical aggression vs child age is shown in Figure 1. The developmental trend shows a rapid increase from 10 months to a peak at about 2 years and 2 months, and a slower decrease thereafter.

Fig 1. Loess plot of average child physical aggression by child age

Predictor measures

Parent/family measures

431 **Parental education and age.** Information about parental age, education, and family
432 composition was obtained from the interviews at child ages 6 months and 1 year. Parental
433 age at child age 6 months was dichotomized into 1 (young, i.e., ≤25 years) and 0 (not
434 young). Mothers and fathers reported on their educational level in six categories: 1 (*9-year*
435 *elementary school or shorter*), 2 (*1-2 years of high school*), 3 (*Higher vocational*), 4 (*3 years*
436 *of high school, upper secondary school*), 5 (*College, university up to 4 years [Bachelor, Nurse,*
437 *Teacher, Engineer]*), or 6 (*University, college, more than 4 years [Master's degree]*). The
438 reports were recoded to represent the total years of education, the six categories taken to
439 represent 9, 10.5, 12, 12, 15.5, and 18 years of education, respectively.
440 **Family composition.** Single parent status, from child age 1 year if a report existed,
441 and from 6 months otherwise, was coded 1 and not being single was coded 0.
442 **Financial stress.** Parents' reports at child age 1 year of having more long-standing
443 financial difficulties in the past year (rent, mortgage, obligations, etc.) We obtained reports
444 from at least one parent for 1096 children, coded 1 if at least one parent reported the
445 presence of such difficulties, and 0 otherwise.
446 **Couple relationship.** We used mothers' questionnaire reports at child age 6 months
447 of the relationship between her and her husband/partner on a shortened and modified 10-
448 item version of the Relationship satisfaction scale [88], answered on a 6-point Likert scale
449 from 1 (*very much agree*) to 6 (*very much disagree*). Cronbach's alpha was .90 for 1084
450 mothers.
451 **Parental depression and anxiety.** A 12-item version of the Hopkins Symptom Check
452 List [89] from the 6-month interview was used as a measure of symptoms of depression and
453 anxiety. The 12 items are a subset of those in the 25-item version, a widely used self-report
454 symptom inventory tapping the most common forms of psychopathology in non-psychiatric

455 populations (i.e., anxiety and depression; [90]). The response format is a Likert-type scale
456 from 1 (*not at all*) to 4 (*extremely*). We computed a mean score of the 12 items. Cronbach's
457 alpha was .88 for 1130 mothers and .87 for 677 fathers who contributed reports.

458 Parenting predictor measures

459 **Parent sensitivity.** Parent sensitivity was measured based on ratings of videotaped
460 parent-child structured interaction tasks at age 1 year. Fathers were the primarily targeted
461 participants at this age, and more fathers took part in the assessment than mothers, but
462 when fathers were not able or willing to come, the mother took part instead. We thus
463 obtained videotapes of father-child interaction for the greater part of the sample, but for
464 mother-child interaction for a smaller part. The interaction tasks were selected based on
465 their capacity to elicit parent and child behaviors associated with subsequent child
466 adjustment [91, 92] following literature reviews and discussions with international experts.
467 The tasks strived at being gender non-stereotypic and included *free-play*, where parents
468 were asked to play with their child as they liked with a provided set of toys (4 min.); *clean*
469 *up*, where parents were asked to put the toys away and told that the child could help but
470 was not required to (2 min.); *structured play*, where two sets of toys were presented, a
471 shape sorter box and a set of stacking rings (selected to be too difficult for most 1-year-olds
472 to manage on their own), and parents were asked to help the child as much as they thought
473 necessary with one toy at a time (2 × 3 min.; the assistant informed the dyad when it was
474 time to switch toys). They were also told that there was no need to complete the tasks (e.g.,
475 put all stacking rings in the correct order). Prior to the observations, parents were informed
476 that they could choose to discontinue the tasks at any time.

477 We used the mean of three global ratings - Sensitivity/responsiveness, Detachment/
478 disengagement (reversed), and Positive regard for the child – from the Norwegian

479 adaptation of the Qualitative Ratings for Parent Child Interaction at 3-15 Months in the
480 NICHD SECCYD study [93, 94] to form a composite index of parent sensitivity. The mean was
481 aggregated across two separately rated interaction segments, one being free-play and clean
482 up, and the other structured play. Each variable was rated on a Likert-type scale of 1 to 5 by
483 trained judges; interrater reliability (single rating, absolute agreement, two-way, random
484 intraclass correlation) ranged from .65 to .74 in the full set of all rating variables [94]. The
485 average was computed if at least three out of six ratings were present, which occurred in
486 717 fathers and 244 mothers. Internal consistency (alpha) for the composite was .88 for
487 fathers and .86 for mothers in 708 fathers and 242 mothers with complete sets of six
488 ratings.

489 **Harsh and positive parenting.** Those mothers and fathers who took part in the 1-
490 year assessment contributed self-reports of parenting practices on a subset of items of the
491 Parenting scale from the Fast Track Project adapted from the Parental Discipline Scale and
492 Parent Praise scales from the Pittsburgh Youth Study [95, 96]. For the same reasons as for
493 parent sensitivity, these self-reports were obtained for a larger proportion of fathers and a
494 smaller proportion of mothers. Items were answered on a frequency scale from 1 (*never*) to
495 5 (*almost always*). We used the mean of three items on parent positive behavior (giving the
496 child a smile, saying something nice about the child to the child, and giving the child a hug,
497 pat on the back, or a kiss) in response to desirable child behavior as the measure of positive
498 parenting practices. We used the mean of two parent negative items (yelling or screaming
499 and slapping or hitting) in response to undesirable child behavior as the measure of harsh
500 parenting practices. Alpha was .60 and .25, respectively, for positive and harsh parenting,
501 for 377 mothers, and alpha was .71 and .24, respectively, for positive and harsh parenting,
502 for 835 fathers. However, one issue with these reliability estimates is that Cronbach's alpha

– a correlational measure – provides an overly pessimistic view of the reliability of these measures considering attenuation due to ceiling and floor effects. We thus in addition conducted supplemental measurement models appropriate for ordered categorical variables. We fit separate confirmatory factor models to the respective items sets for harsh and positive parenting. Because of the floor and ceiling effects, we tested models that specified the items as ordered categorical indicators of their underlying latent constructs using in one case, a logit link function and robust, full-information, maximum likelihood (FIML) estimation and in the second case, a probit link function and WLSMV (weighted least squares mean and variance adjusted) estimation. The confirmatory factor analysis (CFA) model for the harsh parenting with only 2 indicators per latent construct was specified with fixed factor loadings of 1, making it a random intercept model. With 3 items per latent construct for positive parenting, we estimated 2 factor loadings and the factor variance for each latent construct. In the FIML models, the test for the missing completely at random (MCAR) assumption could not be rejected for both constructs, suggesting that the subsets of the sample, father only, mother only and mother and father were random samples from the same underlying population. The latent models indicated that the items functioned as intended, and that the latent factors were well defined: In the FIML models, the test of the underlying model for the ordered categorical indicators (a continuous latent indicator with thresholds) could not be rejected, suggesting that ordered categorical indicator model assumptions were reasonable. For the WLSMV models, the overall chi-square tests of model fit were non-significant (smallest p value was .21), RMSEAs were well under .05 (biggest RMSEA was .020), and CFI/TLI were well above .9, (smallest value was .948). In the fully standardized CFA results, all items loaded significantly (all $p < .001$) and moderately to strongly on their intended constructs, with FIML estimated single-indicator reliabilities for

527 the harsh parenting items for fathers and mothers of .43 and .41, respectively, and
528 estimated single-indicator reliabilities for the positive parenting items for fathers and
529 mothers ranging from .30 to .98, suggesting that equivalent reliabilities for latent construct
530 scores would be .60 for father's harsh parenting, .58 for mother's harsh parenting, .86 for
531 fathers' positive parenting, and .81 for mothers' positive parenting. Correlations of
532 estimated factor scores from latent models with the transformed raw scores we used in the
533 latent class growth models were very high at .95, suggesting that little would be gained by
534 using factor scores rather than the trimmed mean of the items that we had used.

535 Child predictor measures

536 **Child age.** The child's age was calculated as the time difference between the
537 interview date and the child's approximate (±5 days) birth date.

538 **Child gender.** The child's gender was a sampling variable recorded at study inclusion
539 and was coded 0 (*girl*) or 1 (*boy*).

540 **Presence of similar-aged sibling.** Presence of a similar-aged sibling (age difference
541 up to 5 years, thus for the greater part indicating the presence of older siblings) was coded
542 1, absence was coded 0.

543 **Child temperament.** To measure temperament, we included the scales Activity level
544 and Distress to limitations from Rothbart's Infant Behavior Questionnaire-Revised [97] as
545 well as the Soothability scale from the Infant Behavior Questionnaire [98], completed at the
546 6-month interview with a modified three-point response format: 1 (*most of the time*), 2
547 (*sometimes*), or 3 (*rarely or never*), in addition to a *does not apply* option. The mean score of
548 items was used as the scale score. Cronbach's alpha was .71, .76, and .77, respectively, for
549 the three scales.

550 Outcome variables

551 Teachers completed questionnaires in the middle of the second school year. With
552 few exceptions, children were in their 8th year around the time of data collection. We
553 selected the following variables:

554 **Externalizing**. We used the sum of the 12 Externalizing problem behavior items from
555 the Social Skills Improvement System-Rating Scales (SSIS-RS; [99]). Frequency is rated on a
556 four-point scale ranging from 0 (*never*) to 3 (*almost always*). We allowed one item to be
557 missing when calculating the sum; if so, the sum was adjusted as 12 times the item mean
558 and rounded up to the next integer. This sum could be calculated for 901 children (893
559 children had complete data whereas 8 had one missing item). No children had more than
560 one missing item. Cronbach's alpha based on 893 children with complete data was .92 for
561 the scale of 12 items.

562 **Internalizing**. We used the sum of the seven Internalizing behavior items from the
563 SSIS-RS [99]. Frequency is rated on a four-point scale ranging from 0 (*never*) to 3 (*almost*
564 *always*). We allowed one item to be missing when calculating the sum; if so, the sum was
565 adjusted as 7 times the item mean and rounded to the next integer. This sum could be
566 calculated for 901 children (892 children had complete data whereas 9 had one missing
567 item). No children had more than one missing item. Cronbach's alpha based on 892 children
568 with complete data was .83 for the scale of 7 items.

569 **Academic competence**. We used the sum of the ratings for the seven Academic
570 Competence items from the SSIS-RS [99], which are rated on a five-point scale ranging from
571 1 (*lowest 10%*) to 5 (*highest 10%*). We allowed one item to be missing when calculating the
572 sum; if so, the sum was adjusted as 7 times the item mean and rounded to the next integer.
573 This sum could be calculated for 896 children (889 children had complete data whereas 7

had one missing item). Data for one child with four items completed were excluded. Cronbach's alpha based on 889 children with complete data was .96 for the scale of 7 items.

Social skills. We used the sum of the frequency ratings for the 46 Social Skill items of the SSIS-RS [99]. Frequency is rated on a four-point scale ranging from 0 (*never*) to 3 (*almost always*). We allowed up to four items to be missing when calculating the sum; when items were missing, the sum was adjusted as 46 times the item mean and rounded to the next integer. This sum could be calculated for 887 children (839 children had complete data, 38 had one missing item, and 10 children had 2, 3, or 4 missing items). Data for fourteen children with 41 or fewer items completed was excluded. Cronbach's alpha based on 839 children with complete data was .96 for the scale of 46 items.

Statistical analyses

Software

Basic data operations were performed in SPSS, versions 20.0 through 28.0. Rasch analyses of child physical aggression were performed in Winsteps, version 3.90.0 through 4.7.1.0. Standard regression and latent class analyses were carried out using R, version 4.1 and Mplus, versions 8.6 and 8.7, respectively.

Analytic subsample and missing data

The present study used an analytic subsample of 1141 children with reports of physical aggression. Among predictor variables from 6 months representing demography and child temperament, 1.3% missing observations were imputed prior to the main analyses using a single imputation of missing values by means of the Estimation Maximization (EM) algorithm in SPSS) as described in more detail in Nærde et al. [9].

For the remaining early predictor variables and Grade 2 outcomes, missingness was addressed using multiple imputation in R [100] for standard regression analyses. The imputation model included all early predictors and grade 2 outcomes. For all analyses involving latent classes, missingness was handled within the latent-variable framework in Mplus as described below under Statistical Methods.

Attrition analyses revealed that the analytic subsample resembled the entire original study sample very closely with respect to all early predictors (Supporting information C). Analyses also revealed very great similarity between the entire original sample and the subsample of 901 children who had Grade 2 teacher reports with respect to early predictors; a multivariate binomial logistic regression on early predictors of presence of Grade 2 teacher reports, which made up our distal outcomes, suggested that maternal education, single parent, and sibling presence together significantly predicted missingness in Grade 2 teacher reports. These three early predictors were included in the fully adjusted models, including in the modeling of missingness, thus reducing or eliminating the possibility of bias in the fully adjusted models due to any missingness explained by early predictors.

Data transformations

Continuous predictors were centered at their sample means. To prevent potential distortions due to outliers, we trimmed non-binary predictors at about the 2nd and 98th percentiles of their respective distributions. Some of the non-binary predictors as well as the Grade 2 outcomes of externalizing and internalizing were highly skewed, as would be expected in this population, and were square root transformed to make the assumption of within-class normality more reasonable. Descriptive statistics for predictors and outcomes, before and after transformations, are shown in Table 2.

620 **Table 2. Descriptive statistics for predictors and outcomes**

		Raw data						Transformed data					
	N	***M***	***SD***	**Skew**	**Kurt**	**Min**	**Max**	***M***	***SD***	**Skew**	**Kurt**	**Min**	**Max**
							Early predictors						
Maternal education[a]	1141	14.29	2.57	-0.14	-1.06	9	18	0.00	2.57	-0.14	-1.06	-5.29	3.71
Paternal education[a]	1141	13.84	2.59	0.23	-1.09	9	18	0.01	2.59	0.23	-1.09	-4.84	4.17
Young mother	1141	0.14	0.34	2.12	2.49	0	1	--	--	--	--	--	--
Young father	1141	0.06	0.23	3.86	12.95	0	1	--	--	--	--	--	--
Financial stress	1096	0.12	0.33	2.34	3.46	0	1	--	--	--	--	--	--
Couple relationship[b]	1072	5.33	0.65	-1.70	4.18	1.20	6.00	0.70	0.39	-0.15	-0.70	0.00	1.38
Single parent	1141	0.05	0.21	4.32	16.66	0	1	--	--	--	--	--	--
Mat'l depression & anxiety[c]	1116	1.33	0.37	1.95	5.98	1.00	3.75	0.47	0.32	0.07	-0.82	0.00	1.15
Pat'l depression & anxiety[c]	671	1.27	0.33	2.59	13.12	1.00	4.00	0.40	0.31	0.10	-1.11	0.00	1.00
Maternal sensitivity[d]	244	3.90	0.61	-0.79	0.36	1.83	5.00	0.01	0.59	-0.63	-0.31	-1.43	0.93
Paternal sensitivity[d]	717	3.74	0.69	-0.45	-0.17	1.17	5.00	0.00	0.66	-0.32	-0.75	-1.41	1.09

		Raw data						Transformed data					
	N	***M***	***SD***	**Skew**	**Kurt**	**Min**	**Max**	***M***	***SD***	**Skew**	**Kurt**	**Min**	**Max**
Maternal harsh parenting[d]	377	1.42	0.50	1.96	7.73	1.00	5.00	-0.01	0.47	1.28	1.70	-0.42	1.58
Paternal harsh parenting[d]	835	1.40	0.47	1.23	1.77	1.00	4.00	-0.01	0.44	0.82	-0.28	-0.41	1.10
Mat'l positive parenting[d]	377	4.71	0.45	-2.82	13.90	1.00	5.00	0.01	0.40	-1.61	2.25	-1.38	0.29
Pat'l positive parenting[d]	835	4.53	0.54	-2.10	8.50	1.00	5.00	0.02	0.47	-0.86	-0.18	-1.19	0.47
Child gender	1141	0.52	0.50	-0.06	-2.00	0	1	--	--	--	--	--	--
Similar-aged sibling	1141	0.38	0.49	0.47	-1.78	0	1	--	--	--	--	--	--
Child activity level[d]	1141	2.07	0.28	0.07	0.00	1.09	2.83	0.00	0.27	0.11	-0.42	-0.54	0.60
Child distress to limitations[d]	1141	1.69	0.31	0.40	0.13	1.00	2.82	0.00	0.30	0.30	-0.31	-0.56	0.70
Child soothability[d]	1141	2.63	0.29	-0.46	-0.35	1.56	3.00	0.00	0.28	-0.30	-0.89	-0.63	0.37
							Grade 2 outcomes						
Externalizing[e]	901	4.81	5.26	1.68	3.22	0	35	1.83	1.21	0.34	-0.36	0.00	5.92
Internalizing[e]	901	3.01	2.97	1.17	1.49	0	18	1.42	1.00	-0.04	-0.94	0.00	4.24
Academic competence	896	25.03	7.12	-0.40	-0.50	7	35	--	--	--	--	--	--

		Raw data						Transformed data					
	N	***M***	***SD***	**Skew**	**Kurt**	**Min**	**Max**	***M***	***SD***	**Skew**	**Kurt**	**Min**	**Max**
Social skills	887	103.79	19.56	-0.43	-0.45	32	137	--	--	--	--	--	--

Note. Some variables were transformed as marked by letter subscripts. Descriptive statistics for transformed variables are displayed in the right part of the table. Reversing and recentering are linear transformations that do not change the shape of a distribution, the rank order of observations, or the linear relations with other variables. Trimming and square root transformations are non-linear transformations that change the shape of a distribution (make it less skewed and heavy-tailed), the rank order of a few of the observations (only trimming and this is because it produces ties among extremes) and thus potentially alter linear relations with other variables.

[a]Centered. [b]Trimmed (values smaller than the 2nd percentile or larger than the 98th percentile recoded back to the 2nd or 98th percentile values), reversed, recentered to have a minimum score of zero and square root transformed. [c]Trimmed, recentered to have a minimum score of zero and square root transformed. [d]Centered and trimmed. [e]Square root transformed.

Statistical methods

Growth model and latent class (trajectory) analysis (LCGA). Because of the large number and irregular timing of the repeated growth measures within children, we found it simplest to use a 2-level specification for the LCGA with the child's age in years, centered at 2.2, as the time variable. Prior research clearly indicates that the pattern of growth requires at least a 2nd degree polynomial (i.e., quadratic) model to capture the rise to a peak and subsequent decline. Preliminary analyses indicated that a 4th order model fit significantly and substantially better than 2nd or 3rd order models in our sample with up to 21 repeated measures, so we adopted the 4th order model for all analyses (i.e., included terms up to the 4th power of time). The basic 2-level LCGA-growth model specification included 5 child-level growth components with freely estimated means for the 5 growth components in each class, no within-class variability on growth components, and time-specific variability constrained to be equal across both time and classes. The specification of no within class variability in growth is what locates our model in the category of LCGA (latent class growth analysis; [16], which is considered the simplest longitudinal mixture model [101]). Most prominent trajectory studies to date, [including e.g., 8, 16, 21, 25] have employed LCGA models. LCGA models in common circumstances extract more classes than growth mixture models (GMM), which do allow for some variation of trajectories within classes [102].

3-Step approach. For models that included latent classes and either predictors, outcomes, or both, we used the manual 3-step procedure as recommended by Asparouhov and Muthen [84]. In addition to the advantages given by Asparouhov and Muthen, the 3-step approach allows us to use a much simpler way of dealing with missing data in child level predictors because Step 3 eliminates all the repeated growth measurements for a particular child and replaces them with one child level nominal predictor of latent class –

653 pseudo-class – which enables us to switch to a 1-level model. We had several important
654 predictors with substantial amounts of missing data. Mplus requires the use of a
655 computationally intensive approach based on Monte Carlo numerical integration to deal
656 with missing data in the predictors and this is much simpler in a 1-level model.

657 **Prediction of trajectory (class) membership.** To evaluate the overall effect of each
658 predictor on class membership, whether unadjusted and considered singly or fully adjusted
659 for all competing predictors on class membership, we used likelihood ratio tests (LRT)
660 comparing the model with all possible predictor effects to the model that included the
661 predictor but had all predictor effects constrained to zero. We used a pseudo-*R* statistic
662 [103] as the indicator of the strength of association between predictor and class
663 membership.

664 We investigated predictor effects on membership in any class versus another in
665 multinomial logistic regressions in the context of the 3-step approach. The multinomial
666 logistic model is a collection of individual logistic regressions, one less than the number of
667 classes, where one class serves as the reference class and predictors compete to
668 discriminate each of the other classes against the reference class in the individual,
669 simultaneous logistic regressions. The multinomial logistic regression model is estimated on
670 the log odds scale; a contrast coefficient of ±0.69 on the log odds scale corresponds to an
671 odds ratio of 2 or 0.5, thus expressing double the odds of membership in one of the
672 compared classes over the reference class conditional on the predictor. Contrast coefficients
673 express the change in log odds of the outcome for a 1-unit change in the predictor. For
674 interpretability, we standardized predictor coefficients by the predictors' sample standard
675 deviations.

676 In our case with 9 latent classes, the multinomial logistic regression produces 8
677 effects for each class vs the reference class for each predictor. The choice of the reference
678 class is arbitrary and Mplus uses the "last" class by default. The 8 effects can be recombined
679 in simple linear combinations to derive any contrast between any 2 classes, and it is very
680 helpful to examine these alternative contrasts to understand the nature of the effects of a
681 predictor on class membership. With 9 classes, there are 36 non-redundant, pairwise
682 contrasts.

683 **Prediction of Grade 2 outcomes from class (trajectory).** To evaluate the overall
684 effect of class membership on the distal outcomes considered singly, similarly to our
685 approach for comparing predictor effects, we adopted an LRT approach comparing the
686 model with distal outcome means freely estimated in each class to the null hypothesis
687 model with distal outcome means forced to be equal in each class. To summarize the overall
688 effect of latent class on outcomes we computed R as the square root of R^2 by computing the
689 proportional reduction of total outcome variance represented by the residual, within class
690 outcome variance, that is, R^2 = proportion explained variance = 1 - (residual variance/total
691 variance). For means comparisons of outcomes among latent classes, we used standardized
692 mean differences and examined all pairwise contrasts among classes.

693 **Prediction of Grade 2 outcomes from early predictors.** We used a standardized
694 regression weight as the indicator of the strength of association between each early
695 predictor and Grade 2 outcomes. These standard regression analyses were carried out in R
696 using multiple imputations to account for missing data.

697 **Unadjusted versus fully adjusted prediction models.** Comparing unadjusted vs
698 adjusted effects is useful because we gain insight into which predictors are the most
699 proximal and which are more distal. The former remains significant even after full

700 adjustment and the latter becomes nonsignificant, whether overall or just a few class
701 contrasts. We first examined all prediction effects unadjusted for other influences.
702 For each distal outcome, we went on to build fully adjusted models for the influence
703 of early predictors on trajectory, as well as the influence of trajectory and early predictors
704 on the outcome. For these analyses, we retained those early predictors that either had a
705 significant unadjusted effect on trajectory membership or had a significant effect on any of
706 the four distal outcomes after adjusting for all other predictors in separate multi-predictor
707 models. The full models included all possible predictor effects on latent class, all latent class
708 effects on the distal outcome, and all possible direct paths from the predictors to the distal
709 outcome.

710 Results

711 Number and shape of trajectories

712 We considered two- to ten trajectory solutions. We relied on information criteria,
713 the parametric bootstrap likelihood ratio test (BLRT), and substantive considerations to
714 select the number of trajectories that seemed optimal with respect to parsimony, model fit,
715 and the distinctions that the trajectory model made in the data. Table 3 shows information
716 criteria for the sequence of models. Test statistics for the BLRT are given in table SD2 in
717 Supporting information D. Nine trajectories minimized the Bayesian information criterion
718 (BIC), whereas ten trajectories (or possibly more) minimized the sample size adjusted BIC
719 and the Akaike information criterion (AIC). A less often reported criterion, the consistent
720 Akaike information criterion (CAIC; [104]) was minimized by the eight-trajectory solution.
721 The BLRT consistently rejected simpler models with less than 8 trajectories, although due to

the heavy computational burden we did not attempt to test 9 or more trajectories against simpler models.

724 **Table 3. Information criteria, entropy, and parameters for 2 to 10 trajectory models.**

Trajectories	-2log*L*	Parameters	BIC	adjBIC	AIC	CAIC	Entropy
2	51385	12	51499	51461	51408	51511	.83
3	50609	18	50781	50723	50645	50799	.82
4	50212	24	50441	50364	50260	50465	.80
5	50009	30	50294	50199	50069	50324	.77
6	49827	36	50170	50055	49899	50206	.77
7	49680	42	50080	49946	49764	50122	.79
8	49575	48	50032	49880	49671	**50080**	.76
9	49515	54	**50030**	49858	49623	50084	.76
10	49470	60	50041	**49850**	**49590**	50101	.77

725 *Note.* -2log*L*: -2 times the H0 loglikelihood. BIC: Bayesian information criterion. adjBIC: Sample-size adjusted Bayesian criterion. AIC: Akaike information
726 criterion. CAIC: Consistent Akaike information criterion [104]. Numbers in bold indicate minima or maxima that suggest the optimal solution according to
727 each information criterion.

728 The assembled statistical evidence clearly suggested to us that at least 7 trajectories
729 were necessary. Visual inspection of graphs of the two- to ten-trajectory solutions
730 (reproduced in Supporting information D) revealed, among other things, (a) that at least one
731 declining trajectory appeared in all solutions starting with the four-trajectory solution; (b)
732 that the small high-stable trajectory appeared from the seven-trajectory solution and
733 persisted virtually unchanged in shape and proportion through to the ten-trajectory
734 solution; (c) that starting with the eight-trajectory solution, there were two rather than one
735 no- or low-aggression trajectories; and (d) that in going from eight to nine trajectories two
736 declining trajectories appeared rather than one.
737 In considering 7, 8, or 9 trajectories, we eliminated the 7-trajectory model from
738 consideration because it was clearly and strongly rejected by the BLRT, the test that has the
739 best support mathematically [105, 106] and in simulations [104]. We selected the nine-
740 trajectory solution over the eight because it was suggested by the BIC, which is strongly
741 recommended by methodologists as being accurate, or at worst somewhat conservative, in
742 this context [104].
743 Figure 2 shows the 9-class model estimated trajectories, which seem to represent:
744 (1) one *high-stable* trajectory that rises to a very high peak of physical aggression
745 around 2 years, and stays at a high level (Trajectory 9, model-based trajectory
746 proportion 3%);
747 (2) three *variable-start, high-endpoint* trajectories that end up at an intermediately
748 elevated level of physical aggression, but with different early developments: one
749 starting high and declining after a rather high peak (Trajectory 8, 7%), one
750 starting low and quickly attaining the peak and decline levels of the

751 beforementioned (Trajectory 7, 9%), and one starting low with a late and flat
752 peak (Trajectory 6, 17%);
753 (3) two *high-start, low-endpoint* trajectories that decline from their initial highest
754 level, with a late and flat peak followed by another decline to a low level, one
755 following a slightly higher-level path (Trajectory 5, 5%) than the other (Trajectory
756 4, 7%);
757 (4) two *medium-peak, low-endpoint* trajectories that start and end up with a low
758 level of physical aggression, one reaching an early peak of a somewhat high level
759 (Trajectory 3, 16%), and one with a flatter peak at a lower level (Trajectory 2,
760 22%); and
761 (5) one *low-stable* or "*no aggression*" trajectory that starts out and stays at a low
762 level of aggression with just a minor rise to a very low, flat peak, and a decline
763 back to a low level (Trajectory 1, 14%).
764 Details of the 9-class model are given in Supporting information E.
765
766 **Fig 2. Model predicted child physical aggression by trajectory in nine-trajectory solution**
767 **(fourth-grade polynomial growth function)**
768
769 The curls toward the end of the developmental period for many of the trajectories
770 are most likely an artefact of fitting a 4th degree polynomial function to a data set with no
771 planned data collections between 4 and 5 years. The conclusion that levels of physical
772 aggression in Trajectories 4 and 5 rebound at or after 5 years is not well supported or
773 justified due to the lack of observed data between ages 4 and 5 years.

Table 4 shows model-estimated and pseudo-class assigned (i.e., by children's highest estimated trajectory membership probability) proportions of trajectories, and the distribution of boys and girls within trajectories. The estimated and assigned proportions were overall very similar, supporting the model's fit to the empirical distinctions present in the data set. Gender differences in development were clear in that more boys than girls were described by the high-endpoint and high-stable Trajectories 6-9 (from 60% to a maximum of 69% pseudo-class assignments for Trajectory 9), and more girls than boys were described by the no-aggression Trajectory 1 (28% boys by pseudo-class assignments). It is noteworthy that there were no trajectories that exclusively described only one gender; the low-endpoint Trajectories 2-5 described more or less even proportions of boys and girls, and each trajectory, including those with the most uneven gender distribution, described a fair proportion of the less represented gender (1 boy per 2.5 girls in Trajectory 1, and 1 girl per 2.2 boys in Trajectory 9).

787 **Table 4. Trajectory proportions and gender proportions within trajectory.**

	Model estimates		Pseudo-class assignments			
Trajectory	**Trajectory proportion**	**Proportion boys within trajectory[a]**	**Girls**	**Boys**	**Trajectory proportion**	**Proportion boys within trajectory**
1 No aggression	.137	.24	111	44	.136	.28
2 Medium-peak, low-endpoint – flatter	.219	.51	130	126	.224	.49
3 Medium-peak, low-endpoint – early peak	.160	.51	89	98	.164	.52
4 High-start, low-endpoint – lower	.069	.45	39	34	.064	.47
5 High-start, low-endpoint – higher	.054	.58	24	32	.049	.57
6 High-endpoint – late flat peak	.168	.62	80	119	.174	.60
7 High-endpoint – low start	.092	.66	35	64	.087	.65
8 High-endpoint – high start, high peak	.075	.59	35	52	.076	.60
9 High-stable	.027	.70	9	20	.025	.69

788 [a]The model estimated proportion of boys was derived from the set of unstandardized and unadjusted trajectory prediction contrast coefficients for child

789 gender as trajectory-membership predictor (Table 6 displays the corresponding standardized set) and the model estimated trajectory proportions (pseudo-

790 class proportions yielded the same result).

Unadjusted predictor effects on trajectory membership

Overall effects of each predictor in isolation on trajectory membership are shown in the first three data column of Table 5 including an overall, 8-degree-of-freedom chi-square test and the strength of the overall effect as pseudo-*R*.

795 **Table 5. Unadjusted overall prediction effects of early predictors on trajectory and outcomes, and of trajectory on outcomes.**

Predictor	Prediction to trajectory[a]			Prediction to Grade 2 outcomes[b]			
						Academic	
	χ^2	*p*	Pseudo R^a	Externalizing	Internalizing	competence	Social skills
Maternal education	17.1	.029	.12*	-.13***	-.12***	.22***	.18***
Paternal education	14.5	.070	.11	-.12***	-.09**	.17***	.15***
Young mother	13.7	.090	.11	.09**	.11**	-.01	-.09*
Young father	5.9	.660	.07	.09*	.13***	-.05	-.08*
Financial stress	12.2	.144	.10	.16***	.15***	-.13***	-.18***
Couple relationship (reversed)	23.3	.003	.14**	.09*	.12***	-.08*	-.06
Single parent	13.9	.086	.11	.16***	.13***	-.14***	-.15***
Maternal depression & anxiety	40.8	<.001	.19***	.07*	.09**	-.09**	-.06
Paternal depression & anxiety	16.1	.041	.12*	.13**	.08	.00	-.17***
Maternal sensitivity	4.4	.818	.06	-.12	-.13*	.11	.09
Paternal sensitivity	11.4	.179	.10	-.12**	-.07	.10*	.11**
Maternal harsh parenting	46.5	<.001	.20***	.13*	.04	-.09	-.09

Predictor	Prediction to trajectory[a]			Prediction to Grade 2 outcomes[b]			
						Academic	
	χ^2	*p*	Pseudo *R*[a]	Externalizing	Internalizing	competence	Social skills
Paternal harsh parenting	31.7	<.001	.17***	.06	.01	-.11**	-.09*
Maternal positive parenting	6.5	.594	.08	.02	-.02	-.06	-.03
Paternal positive parenting	7.5	.485	.08	.03	.04	.04	-.02
Child gender	54.7	<.001	.22***	.30***	.13***	-.13***	-.30***
Similar-aged sibling	66.6	<.001	.24***	-.02	-.07*	-.02	.04
Child activity level	32.3	<.001	.17***	.09**	.09**	-.03	-.08*
Child distress to limitations	38.0	<.001	.18***	-.04	.00	.02	.04
Child soothability	29.5	<.001	.16***	.09**	.03	-.07*	-.11**
Trajectory[c]	--	--	--	.27***	.11	.19***	.19**

796 *p<.05 **p<.01 ***p<.001.

797 [a]Overall prediction to trajectory was assessed by comparing a model with all class discrimination effects of each single predictor freely estimated to a

798 reduced model where all the class discrimination effects are zeroed out in an 8 degrees of freedom chi-square test; the effect size is given as pseudo-R

799 estimated as the square root of Nagelkerke's pseudo R^2 derived from robust chi-square, which was identical to the effect measure *w* derived from robust

chi square. [b]Predictor-to-outcome associations are given as fully standardized regression coefficients. Coefficients were standardized by sample means (Table 2). [c]Prediction of outcomes from trajectory membership is given as *R,* the square root of R^2 computed as the proportional reduction in total outcome variance represented by the residual within class variance, that is, R^2 = proportion explained variance = 1 - (residual variance/total variance).

Eleven out of the 20 predictors had significant overall effects on trajectory membership with pseudo-*R* ranging from .12 to .24. Similar-aged sibling and child gender were the two strongest predictors, and also significant at this level were maternal education, couple relationship, both parents' depression and anxiety, both parents' harsh parenting, and all three child-temperament variables. Trajectory predictor coefficients (with Trajectory 1 [no aggression], as the reference trajectory) with all possible pairwise contrast tests are shown in the upper part of Table 6. Pairwise contrasts were not corrected for mass significance; at $p<.05$, 1.8 pairwise contrasts (i.e., 3.6 superscripts as each pair of trajectories are superscripted for a contrast) would be expected to be significant by chance in the case of a null association between a predictor and the 9-trajectory solution. In the presence of a significant overall effect, the pairwise contrast tests are offered as a guide to interpret the predictor-trajectory effect direction. In the absence of a significant overall effect, significant uncorrected pairwise contrasts do not suggest an effect unless clearly exceeding the chance expected number (i.e., at least 3 significant contrasts) and suggesting a difference involving one – or possibly a few – categories in isolation against other classes, which is an effect pattern that may be present even in the case of an insignificant overall test. This ability to detect effects that are limited to only a few groups but still may be large enough to be important represents a strength of the latent class approach.

821 **Table 6. Trajectory predictor coefficient contrasts for unadjusted effects of early predictors on trajectory and means comparison contrasts for**
822 **unadjusted effects of trajectory on outcomes.**

Variable	Overall effect[a]	Trajectory 1	2	3	4	5	6	7	8	9
		Trajectory prediction coefficients with pairwise contrasts for effects of early predictors on trajectory[b]								
Maternal education	.12*	0.00	0.15[489]	0.00	-0.27[267]	0.12	0.13[489]	0.21[489]	-0.24[267]	-0.40[267]
Paternal education	.11	0.00[9]	-0.04	-0.10	-0.17	-0.02	-0.03[9]	0.21[89]	-0.31[7]	-0.60[167]
Young mother	.11	0.00	-0.02	0.03	0.23[67]	-0.04	-0.29[48]	-0.30[4]	0.12[6]	-0.03
Young father	.07	0.00	-0.03	-0.07	-0.08	-0.07	-0.23[9]	-0.17	-0.06	0.18[6]
Financial stress	.10	0.00[9]	0.04[9]	0.20	0.29	0.23	0.09[9]	0.26	0.29	0.46[126]
Couple relationship (reversed)	.14**	0.00[6]	-0.13[35689]	0.21[24]	-0.22[3689]	0.23[2]	0.28[124]	0.17	0.24[24]	0.45[24]
Single parent	.11	0.00	-0.50	0.02	0.06	-0.12	-0.05	0.12	0.15	0.06
Maternal depression & anxiety	.19***	0.00[36789]	0.09[3789]	0.45[12]	0.30[9]	0.33	0.32[19]	0.57[12]	0.63[12]	0.84[1246]
Paternal depression & anxiety	.12*	0.00[8]	-0.35[48]	0.05	0.29[2]	0.04	0.15	0.10	0.47[12]	0.18
Maternal sensitivity	.06	0.00	0.09	-0.11[5]	-0.09[5]	0.68[34678]	-0.07[5]	-0.10[5]	-0.08[5]	0.22
Paternal sensitivity	.10	0.00	-0.03	-0.13[4]	0.57[3569]	-0.22[4]	-0.24[4]	0.08	0.03	-0.22[4]

Variable	Overall effect[a]	Trajectory 1	2	3	4	5	6	7	8	9
Maternal harsh parenting	.20***	0.00^{38}	$-0.64^{3456789}$	0.88^{12}	0.68^{2}	0.72^{2}	0.65^{28}	0.64^{2}	1.15^{1269}	0.47^{28}
Paternal harsh parenting	.17***	0.00^{3679}	0.23^{79}	0.51^{19}	0.45^{9}	0.48^{9}	0.54^{19}	0.75^{12}	0.47^{9}	$1.00^{1234568}$
Maternal positive parenting	.08	0.00	-0.19	-0.43	-0.10	-0.46	-0.13	-0.38	-0.39	0.00
Paternal positive parenting	.08	0.00^{28}	-0.43^{1}	-0.29	-0.35	-0.25	-0.22	-0.21	-0.43^{1}	-0.46
Child gender	.22***	$0.00^{23456789}$	0.59^{1}	0.59^{1}	0.47^{17}	0.73^{1}	0.80^{1}	0.90^{14}	0.76^{1}	0.98^{1}
Similar-aged sibling	.24***	0.00^{356789}	0.33^{789}	0.37^{1789}	0.07^{6789}	0.55^{1}	0.53^{1479}	0.90^{12346}	0.80^{1234}	1.00^{12346}
Child activity level	.17***	0.00^{4589}	-0.04^{4589}	0.06^{48}	0.45^{1236}	0.35^{126}	-0.09^{4589}	0.18	0.50^{1236}	0.37^{126}
Child distress to limitations	.18***	$0.00^{3456789}$	0.32^{58}	0.56^{1}	0.63^{1}	0.75^{12}	0.53^{1}	0.54^{1}	0.68^{12}	0.67^{1}
Child soothability	.16***	0.00^{23568}	-0.34^{18}	-0.39^{18}	-0.17^{8}	-0.58^{1}	-0.45^{18}	-0.20^{8}	$-0.76^{1234679}$	-0.29^{8}
		Standardized means comparisons with pairwise contrasts for effects of trajectory on distal outcomes[c]								
Externalizing	.27***	$0.00^{23456789}$	0.32^{167}	0.37^{17}	0.47^{17}	0.48^{17}	0.71^{12}	0.99^{12345}	0.63^{1}	0.75^{1}
Internalizing	.11	0.00	0.06	0.19	0.25	-0.03	0.16	0.17	0.31	0.46
Academic competence	.19***	0.00^{9}	-0.19^{9}	-0.24^{9}	-0.11^{9}	0.08^{79}	-0.19^{9}	-0.34^{59}	-0.38^{9}	$-1.09^{12345678}$

Variable	Overall effect[a]	Trajectory 1	2	3	4	5	6	7	8	9
Social skills	.19**	0.00[679]	-0.06[679]	-0.15[9]	-0.43	-0.35	-0.42[12]	-0.45[12]	-0.33	-0.80[123]

*p<.05 **p<.01 ***p<.001 for overall effects.

[a]Pseudo-R/R estimate with overall significance test from Table 5. [b]Trajectory prediction contrast coefficients are fully standardized logistic regression effects on the log odds scale (the effect on the log odds of the class contrast for a change in the predictor of one standard deviation) given relative to membership in Trajectory 1 (the "no aggression" trajectory). Coefficients were standardized by sample means (Table 2). Contrast coefficients between any pair of trajectories can be derived by subtraction (i.e., for maternal education between Trajectories 2 and 4: 0.15- [-0.27]=0.42, which is significant at p<.05 as indicated by superscripts). A contrast coefficient of 0.69 is equivalent to an odds ratio of 2.0, and a coefficient of -0.69 is equivalent to an odds ratio of 0.5. [c]The entries are standardized mean differences relative to Trajectory 1. Contrast coefficients between any pair of trajectories can be derived by subtraction.

[123456789]Number superscripts indicate pairwise contrast tests significant at p<.05 between the indexed trajectory and the trajectory with the superscripts' number; significant contrasts are superscripted for both involved trajectories. Contrast tests are not adjusted for mass significance; the chance expected number out of 36 pairwise contrasts at p<.05 would be 1.8 per variable, or 3.6 numbered subscripts per table row, in the case of a null association between a predictor/outcome and the 9-trajectory solution. In the presence of a significant overall effect, contrast tests are offered as a guide to effect direction interpretation; in the absence of a significant overall effect, significant contrast tests should not be taken as implying an effect unless clearly exceeding the chance expected number (i.e., at least 3 significant contrasts) and suggesting a difference involving one or possibly a few isolated trajectories.

The contrast tests for predictors with a significant overall effect on trajectory membership generally indicated effects in the expected direction, with risk factors predicting trajectories with higher overall level, peak, or endpoints, and protective factors predicting in the opposite direction. Out of the nine predictors without a significant overall effect, there were four for which the number of significant pairwise contrast tests clearly exceeded the chance expected number and suggested a difference involving one or a few categories in isolation; these were *paternal education* (as education went up, a child was less likely to be in Trajectory 9 vs. Trajectories 1, 6, or 7, and more likely to be in Trajectory 7 vs. Trajectory 8), *financial stress* (as stress went up, a child was more likely to be in Trajectory 9 compared to Trajectories 1, 2, and 6), *maternal sensitivity* (as sensitivity went up, a child was more likely to be in Trajectory 5 compared to five other trajectories) and *paternal sensitivity* (as sensitivity went up, a child was more likely to be in Trajectory 4 compared to four other trajectories). (For young mothers, the number of significant contrasts exceeded chance expectations, but suggested no clear effect pattern.)

Unadjusted effects of trajectory membership on grade 2 outcomes

Overall effects of physical-aggression trajectory on outcomes in isolation tested with a likelihood ratio chi-square with 8 degrees of freedom are shown in the lowermost row of Table 5. Trajectory significantly predicted externalizing (R^2 = 7.4%), academic competence (R^2 = 3.6%), and social skills (R^2 = 3.7%). Uncorrected for mass significance, pairwise means comparisons among outcomes by trajectory with contrast tests are shown in the lower part of Table 6, and should be interpreted analogous to the predictor-trajectory contrasts tests, That is, in the presence of a significant overall effect as a guide to effect direction interpretation, and in the absence of a significant overall effect suggesting no effect unless

clearly exceeding the chance expected number (i.e., at least 3) and revealing a difference involving one or possibly a few isolated trajectories. Regarding externalizing, the no-aggression trajectory (1) had the lowest mean, contrasting to all other trajectories, of which Trajectory 7 had the highest mean level. For academic competence, however, it was the high-stable trajectory (9) that contrasted to all other trajectories.

Unadjusted predictor effects on grade 2 outcomes

When considering the effects of predictors in isolation on any of the four outcomes, 17 predictors—all except maternal and paternal positive parenting and child distress to limitations—influenced one or more of the outcomes. The effects of the significant predictors on each outcome are shown in the rightmost four data columns of Table 5. The significant predictor-outcome associations were in the expected direction, with risk factors predicting poorer Grade 2 outcomes and protective factors predicting in the opposite direction, with the following exceptions: Soothability predicted slightly to (a) more externalizing, (b) less academic competence, and (c) lower social skills; having a similar-aged sibling predicted to less internalizing. The consistent prediction to all four Grade 2 outcomes from the parent and family variables of parental educational level, financial stress, single parent, and maternal depression and anxiety is noteworthy, as well as the finding that child gender (but not the other child variables) predicted strongly and consistently to each outcome.

Fully adjusted models predicting each distal outcome from early predictors and trajectory

For each distal outcome, we went on to build fully adjusted models for the influence of early predictors on trajectory, as well as the influence of trajectory and early predictors

882 on the outcome. For these analyses, we retained those 14 early predictors that either had a
883 significant unadjusted effect on trajectory membership (Table 5) or had a significant effect
884 on any of the four distal outcomes after adjusting for all other predictors in separate multi-
885 predictor models. Paternal education, young father, maternal and paternal sensitivity, and
886 maternal and paternal positive parenting did not have a significant overall effect on
887 trajectory (unadjusted) or any of the Grade 2 outcomes (adjusted for other predictors) and
888 were not included in the fully adjusted models. We did not allow significant predictor effects
889 to be in the opposite direction of effects in the unadjusted results. This occurred for the
890 effect of paternal depression and anxiety on academic competence (unadjusted
891 standardized regression coefficient -0.002 [N.S.]; freely estimated fully adjusted
892 standardized regression coefficient 0.139 [*p*=.018]); the final fully adjusted model for this
893 outcome was thus re-estimated with the effect of this predictor fixed to zero. The findings
894 are shown in Table 7.

895 **Table 7. Fully adjusted overall prediction effects of early predictors on trajectory and outcomes, and of trajectory on outcomes.**

Predictor	Prediction to trajectory[a]			Prediction to Grade 2 outcomes[b]			
						Academic	
	χ^2	*p*	Pseudo *R*[a]	Externalizing	Internalizing	competence	Social skills
Maternal education	8.2	.417	.08	-.07	-.04	.18***	.10**
Young mother	5.2	.732	.07	.04	.04	.08	-.01
Financial stress	5.5	.704	.07	.09*	.10*	-.07	-.09*
Couple relationship (reversed)	7.8	.453	.08	.03	.08*	-.02	.00
Single parent	12.4	.135	.10	.09*	.07	-.09*	-.07
Maternal depression & anxiety	10.9	.205	.10	.00	.02	-.02	.00
Paternal depression & anxiety	4.2	.834	.06	.03	.00	.00[c]	-.09
Maternal harsh parenting	27.8	.001	.16***	.06	.00	-.10	-.05
Paternal harsh parenting	20.1	.010	.13**	.02	-.01	-.09	-.05
Child gender	49.2	<.001	.21***	.25***	.11**	-.10**	-.27***
Similar-aged sibling	90.3	<.001	.28***	-.06	-.06	.02	.06
Child activity level	15.2	.055	.12	.05	.06	.03	-.02

Predictor	Prediction to trajectory[a]			Prediction to Grade 2 outcomes[b]			
						Academic	
	χ^2	*p*	Pseudo *R*[a]	Externalizing	Internalizing	competence	Social skills
Child distress to limitations	10.7	.222	.10	-.10*	-.05	.01	.08*
Child soothability	11.5	.174	.10	.10*	.04	-.08*	-.10**
Trajectory[d]	--	--	--	.22***	.09	.16	.12

896 *p<.05 **p<.01 ***p<.001.

897 [a]Comparisons are of a baseline model where all class discrimination effects of all early predictors are included to a reduced model where all the class
898 discrimination effects for a single early predictor are zeroed out completely; prediction to trajectory membership is given as pseudo–*R* as the effect
899 measure *w* derived from robust chi square. [b]Predictor-to-outcome associations are given as fully standardized regression coefficients. [c]The effect of
900 paternal depression and anxiety changed signs from the unadjusted results and was fixed at 0. [d]Prediction of outcomes from trajectory membership is given
901 as *R,* the square root of R^2 computed as the proportional reduction in total outcome variance represented by the residual within class variance, that is, R^2 =
902 proportion explained variance = 1 - (residual variance/total variance).

903 In these adjusted models, the number and magnitude of significant effects, as
904 expected, generally shrank relative to the unadjusted effects (Table 5), highlighting that
905 these adjusted models identify the most proximal predictors of the trajectories and
906 outcomes. Important but more distal influences that primarily work through the most
907 proximal predictors will generally not be significant, except possibly in very large samples.
908 Significant adjusted overall predictors of trajectory included similar-aged sibling
909 presence (the effect of which was stronger in the fully adjusted model than in unadjusted
910 models, but which did not have a direct effect on any Grade 2 outcome), child gender
911 (which also had direct effects on all Grade 2 outcomes), and maternal and paternal harsh
912 parenting (with no direct effects on Grade 2 outcomes). The upper part of Table 8 displays
913 fully adjusted trajectory prediction contrast coefficients. For the predictors with significant
914 effects, the contrast patterns resembled those in the unadjusted results, although fewer
915 contrasts reached significance. Out of the nine predictors without a significant overall effect,
916 there were three for which the number of significant pairwise contrast tests clearly
917 exceeded the chance expected number and suggested a difference involving one or a few
918 categories in isolation after adjustment for competing predictors; these included maternal
919 perceptions of poor relationship quality, child distress to limitations, and child soothability.
920 As maternal perceptions of poor relationship quality went up, a child was significantly more
921 likely to be in Trajectories 1, 3, or 6 compared to Trajectory 4. As distress to limitations went
922 up, a child was significantly more likely to be in Trajectories 3, 5, or 6 compared to
923 Trajectory 1. As soothability went up, a child was significantly more likely to be in
924 Trajectories 1, 3, and 4 compared to Trajectory 8.

925 **Table 8. Trajectory predictor coefficient contrasts for fully adjusted effects of early predictors on trajectory and means comparison contrasts for fully**
926 **adjusted effects of trajectory on outcomes.**

Variable	Overall effect[a]	Trajectory 1	2	3	4	5	6	7	8	9
		Trajectory prediction coefficients with pairwise contrasts for effects of early predictors on trajectory[b]								
Maternal education	.08	0.00	0.09	0.10	-0.14	0.18	0.12	0.36	-0.13	-0.24
Young mother	.07	0.00	-0.05	0.16	0.18	0.08	-0.17	-0.14	0.18	-0.10
Financial stress	.07	0.00	-0.07	0.12	0.11	0.17	0.07	0.25	0.10	0.27
Couple relationship (reversed)	.08	0.00[4]	-0.21	0.06[4]	-0.40[136]	0.05	0.17[4]	-0.03	-0.13	0.07
Single parent	.10	0.00	-0.29[7]	0.24	0.14	0.07	0.15	0.34[2]	0.27	0.15
Maternal depression & anxiety	.10	0.00[79]	0.27	0.34	0.24	0.15	0.19	0.51[1]	0.41	0.71[1]
Paternal depression & anxiety	.06	0.00	-0.10	-0.26	0.06	-0.13	-0.14	-0.17	0.03	-0.03
Maternal harsh parenting	.16***	0.00[38]	-0.66[38]	1.18[129]	0.62	0.66	0.72	0.57	1.16[129]	0.34[38]
Paternal harsh parenting	.13**	0.00[3679]	0.20[7]	0.68[1]	0.51	0.55	0.62[1]	0.76[12]	0.53	0.80[1]
Child gender	.21***	0.00[2356789]	0.64[1]	0.53[1]	0.39[79]	0.70[1]	0.80[1]	1.04[14]	0.69[1]	1.04[14]
Similar-aged sibling	.28***	0.00[356789]	0.36[789]	0.38[1789]	0.25[789]	0.59[1]	0.49[1789]	1.03[12346]	0.97[12346]	1.03[12346]
Child activity level	.12	0.00	0.03	-0.26[8]	0.12	0.12	-0.33[8]	0.00	0.20[36]	0.10

Variable	Overall effect[a]	Trajectory 1	2	3	4	5	6	7	8	9
Child distress to limitations	.10	0.00^{356}	0.17	0.48^{1}	0.54	0.53^{1}	0.45^{1}	0.42	0.32	0.46
Child soothability	.10	0.00^{8}	-0.28	-0.05^{8}	0.02^{8}	-0.28	-0.21	0.03	-0.49^{134}	-0.07
		Standardized means comparisons with pairwise contrasts for effects of trajectory on distal outcomes[c]								
Externalizing	.22***	$0.00^{23456789}$	0.33^{17}	0.35^{17}	0.40^{1}	0.47^{1}	0.62^{1}	0.88^{123}	0.52^{1}	0.53^{1}
Internalizing	.09	0.00	0.17	0.19	0.23	-0.02	0.14	0.16	0.29	0.35
Academic competence	.16	0.00^{9}	-0.31	-0.20^{9}	-0.02^{9}	0.07^{9}	-0.09^{9}	-0.22	-0.27	-0.84^{13456}
Social skills	.12	0.00	-0.12	-0.13	-0.36	-0.32	-0.31	-0.27	-0.22	-0.55

927 *p<.05 **p<.01 ***p<.001 for overall effects.

928 [a]Pseudo-R/R estimate with overall significance test from Table 7. [b]Trajectory prediction contrast coefficients are fully standardized logistic regression effects

929 on the log odds scale (the effect on the log odds of the class contrast for a change in the predictor of one standard deviation) given relative to membership

930 in Trajectory 1 (the "no aggression" trajectory). Coefficients were standardized by sample means (Table 2). Contrast coefficients between any pair of

931 trajectories can be derived by subtraction (i.e., for couple relationship [reversed] between Trajectories 6 and 4: 0.17- [-0.40]=0.57, which is significant at

932 p<.05 as indicated by superscripts). A contrast coefficient of 0.69 is equivalent to an odds ratio of 2.0, and a coefficient of -0.69 is equivalent to an odds

ratio of 0.5. [c]The entries are standardized mean differences relative to Trajectory 1. Contrast coefficients between any pair of trajectories can be derived by subtraction.

[123456789] Number superscripts indicate pairwise contrast tests significant at $p<.05$ between the indexed trajectory and the trajectory with the superscripts' number; significant contrasts are superscripted for both involved trajectories. Contrast tests are not adjusted for mass significance; the chance expected number out of 36 pairwise contrasts at $p<.05$ would be 1.8 per variable, or 3.6 numbered subscripts per table row, in the case of a null association between a predictor/outcome and the 9-trajectory solution. In the presence of a significant overall effect, contrast tests are offered as a guide to effect direction interpretation; in the absence of a significant overall effect, significant contrast tests should not be taken as implying an effect unless clearly exceeding the chance expected number (i.e., at least 3 significant contrasts) and suggesting a difference involving one or possibly a few isolated trajectories.

We can thus infer that the set of predictors that were significantly related to trajectory membership (either overall or for some trajectory contrasts) prior to adjustment, but not after, potentially have important *indirect* or *mediated effects* on trajectory membership. These predictors include paternal depression and anxiety, maternal education, financial stress, and young mother status. It is beyond this work's scope, however, to identify and quantify specific mediational pathways.

In the adjusted models, trajectory had a significant overall effect on externalizing (R^2 = 5.0%), but not on academic competence (R^2 = 2.6%), social skills (R^2 = 1.5%) or internalizing (R^2 = 0.8%). Adjustment for competing predictors eliminated the significant unadjusted overall effect of trajectory on academic competence and social skills. Out of the outcomes with no significant overall effect, the number of pairwise contrast tests exceeded the chance expected number for academic competence, suggesting a difference involving specifically Trajectory 9 (high-stable), which had the lowest mean, contrasting with Trajectories 1, 3, 4, 5, and 6. Thus, after adjustment for early predictors, trajectory effects remained significant and important for externalizing (overall) and for academic competence (high-stable in isolation contrasting with five other trajectories), were eliminated for social skills, and remained unrelated to internalizing.

There were direct effects on one or more Grade 2 outcomes of several early predictors including child gender (effects on all 4 outcomes), financial stress (externalizing, internalizing, and social skills), maternal education (academic competence and social skills), single parent status (externalizing and academic competence), couple relationship (internalizing), child distress to limitations (three outcomes, all in unexpected direction), and soothability (two outcomes, both in unexpected direction).

964 # Discussion

965 The current study extends previous research about the variation in normative
966 development of physical aggression from infancy to preschool age, how this development
967 matters for social, behavioral, and academic functioning in early school age, and the relative
968 importance of early characteristics related to the child, the parents, and the family on the
969 course and outcomes of aggression development. This is one of few studies extensively
970 testing these longitudinal trajectories in comprehensive models unraveling the relative
971 importance of children's developmental patterns of physical aggression for their subsequent
972 behavior and functioning, and the only one to simultaneously considering the direct and
973 indirect impact of a range of early risk and protective factors.

974 We found that nine latent trajectories represented the development of child physical
975 aggression from 1 to 5 years in a Norwegian population-based sample. The developmental
976 patterns could be characterized as *high-stable*, *variable-start and high endpoint*, *high-start*
977 *and low-endpoint*, *medium-peak and low-endpoint*, and *low stable* (*no aggression*),
978 respectively (Figure 2). Before adjustment for competing early risk and protective factors,
979 these trajectories had significant and moderate overall effects on externalizing, social skills,
980 and academic competence in Grade 2, but not on internalizing (R^2 of 7.4%, 3.7%, 3.6% and
981 1.2%, respectively). Despite the moderate overall effect size, significant individual mean
982 contrasts on outcomes among trajectory groups ranged from medium to very large effects.

983 The presence of a similar-aged sibling and child gender showed the strongest
984 unadjusted association with trajectory membership. Whereas each trajectory described
985 both boys and girls, gender differences manifested themselves in a larger proportion of girls
986 in the no-aggression trajectory, and a larger proportion of boys in the high-stable and high-

987 endpoint trajectories. In this respect, our findings are in accordance with earlier trajectory-
988 or group-based approaches [11] in suggesting that gender differences may result from more
989 boys than girls making up the small group of children who continues to deploy aggression at
990 high rates. The strongest and most consistent unadjusted predictors of Grade 2 outcomes
991 were child gender, maternal and paternal educational level, single parent status, and
992 financial stress.

993 In the fully adjusted model, which represents the most rigorous empirical test of
994 these influences to date, the effects of the aggression trajectories on the outcomes
995 simultaneously competed with those early predictors identified as having significant
996 unadjusted effects on trajectory membership or significant adjusted effects on any
997 outcome. While the significant direct effects were fewer and smaller, the ones remaining
998 represent the most proximal predictors of physical aggression development and Grade 2
999 functioning. Trajectory membership had significant and important effects on teacher rated
1000 externalizing problems as well as on academic competence when the children were in their
1001 8th year. Significant adjusted overall predictors of trajectory membership included the
1002 presence of a similar-aged sibling, child gender, and early maternal and paternal harsh
1003 parenting. Five other predictors (i.e., child temperament, maternal depression and anxiety,
1004 maternal perceptions of poor relationship quality, and single parent status) that fell short of
1005 an overall significant effect still had at least one significant class contrast. In addition to
1006 indirect effects through trajectory membership, child gender also had direct effects on all
1007 Grade 2 outcomes. Finally, paternal depression, maternal education, financial stress, and
1008 young motherhood all had potential indirect or mediated effects on class membership, as
1009 these predictors were significantly related to class membership *before* adjustment, but not

1010 *after*. In the following, we discuss the findings and the empirical contributions from our
1011 study in more detail, focusing mainly on the adjusted results.
1012 In contrast to our 9 trajectories, the most comparable extant latent trajectory studies
1013 have typically found three to five trajectories [2, 8, 10, 13, 17, 19, 21, 24, 25]. Compared to
1014 existing studies, our study started early encompassing the period of peak change (at about 2
1015 years), had frequent assessment (up to 21), used a vertically (age) scaled measurement
1016 instrument, and utilized 4th order polynomial growth curves, all qualities that would tend to
1017 increase significant distinctions among trajectories. The present study is indeed the only one
1018 getting close to the recommendation made by Tremblay [1] of assessments at least every
1019 two to three months from birth to 3 years, and ideally every 6 months from 3 years onward.
1020 Overall, these differences provided the power to make finer distinctions and detect more
1021 fine-grained developmental patterns of physical aggression.
1022 Even with the dissimilarities between our sample and findings and those from
1023 previous studies, there are also important similarities. Like other research [15, 17, 21, 24,
1024 25, 27], we identified one high-stable trajectory group (Trajectory 9) with the lowest
1025 prevalence (3%) and poor outcomes (the most severe on academic competence and social
1026 skills; but not on externalizing, although the difference between Trajectory 9 and the one
1027 with the poorest externalizing outcome was not significant). The proportion of children
1028 belonging to the typically identified high-stable trajectory, however, tends to vary across
1029 studies. For instance, in accord with our findings, the SECCYD study [21] found a high-stable
1030 trajectory group from 2 to 9 years constituting 3% of their sample (N=1195), having the
1031 poorest academic and social functioning through age 12 [13]. Also, Wildeboer et al. [17]
1032 found a high-increasing trajectory across 1.5, 3, and 6 years consisting of 3.2% of their
1033 sample (N=4781). Moreover, Joussemet et al. [25] reported about a high trajectory group

1034 from kindergarten to sixth grade consisting of 6% of their sample (N=1508), whereas
1035 Tremblay et al. [10] found that 13.9% of the sample (N=572), followed a rising trajectory of
1036 high physical aggression from 17 to 42 months. Finally, Côté et al. [8] found that as much as
1037 16.6% of their large sample (N=10658) followed a high stable trajectory between 2 and 11
1038 years.

1039 With respect to the lower end of the physical aggression spectrum, the SECCYD study
1040 [21], found a very low trajectory describing 45% of the study children, and another 25%
1041 described by a low group. In our sample, the *low-stable or no aggression* (1), *medium-peak*
1042 *and low endpoint* (2,3), and *high-start and low-endpoint* (4,5), all with low aggression
1043 endpoints at age 5, albeit with differing developmental patterns up to then, may
1044 quantitatively (in all 64%) correspond to a couple of larger very-low or low aggression
1045 trajectories like those identified in the SECCYD [21]. Such a comparison may be even more
1046 reasonable considering that the SECCYD, like most previous research, did not start until 2
1047 years, when physical aggression probably has begun to diminish among most children and
1048 particularly those engaging in low levels or frequency of aggression.

1049 Our set of nine trajectories revealed more diversity in predictor-trajectory patterns
1050 than a less fine-grained classification into fewer developmental patterns might have
1051 enabled. That is, the somewhat complex solution offered a perspective that was
1052 differentiated enough to view not only a few main developmental trends, but also to
1053 identify some more subtle distinctions in developmental patterns and their relation to
1054 predictors that supplemented and nuanced the main findings. Overall, all child
1055 characteristics (gender and three temperament dimensions) predicted trajectory
1056 membership in the unadjusted results, along with similar-aged siblings, maternal education,
1057 couple relationship, both parents' depression and anxiety, and both parents' harsh

1058 parenting practices, generally in the expected direction (Table 6). These effects were most
1059 clearly seen in significant predictor-coefficient contrasts between, on the one hand, the *low-*
1060 *stable or no aggression* trajectory (1) and to a lesser extent the *medium-peak and low-*
1061 *endpoint* trajectory (2), and on the other hand, the *high-stable* trajectory (9) and to a lesser
1062 extent the *variable-start and high-endpoint* trajectories (6, 7, 8).

1063 Exploration of the finer distinctions in predictor-trajectory effect patterns may
1064 provide some clues for what make children follow an intermediate-endpoint rather than a
1065 stable-high trajectory. While the intermediate- and high endpoint Trajectories 6-9 shared
1066 many predictors – notably, there were no significant trajectory prediction contrasts among
1067 them for child gender, child distress to limitation, couple relationship, paternal depression
1068 and anxiety, or paternal or maternal sensitivity – the predictor-trajectory contrasts seemed
1069 to suggest some possibly important effect-pattern differences. High maternal, but low
1070 paternal, harsh parenting (compared to Trajectory 9), low child soothability (compared to
1071 Trajectories 9, 6, and 7) and high child activity level (like Trajectory 9 but unlike Trajectory 6)
1072 was associated with a child being in Trajectory 8. This could suggest that Trajectory 8 was
1073 somewhat more driven by child temperament and heated mother-child interactions, which
1074 is in line with the established knowledge that children's temperamental characteristics are
1075 associated with the development of externalizing problems [72], and with the mentioned
1076 differential susceptibility model postulating that children with difficult temperament are
1077 more vulnerable to negative parenting behavior than those with easy temperament [47, 70,
1078 73-75]. Moreover, the suggestion of a possible compensatory effect of father's low harsh
1079 parenting in a context with high maternal harsh parenting underscores the importance of
1080 including fathers in developmental research. This is in accord with the family systems
1081 perspective [107, 108] highlighting the need to consider how the behavior of both parents

1082 jointly impact children's socioemotional development. Also, higher maternal and paternal
1083 education (compared to Trajectories 9 and 8), and higher child soothability (compared to
1084 Trajectory 8), was associated with membership in Trajectories 6 and 7. This may indicate
1085 that parent resourcefulness and child self-regulation can function as protective factors
1086 against the high-stable aggression pattern even in the presence of many other shared risk
1087 factors. The association between self-regulation and externalizing is well-documented in the
1088 literature [109-111], as is the link between socioeconomic status and parenting [112, 113].
1089 Again, the results indicate that both maternal and paternal level of education matters for
1090 development. In addition, Trajectory 6 membership was associated with low financial stress,
1091 sibling absence, low infant activity level, and low maternal depression and anxiety
1092 (compared to Trajectory 9 and for child activity also to Trajectory 8), suggesting several
1093 other possible contributing factors that might elevate physical aggression.
1094 As regards the trajectory-outcome effect patterns, results from the unadjusted
1095 models consistently suggested poorer outcomes on externalizing, social skills, and academic
1096 competence associated with the higher level and endpoint Trajectories 6-9. Still, there was a
1097 mean difference among these trajectories for academic competence only. In particular,
1098 Trajectory 9 had significantly worse outcomes than Trajectories 6-8 (except for
1099 externalizing, for which Trajectory 7 had the highest level), higher parental education
1100 predicting Trajectory 6 and 7 memberships compared to 9 and 8 possibly suggesting a
1101 contributing explanation.
1102 The prediction patterns for the two declining-aggression trajectories 4 and 5 may be
1103 particularly interesting to note. More difficult child temperament (i.e., higher activity level,
1104 more distress to limitations, and less soothability) compared to Trajectory 1, and one
1105 parent's high sensitivity (mother's for Trajectory 5 and father's for Trajectory 4,

1106 underscoring the importance of including both parents), predicted membership in these
1107 trajectories. This might suggest that parental sensitivity (and perhaps a good couple
1108 relationship, in the case of the lower declining Trajectory 4) moderates or mitigates
1109 physical-aggression development in children displaying difficult infant temperament, which
1110 is in accord with the differential susceptibility model [47, 70, 73-75]. In line with this, high
1111 infant activity level in combination with parental sensitivity to infant behavior may also
1112 hypothetically be elements explaining the high starting point of parent-reported child
1113 aggression in these trajectories.

1114 Of interest in an opposite manner might be pairs of trajectories for which the
1115 prediction pattern seemingly had less to offer in terms of possible explanations for
1116 differences in physical-aggression development. For instance, Trajectory 3 (low aggression,
1117 with an early rise and moderate peak) shared some predictors with trajectories with
1118 consistently higher aggression over time and was in fact indistinguishable from Trajectory 6
1119 on any predictor and contrasted with Trajectory 7 only with respect to sibling absence. It is
1120 tempting to hypothesize that groups of children whose development were described by
1121 Trajectories 3, 6, and possibly 7, did not differ dramatically on critical predictors at the
1122 outset, and that explanations outside the present set of predictors – perhaps involving time-
1123 varying or intervening influences – must be sought for what differentiated these
1124 developmental paths.

1125 Among predictors identified by previous research, neither being a single parent nor a
1126 young mother was significantly related to trajectory. This is probably due to the low
1127 proportion of single parents in our sample when the children were 1 year old, as well as the
1128 low proportion of young mothers (≤ 25 years), given the low prevalence of teenage mothers
1129 in Norway and the high mean age for primiparous women (i.e., 30.1 years in 2021 [114]).

1130 There were notably fewer statistically significant predictor-trajectory, trajectory-outcome,
1131 and predictor-outcome relationships in the *fully adjusted* set of results than in the
1132 unadjusted. While the differences between these sets of results may seem dramatic, it is
1133 crucial to acknowledge that the fully adjusted model, which is the benchmark test for
1134 identifying the most proximal predictors, constituted a rigorous test of all possible
1135 simultaneous effects. Importantly, net of all other effects, trajectory still predicted Grade 2
1136 externalizing problems and academic competence. Within this context, it may be worth
1137 noting that the high-stable trajectory specifically predicted to poor academic competence,
1138 while the no-aggression trajectory was the one that contrasted most starkly to all other
1139 groups with respect to externalizing problems.

1140 Our results both support and extend previous research in strongly suggesting that
1141 early maternal and paternal harsh parenting practices, in addition to child gender and
1142 similar-aged sibling presence, consistently predict physical-aggression development from 1
1143 to 5 years. This is the first study applying a trajectory approach addressing developmental
1144 patterns of physical aggression from infancy to preschool age to link early *paternal* harsh
1145 parenting, as well as that of *maternal*, with this development. Previous research [8, 10, 14,
1146 25] has not been able to explore this association due to lack of data on fathers parenting
1147 behavior. The results from the current study underscore the significance of fathers' harsh
1148 parenting behavior in infancy for children's subsequent development of physical aggression
1149 and is in line with a limited but growing literature highlighting the significance of fathering
1150 for children's and adolescents' development [58, 59, 64, 115].

1151 The importance of harsh parenting in predicting the development of children's
1152 physical aggression corresponds with findings from previous research that has included this
1153 process-oriented predictor albeit with mothers only [8, 10, 14, 25], and generally with the

1154 mentioned theories of development of aggression and externalizing problems, particularly
1155 Patterson's [53] coercion theory. In line with the established knowledge that parenting
1156 behaviors typically associated with the development of externalizing behaviors tend to co-
1157 appear with parental psychopathology [55-57], it is theoretically and empirically sound to
1158 view the pattern of predictor-trajectory influences in our unadjusted and adjusted results in
1159 a perspective where factors like parental depression and anxiety, couple relationship
1160 troubles, and difficult child temperament do not in *themselves* produce an effect on
1161 physical-aggression development, unless they are associated with an increase in harsh
1162 maternal and paternal parenting behavior. At the same time, we would like to add that this
1163 is in fact a bit of an over-simplification of all the predictor-trajectory results. As mentioned
1164 earlier, there were some significant contrast patterns in the fully adjusted models for
1165 predictors that did not have a significant overall adjusted effect (i.e., child distress to
1166 limitations and soothability and maternal perceptions of poor relationship quality). These
1167 effects tend to be notable in magnitude but limited in scope, that is, they affect only a few
1168 classes or small classes.

1169 All in all, the fully adjusted findings add importantly to previous findings. The few
1170 studies that have examined trajectory outcomes as well as predictors [14, 17, 21, 24]
1171 included a more limited number of social risk variables as covariates in their final set of
1172 analyses, and thus did not consider competing effects on trajectory and outcomes of the
1173 *complete* set of predictors, as was done in the present study.

1174 As regards *predictor-outcome effects* in both sets of findings, it is noteworthy that
1175 early financial stress had a consistent impact on three Grade 2 teacher-rated outcomes
1176 (externalizing, internalizing, and social skills) in the fully adjusted model. This is even more
1177 notable considering the crude measure we used (i.e., parents' dichotomous reports of

1178 having more long-standing financial difficulties in the past year [rent, mortgage, obligations,
1179 etc.]). Family financial stress when the child is about one year old may thus have long-lasting
1180 (or continuing) effects on child social, behavioral, and academic functioning. Importantly
1181 though, this effect may not relate specifically to the development of children's physical
1182 aggression (the predictor had no significant effect on *trajectory* in any set of findings). This
1183 result highlights the well-established finding linking socioeconomic conditions to a range of
1184 child behavioral outcomes [40, 42, 116, 117]. At the same time, some previous research has
1185 found low family income to act as a risk factor for more frequent and stable use of physical
1186 aggression across early childhood and adolescence [2, 8, 10]. We have no clear opinion as to
1187 why our results do not replicate earlier findings on this point, but possible explanations
1188 could be differences between countries/cultures with a smaller income range and a smaller
1189 proportion of families in serious financial hardship in Scandinavia than in North America and
1190 Canada, variations in age span, the use of financial stress as opposed to low family income
1191 as an indicator of financial risk, and the crude nature of our measure.
1192 Another note concerning predictor-outcome effects concerns the two child
1193 temperament variables of distress to limitations and soothability. While predicting
1194 aggression trajectories as expected in the unadjusted results, these also predicted three
1195 outcomes in the *unexpected direction* in the unadjusted and adjusted findings. We have no
1196 strong and plausible hypothesis concerning these findings. Possible explanations include
1197 interactional or transactional associations among child temperament, child behavior, and
1198 parenting [72]. For example, that difficult temperament in a population-based and for some
1199 part resourceful sample like the present might elicit child-parent behavior transactions
1200 bolstering or mitigating the temperamental risk, which could eventually lead to better

1201 outcomes for some children having difficult infant temperament, to the extent of producing
1202 a paradoxical overall mean effect.

1203 Finally, the weak associations between observed parental sensitivity and aggression
1204 trajectory, and the lack of association with Grade 2 outcomes, merit some attention, not
1205 least considering that self-reported harsh parenting practices in response to undesirable
1206 child behavior for mothers and fathers alike was one of the strongest predictors of
1207 trajectory membership. As previously noted by Nordahl [94], null findings in the current
1208 sample may partly relate to the fact that the parents overall displayed low levels of aversive
1209 or negative parenting behaviors with their 1-year-olds during the interaction session. Thus,
1210 limited variability in the observed parenting behavior may have contributed to the finding of
1211 weak associations. For one thing, fathers from a population based (versus clinical) sample
1212 may essentially engage in positive behavior during interaction with their infants, and
1213 particularly when being observed. Moreover, the observed interactions were relatively short
1214 in duration and the tasks were not designed to induce conflict or elicit negative or harsh
1215 parenting behaviors. This made it less likely that parents would experience the session as
1216 very challenging and/or that it would trigger much undesirable child behavior. Whereas
1217 results from the NICHD SECCYD [21] indicated that less observed sensitive and involved
1218 maternal parenting was associated with higher and more stable trajectories of physical
1219 aggression from toddlerhood to middle childhood, that study had access to more
1220 comprehensive observational data at several ages across infancy and third grade.

1221 This study has some notable strengths. We apply rigorous statistical tests using very
1222 recent advances in latent class methodology of the early predictors and distal outcomes of
1223 developmental trajectories of physical aggression, using a rich longitudinal data set. The
1224 latent trajectory approach, among other things, also allows us to detect important effects

that are limited to only a few groups. The sample is relatively large and resembles the general Norwegian population fairly well. The participation rate is furthermore high and the attrition for the most part low. We also include frequent reports on a relevant measure of child physical aggression from both mothers and fathers, covering the developmental periods when such behavior typically emerges, peaks, and decreases. Teachers, being independent informants from a different setting, report the Grade 2 outcomes. Moreover, our data set contains a comprehensive number of early predictors and all of those used in previous research, including some with reports from both parents. The only relevant predictor from other research that we were not able to include was parents' antisocial behavior in their own youth, which was not measured in our sample. We also test all possible predictor-outcome influences simultaneously. The strength of the fully adjusted set of findings is that it highlights effects that stand out as proximal, strong and consistent enough to withstand the competition with other possibly relevant associations. Particularly, in contrast to previous sets of findings in which all links between the included early predictors, the trajectories as well as the outcomes were not considered simultaneously, trajectory-outcome associations are not overstated.

The study is also limited in a few ways. Some of our predictors were quite imperfectly measured. In particular, the estimated reliability of self-reported harsh and positive parenting was less than desirable even when estimated by latent-variable models that were not attenuated by floor and ceiling effects. Also, due to study design, only subsets of fathers and mothers provided such self-reports at 1 year. Under such circumstances, null findings (e.g., no trajectory or outcome prediction from positive parenting practices) are uninformative, whereas significant results (e.g., significant prediction to trajectory and distal outcomes from harsh parenting) signal true effects that are large enough to be detected

1249 with the given sample, although attenuated and less precisely estimated. The findings from
1250 sensitivity analyses showing that latent-variable constructs predicted trajectory just in the
1251 same way as the manifest variables further supported the unlikeliness that low predictor
1252 reliability artefactually distorted these effects.

1253 As mentioned, there is evidence that individual differences in the early use of
1254 physical aggression is associated with both social and genetic factors [33]; however, our
1255 design did not permit the disentangling of genetic predispositions toward aggressive
1256 behavior in children and parents. Moreover, in line with most others, we use parents as
1257 informants about children's aggressive behavior. Our design and analytical strategy, with a
1258 range of predictors and outcomes and nine distinct developmental trajectories of physical
1259 aggression, furthermore, reduced the possibility for investigating interaction effects among
1260 predictors, or among predictors and trajectories in predicting outcomes. Possible interaction
1261 effects worth studying include moderating effects of gender on predictor-trajectory,
1262 trajectory-outcome, or predictor-outcome patterns, as well as interactions between
1263 parenting and child gender and sibling status. For example, it is easy to imagine that sibling
1264 squabbling could amplify harsh parenting. The choice to study a nine-trajectory model (in
1265 the normative absence of a known true number of groups) also has consequences for what
1266 results emerge as significant: A model with fewer trajectories, all other things being equal,
1267 would have greater statistical power to detect predictor-trajectory or trajectory-outcome
1268 associations. Of course, this holds only to the degree that the fewer trajectories still
1269 captured the critical differences in development that mattered for those relationships.
1270 Otherwise, fewer groups could obscure associations to the degree that the broader
1271 trajectories encompassed subgroups with differing predictor or outcome patterns.

Also related to analytical strategy, the merits of testing all influences in fully adjusted models are obvious, but this approach also brings with it the limitation that the results are dependent on the particular variable universe of the present study. The included predictors were all time-invariant in the analyses, measured only *prior to* or *at* the starting point for the trajectory model. This precluded us from investigating the importance of time-varying predictors or intermediate-timepoint predictors such as parental harsh parenting and financial stress at later child ages for the subsequent development of physical aggression and school-age outcomes. Then again, including such variables would have brought with it even more complexity and have warranted more interpretive caution with respect to cause-effect directions.

Sibling presence merits a special note. We measured the presence of a similar-aged sibling only at study start. This predictor, which constitutes one of the strongest and most consistent predictors of physical aggression, is thus of having an *older* sibling. Our finding is in line with those from other studies [10] exploring the links between older siblings and aggression trajectories (but not with outcomes). A fair proportion of the children we studied, however, acquired younger siblings as the study progressed, and any influences of the presence of younger siblings is unmeasured in our study. The overall predictive effect of sibling presence may thus potentially have been underestimated.

As regards outcomes, we studied each outcome in isolation, not addressing possible comorbidity or overlap. However, physical aggression trajectories may well predict specifically to certain multivariate outcome configurations or profiles (e.g., predictors of a multi-problem outcome profile might be different from what precedes isolated problematic outcomes).

Future research should focus on the potentially intricate interplay of predictors, including gender-moderating effects and relevant interaction effects among predictors (e.g., of parental depression, harsh parenting, child gender and number of siblings). This should also include the study of time-varying predictors (e.g., of financial or other stress across the developmental period, or of language and motor development in tandem with aggressive-behavior development), and predictors that may have an influence at a particular time or age (e.g., transition to day-care or to school). Research should also address trajectory and predictor effects on profiles of outcomes rather than single outcomes in isolation.

Understanding common and less common developmental patterns of physical aggression, the link with subsequent social, behavioral, and academic functioning, and markers for the various patterns, is crucial for early prevention and intervention efforts. The results from our study confirm and expand previous research in underscoring the importance of addressing harsh parenting behavior to influence the developmental pathways of physical aggression during infancy and toddlerhood. Our findings also support the notion that fathers matter and must be included both in research and early prevention and intervention efforts.

Acknowledgements

We are very grateful to the participating families and to the research assistants who contributed to the data collection. We thank Elisabeth Askeland and Anett Apeland for their contribution to the development of the physical aggression measure.

References

1. Tremblay RE. The development of aggressive behavior during childhood: What have we learned in the past century? International Journal of Behavioral Development. 2000;24(2). https://doi.org/10.1080/016502500383232
2. Teymoori A, Côté S, Jones B, Nagin, D, Boivin, M, Vitaro F, et al. Risk factors associated with boys' and girls' developmental trajectories of physical aggression from early childhood through early adolescence. JAMA network open. 2018;1(8), Article e186364. https://doi.org/10.1001/ jamanetworkopen. 2018.6364
3. Alink LRA, Mesman J, Van Zeijl J, Stolk MN, Juffer F, Koot, HM, et al. The early childhood aggression curve: Development of physical aggression in 10- to 50-month-old children. Child Development. 2006;77(4), 954–966. https://doi.org/10.1111 / j.1467-8624.2006.00912.x
4. Lorber MF, Del Vecchio T, Slep AMS. The development of individual physically aggressive behaviors from infancy to toddlerhood. Developmental Psychology. 2018; 54(4), 601-612. https://doi.org/10.1037/dev0000450
5. Hay DF. The early development of human aggression. Child Development Perspectives. 2017;11(2), 102-106. https://doi.org/10.1111/cdep.12220
6. Lorber MF, Del Vecchio T, Slep AMS. The emergence and evolution of infant externalizing behavior. Development and Psychopathology. 2015;27(3), 663-680. https://doi.org/10.1017/S0954579414000923
7. Lorber MF, Del Vecchio T, Slep AM, Scholer SJ. Normative trends in physically aggressive behavior: Age-aggression curves from 6 to 24 months. The Journal of Pediatrics. 2019;206, 197-203. https://doi.org/10.1016/j.jpeds.2018.10.025

8. Côté SM, Vaillancourt T, LeBlanc JC, Nagin DS, Tremblay RE. The development of physical aggression from toddlerhood to pre-adolescence: A nationwide longitudinal study of Canadian children. Journal of Abnormal Child Psychology. 2006;34(1), 68–82. https://doi.org/10.1007/s10802-005-9001-z

9. Nærde A, Ogden T, Janson H, Zachrisson HD. Normative development of physical aggression from 8 to 26 months. Developmental Psychology. 2014;50(6), 1710–1720. https://doi.org/10.1037/a0036324

10. Tremblay RE, Nagin DS, Séguin JR, Zoccolillo M, Zelazo PD, Boivin M, et al. Physical aggression during early childhood: Trajectories and predictors. Pediatrics. 2004; 114(1), 43–50. https://doi.org/10.1542/peds.114.1.e4

11. Baillargeon RH, Zoccolillo M, Keenan K, Côté S, Pérusse D, Wu HX, et al. Gender differences in physical aggression: A prospective population-based survey of children before and after 2 years of age. Developmental Psychology. 2007;43(1), 13-26. https://doi.org/10.1037/0012-1649.43.1.13

12. Broidy LM, Nagin DS, Tremblay RE, Bates J, Brame B, Dodge KA, et al. Developmental trajectories of childhood disruptive behaviors and adolescent delinquency: A six-site, cross-national study. Developmental Psychology. 2003;39(2), 222–245. https://doi.org/10.1037/0012-1649.39.2.222

13. Campbell SB, Spieker S, Burchinal M, Poe MD, The NICHD Early Child Care Research Network. Trajectories of aggression from toddlerhood to age 9 predict academic and social functioning through age 12. Journal of Child Psychology and Psychiatry. 2006; 47(8), 791–800. https://doi.org/10.1111/j.1469-7610.2006.01636.x

14. Campbell SB, Spieker S, Vandergrift N, Belsky J, Burchinal M, The NICHD Early Child Care Research Network. Predictors and sequelae of trajectories of physical

aggression in school-age boys and girls. Development and Psychopathology. 2010; 22(1), 133–150. https://doi.org/10.1017/S0954579409990319

15. Fontaine N, Carbonnea R, Barker ED, Vitaro F, Hébert M, Côté SM, et al. Girls' hyperactivity and physical aggression during childhood and adjustment problems in early adulthood: A 15-year longitudinal study. Archives of General Psychiatry. 2008;65(3), 320-328. https://doi.org/10.1001/archgenpsychiatry.2007.41

16. Nagin D, Tremblay RE. Trajectories of boys' physical aggression, opposition, and hyperactivity on the path to physically violent and nonviolent juvenile delinquency. Child Development. 1999;70(5), 1181–1196. https://doi.org/10.1111/1467-8624.00086

17. Wildeboer A, Thijssen S, Van IJzendoorn MH, Van Der Ende J, Jaddoe VW, Verhulst FC, et al. Early childhood aggression trajectories: Associations with teacher-reported problem behaviour. International Journal of Behavioral Development. 2015;39(3), 221-234. https://doi.org/10.1177/ 0165025414562239

18. Tremblay RE, Japel CJ, Perusse D, Mcduff P, Boivin M, Zoccolillo M, et al. The search for the age of "onset" of physical aggression: Rousseau and Bandura revisited. Criminal Behaviour and Mental Health. 1999;9(1), 8–23. https://doi.org/10.1002/cbm.288

19. Côté SM, Vaillancourt T, Barker ED, Nagin D, Tremblay RE. The joint development of physical and indirect aggression: Predictors of continuity and change during childhood. Development and psychopathology. 2007;19(1), 37-55. https://doi.org/ 10.1017/S0954579407070034

20. Côté SM, Boivin M, Nagin DS, Japel C, Xu Q, Zoccolillo M, et al. The role of maternal education and nonmaternal care services in the prevention of children's physical

aggression problems. Archives of General Psychiatry. 2007;64(11), 1305-1312. doi:10.1001/archpsyc.64.11.1305

21. NICHD Early Child Care Research Network. Trajectories of physical aggression from toddlerhood to middle childhood: Predictors, correlates and outcomes. Monographs of the Society for Research in Child Development. 2004;69(4), vii-129. https://doi.org/10.1111/j.0037-976x.2004.00312.x

22. Jennings WG, Reingle JM. On the number and shape of developmental/life-course violence, aggression, and delinquency trajectories: A state-of-the-art review. Journal of Criminal Justice. 2012; 40(6), 472-489. https://doi.org/10.1016/ j.jcrimjus. 2012.07.001

23. Archer J, Côté S. Sex differences in aggressive behavior: a developmental and evolutionary perspective. In: Tremblay RE, Hartup WW, Archer J, editors. Developmental Origins of Aggression. New York: Guilford Press; 2005. p. 425–43.

24. Jambon M, Madigan S, Plamondon A, Jenkins J. Developmental trajectories of physical aggression and prosocial behavior in early childhood: Family antecedents and psychological correlates. Developmental Psychology. 2019;55(6), 1211-1225. https://doi.org /10.1037/dev0000714

25. Joussemet M, Vitaro F, Barker ED, Côté S, Nagin DS, Zoccolillo M, Tremblay RE. Controlling parenting and physical aggression during elementary school. Child Development. 2008; 79(2), 411-425. https://doi.org/10.1111/j.1467-8624.2007.01133.x

26. Nagin DS, Tremblay RE. Parental and early childhood predictors of persistent physical aggression in boys from kindergarten to high school. Archives of General psychiatry. 2001; 58(4), 389-394. doi:10.1001/archpsyc.58.4.389

27. van Adrichem DS, Huijbregts SC, van der Heijden KB, Van Goozen SH, Swaab H. The role of inhibitory control, attention and vocabulary in physical aggression trajectories from infancy to toddlerhood. Frontiers in Psychology. 2020; 11, 1079. https://doi.org/10.3389/fpsyg.2020.01079

28. Orri M, Boivin M, Chen C, Ahun MN, Geoffroy MC, Ouellet-Morin I, et al. Cohort profile: Quebec longitudinal study of child development (QLSCD). Social Psychiatry and Psychiatric Epidemiology. 2021; 56(5), 883-894. https://doi.org/10.1007/s00127-020-01972-z

29. Hay DF, Waters CS, Perra O, Swift N, Jones R, Jones I, et al. Precursors to aggressive conduct problems are evident by six months of age. Developmental Science. 2014; 17(3), 471–480. doi:10.1111/desc.12133

30. Brendgen M, Dionne G, Girard A, Boivin M, Vitaro F, Pérusse D. Examining genetic and environmental effects on social aggression: A study of 6-year-old twins. Child Development. 2005; 76(4), 930-946. https://doi.org/10.1111/j.1467-8624.2005.00887.x

31. Dionne G, Tremblay R, Boivin M, Laplante D, Pérusse D. Physical aggression and expressive vocabulary in 19-month-old twins. Developmental Psychology. 2003; 39(2), 261–273. https://doi.org/10.1037/0012-1649.39.2.261

32. Provençal N, Booij L, Tremblay RE. The developmental origins of chronic physical aggression: biological pathways triggered by early life adversity. Journal of Experimental Biology. 2015;218(1), 123-133. doi:10.1242/jeb.111401

33. Tremblay RE, Vitaro F, Côté SM. Developmental origins of chronic physical aggression: a bio-psycho-social model for the next generation of preventive

interventions. Annual Review of Psychology. 2018;69(1), 383-407. https://doi.org/10.1111/ j.1469-7610.2010.02211.x

34. Tremblay RE. Developmental origins of chronic physical aggression: An international perspective on using singletons, twins and epigenetics. European Journal of Criminology. 2015;12(5), 551-561. https://doi.org/10.1177/1477370815600617

35. Brumley LD, Jaffee SR. Defining and distinguishing promotive and protective effects for childhood externalizing psychopathology: a systematic review. Social Psychiatry and Psychiatric Epidemiology. 2016;51(6), 803-815. https://doi.org/10.1007/ s00127-016-1228-1

36. Bronfenbrenner U. The ecology of human development: Experiments by nature and design. Harvard university press, 1979

37. Labella MH, Masten AS. Family influences on the development of aggression and violence. Current Opinion in Psychology. 2018;19, 11-16. https://doi.org/10.1016 /j.copsyc.2017.03.028

38. Huijbregts SC, Seguin JR, Zoccolillo M, Boivin M, Tremblay RE. Maternal prenatal smoking, parental antisocial behavior, and early childhood physical aggression. Development and Psychopathology. 2008;20(2), 437-453. DOI: https://doi.org /10.1017/S0954579408000217

39. Kinge JM, Øverland S, Flatø M, Dieleman J, Røgeberg O, Magnus MC, et al. Parental income and mental disorders in children and adolescents: prospective register-based study. International Journal of Epidemiology. 2021;50(5), 1615-1627. doi: 10.1093/ije /dyab066

40. Peverill M, Dirks MA, Narvaja T, Herts KL, Comer JS, McLaughlin KA. Socioeconomic status and child psychopathology in the United States: A meta-analysis of

population-based studies. Clinical Psychology Review. 2021;83, 101933. https://doi.org/10.1016/j.cpr.2020.101933

41. Conger RD, Ge X, Elder Jr GH, Lorenz FO, Simons RL. Economic stress, coercive family process, and developmental problems of adolescents. Child Development. 1994; 65(2), 541-561. https://doi.org/10.1111/j.1467-8624.1994.tb00768.x

42. Conger RD, Conger KJ, Martin MJ. Socioeconomic status, family processes, and individual development. Journal of Marriage and Family. 2010;72(3), 685-704. https://doi.org/10.1111/j.1741-3737.2010.00725.x

43. Masarik AS, Conger RD. Stress and child development: A review of the Family Stress Model. Current Opinion in Psychology. 2017;13, 85-90. https://doi.org/10.1016/ j.copsyc.2016.05.008

44. Mazza JR, Pingault JB, Booij L, Boivin M, Tremblay R, Lambert J, et al. Poverty and behavior problems during early childhood: The mediating role of maternal depression symptoms and parenting. International Journal of Behavioral Development. 2017;41(6), 670-680. https://doi.org/10.1177/0165025416657615

45. Dunn J, Munn P. Sibling quarrels and maternal intervention: Individual differences in understanding and aggression. Journal of Child Psychology and Psychiatry. 1986; 27(5), 583-595. https://doi.org/10.1111/j.1469-7610.1986.tb00184.x

46. Garcia MM, Shaw DS, Winslow EB, Yaggi KE. Destructive sibling conflict and the development of conduct problems in young boys. Developmental Psychology. 2000; 36(1), 44–53. http://dx.doi.org/10.1037/0012-1649.36.1.44

47. Mesman J, Stoel R, Bakermans-Kranenburg MJ, van IJzendoorn MH, Juffer F, Koot HM, et al. Predicting growth curves of early childhood externalizing problems: Differential susceptibility of children with difficult temperament. Journal of

Abnormal Child Psychology. 2009;37(5), 625-636. https://doi.org/10.1007/s10802-009-9298-0

48. Olson SL, Ip KI, Gonzalez R, Beyers-Carlson EE, Volling, BL. Development of externalizing symptoms across the toddler period: The critical role of older siblings. Journal of Family Psychology. 2020;34(2), 165-174. https://doi.org/10.1037/fam0000581

49. Patterson GR. Siblings: Fellow travelers in coercive family processes. In: Blanchard RJ, editor. Advances in the study of aggression New York: Academic Press; 1984. p. 174-213.

50. Patterson G. R. The contribution of siblings to training for fighting: A microsocial analysis. In: Olweus D, Block J, Radke-Yarrow M, editors. Development of antisocial and prosocial behavior: Research, theories, and issues. Orlando, Florida: Academic; 1986, p. 235–261.

51. Caspi, J. Sibling aggression: Assessment and treatment. New York, NY: Springer; 2012.

52. Eriksen S, Jensen V. A push or a punch: Distinguishing the severity of sibling violence. Journal of Interpersonal Violence. 2009; 24(1), 183-208. https://doi.org/10.1177/0886260508316298

53. Patterson GR. The early development of coercive family processes. In: Reid JB, Patterson GR, Snyder J, editors. Antisocial behavior in children and adolescents: A developmental analysis and model for intervention. American Psychological Association; 2002. p. 25–44.

54. Shaw DS, Winslow EB, Owens EB, Vondra JI, Cohn JF, Bell RQ. The development of early externalizing problems among children from low-income families: A

transformational perspective. Journal of Abnormal Child Psychology. 1998;26(2), 95-107. https://doi.org/10.1023/A:1022665704584

55. Goodman SH, Rouse MH, Connell AM, Broth MR, Hall CM, Heyward D. Maternal depression and child psychopathology: A meta-analytic review. Clinical Child and Family Psychology Review. 2011;14(1), 1–27. https://doi.org/10.1007/s10567-010-0080-1

56. Wilson S, Durbin CE. Effects of paternal depression on fathers' parenting behaviors: A meta-analytic review. Clinical Psychology Review. 2010;30(2), 167-180. https://doi.org/10.1016/j.cpr.2009.10.007

57. Cheung K, Theule J. Paternal depressive symptoms and parenting behaviors: An updated meta-analysis. Journal of Child and Family Studies. 2019;28(3), 613-626. https://doi.org/10.1007/s10826-018-01316-1

58. Yang, P, Schlomer GL, Lippold MA. Mothering versus fathering? Positive parenting versus negative parenting? Their relative importance in predicting adolescent aggressive behavior: A longitudinal comparison. Developmental psychology. 2022; 59(1), 69-83. https://doi.org/10.1037 /dev0001442.

59. Cabrera NJ, Volling BL, Barr R. Fathers are parents, too! Widening the lens on parenting for children's development. Child Development Perspectives. 2018;12(3), 152-157. https://doi.org/10.1111/cdep.12275

60. Cabrera NJ. Father involvement, father-child relationship, and attachment in the early years. Attachment & Human development. 2020;22(1), 134-138. https://doi. org /10.1080/14616734.2019.1589070

1529 61. Haas LL, Hwang CP. Fatherhood and social policy in Scandinavia. In: Shwalb DW,
1530 Shwalb BJ, Lamb ME, editors. Fathers in cultural contexts. New York, NY: Routledge;
1531 2013. p. 303–331.
1532 62. Rege M, Solli IF. The impact of paternity leave on long-term father involvement.
1533 CESifo Working Paper Series No. 3130. 2010. Available at SSRN:
1534 https://ssrn.com/abstract =1649344
1535 63. Sullivan O, Coltrane S, McAnnally L, Altintas E. Father-friendly policies and time-use
1536 data in a cross-national context: Potential and prospects for future research. The
1537 Annals of the American Academy of Political and Social Science, 2009; 624(1), 234–
1538 254. https:// doi.org/10.1177/0002716209335138
1539 64. Rodrigues M, Sokolovic N, Madigan S, Luo Y, Silva V, Misra S. et al. Paternal
1540 sensitivity and children's cognitive and socioemotional outcomes: A meta-analytic
1541 review. Child Development. 2021;92(2), 554-577. https://doi.org/10.1111/
1542 cdev.13545
1543 65. Card NA, Stucky BD, Sawalani GM, Little TD. Direct and indirect aggression during
1544 childhood and adolescence: A meta-analytic review of gender differences,
1545 intercorrelations, and relations to maladjustment. Child Development 2008;79(5),
1546 1185–1229. https://doi.org/10.1111 /j.1467-8624.2008.01184.x
1547 66. Hay DF. The gradual emergence of sex differences in aggression: alternative
1548 hypotheses. Psychological Medicine. 2007;37(11), 1527-1537. DOI: https://doi.org
1549 /10.1017/S0033291707000165
1550 67. Hay DF, Nash A, Caplan M, Swartzentruber J, Ishikawa F, Vespo JE. The emergence of
1551 gender differences in physical aggression in the context of conflict between young

peers. British Journal of Developmental Psychology. 2011;29(2), 158–175. doi:10.1111/j.2044-835X.2011.02028.x

68. Hay DF. The beginnings of aggression in infancy. In: Tremblay RE, Hartup WW, Archer J, editors. Developmental origins of aggression. New York: Guilford Press; 2005. p. 107–132.

69. Côté SM. Sex differences in physical and indirect aggression: A developmental perspective. European Journal on Criminal Policy and Research. 2007;13(3), 183-200. https://doi.org/10.1007/s10610-007-9046-3

70. van Zeijl J, Mesman J, Stolk MN, Alink LR, Van IJzendoorn MH, Bakermans-Kranenburg MJ, et al. Differential susceptibility to discipline: the moderating effect of child temperament on the association between maternal discipline and early childhood externalizing problems. Journal of Family Psychology. 2007;21(4), 626-636. https://doi.org/10.1037/0893-3200.21.4.626

71. Vitaro F, Barker ED, Boivin M, Brendgen M, Tremblay RE. Do early difficult temperament and harsh parenting differentially predict reactive and proactive aggression? Journal of Abnormal Child Psychology. 2006;34(5), 681–691. https://doi.org/ 10.1007/s10802-006-9055-6

72. Slagt M, Dubas JS, Deković M, van Aken, MA. Differences in sensitivity to parenting depending on child temperament: A meta-analysis. Psychological Bulletin. 2016; 142(10), 1068-1110. https://doi.org/10.1037/bul0000061

73. Belsky J. Variation in susceptibility to environmental influence: An evolutionary argument. Psychological Inquiry. 1997;8(3), 182-186. DOI: 10.1207/s15327965pli0803_3

74. Belsky J. Theory testing, effect-size evaluation, and differential susceptibility to rearing influence: The case of mothering and attachment. Child Development. 1997; 68(4), 598-600. https://doi.org/10.1111/j.1467-8624.1997.tb04221.x

75. Belsky J, Hsieh KH, Crnic K. Mothering, fathering, and infant negativity as antecedents of boys' externalizing problems and inhibition at age 3 years: Differential susceptibility to rearing experience? Development and Psychopathology. 1998;10(2), 301-319. DOI: https://doi.org/10.1017/S095457949800162X

76. Campbell SB, Shaw DS, Gilliom M. Early externalizing behavior problems: Toddlers and preschoolers at risk for later maladjustment. Development and Psychopathology 2000; 12(3), 467– 488. DOI: https://doi.org/10.1017/ S0954579400003114

77. Brame B, Nagin DS, Tremblay RE. Developmental trajectories of physical aggression from school entry to late adolescence. The Journal of Child Psychology and Psychiatry and Allied Disciplines. 2001;42(4), 503-512. DOI: https://doi.org/10.1017/ S0021963001007120

78. Cleverley K, Szatmari P, Vaillancourt T, Boyle M, Lipman E. Developmental trajectories of physical and indirect aggression from late childhood to adolescence: Sex differences and outcomes in emerging adulthood. Journal of the American Academy of Child & Adolescent Psychiatry. 2012;51(10), 1037-1051. https://doi.org/ 10.1016/j.jaac.2012.07.010

79. Kokko K, Tremblay RE, Lacours, E, Nagin DS, Vitaro F. Trajectories of prosocial behavior and physical aggression in middle childhood: Links to adolescent school dropout and physical violence. Journal of Research on Adolescence. 2006;16(3), 403-428. https://doi.org/10.1111/j.1532-7795.2006.00500.x

80. Kokko K, Pulkkinen L, Huesmann LR, Dubow EF, Boxer P. Intensity of aggression in childhood as a predictor of different forms of adult aggression: A two-country (Finland and the United States) analysis. Journal of Research on Adolescence. 2009; 19(1), 9-34. https://doi.org/10.1111/j.1532-7795.2009.00579.x

81. Reef J, Diamantopoulou S, van Meurs I, Verhulst F, van der Ende J. Predicting adult emotional and behavioral problems from externalizing problem trajectories in a 24-year longitudinal study. European Child & Adolescent Psychiatry. 2010;19(7), 577-585. https://doi.org/10.1007/s00787-010-0088-6

82. Temcheff CE, Serbin LA, Martin-Storey A, Stack DM, Hodgins S, Ledingham J, et al. Continuity and pathways from aggression in childhood to family violence in adulthood: A 30-year longitudinal study. Journal of Family Violence. 2008;23(4), 231-242. https://doi.org/10.1007/s10896-007-9147-2

83. Carbonneau R, Boivin M, Brendgen M, Nagin D, Tremblay RE. Comorbid development of disruptive behaviors from age 1½ to 5 years in a population birth-cohort and association with school adjustment in first grade. Journal of Abnormal Child Psychology. 2016;44(4), 677-690. https://doi.org/10.1007/s10802-015-0072-1

84. Asparouhov T, Muthén B. Auxiliary variables in mixture modeling: Three-step approaches using Mplus. Structural Equation Modeling. 2014;21(3), 329-341. https://doi.org/10.1080/10705511.2014.915181

85. Nærde A, Janson H, Ogden T. BONDS (The Behavior Outlook Norwegian Developmental Study): A prospective longitudinal study of early development of social competence and behavior problems. Oslo: The Norwegian Center for Child Behavioral Development; 2014. Report. ISBN 978-82-93406-00-6.

86. Bond TG, Fox CM. Applying the Rasch model: Fundamental measurement in the human sciences (3rd ed.). Routledge; 2015.

87. Linacre JM. Estimating Rasch measures with known polytomous (or rating scale) item difficulties: Anchored Maximum Likelihood Estimation (AMLE). Rasch Measurement Transactions. 1998;12(2), 638. https://www.rasch.org/rmt/rmt122q.htm

88. Blum JS, Mehrabian A. Personality and temperament correlates of marital satisfaction. Journal of Personality. 1999; 67(1), 93–125.

89. Winokur A, Winokur DF, Rickels K, Cox DS. Symptoms of emotional distress in a family planning service: Stability over a four-week period. British Journal of Psychiatry. 1984;144(4), 395-399. https://doi.org/10.1192/bjp.144.4.395

90. Tambs K, Moum T. How well can a few questionnaire items indicate anxiety and depression? Acta Psychiatrica Scandinavica. 1993;87(5), 364–367. https://doi.org/10.1111/j.1600-0447.1993.tb03388.x

91. Aspland H, Gardner F. Observational measures of parent-child interaction: An introductory review. Child and Adolescent Mental Health. 2003;8(3),136–143. http://dx.doi.org/10.1111/1475-3588.00061

92. Snyder J, Stoolmiller M, Wilson M, Yamamoto M. Child anger regulation, parental responses to children's anger displays, and early child antisocial behavior. Social Development. 2003;12*(3)*, 335–360. http://dx.doi.org/10.1111/1467-9507.00237

93. Cox MJ, Crnic K. Qualitative ratings for parent–child interaction at 3–12 months of age. Unpublished manuscript. Department of Psychology, University of North Carolina at Chapel Hill; 2003.

94. Nordahl KB. Early father-child interaction in a father-friendly context: Gender differences, child outcomes, and predictive factors related to fathers' parenting behaviors with one-year-olds [dissertation]. University of Bergen, Norway; 2014.

95. Loeber R, Stouthamer-Loeber M, Van Kammen W, Farrington D. Initiation, escalation and desistance in juvenile offending and their correlates. Journal of Criminal Law and Criminology. 1991;82(1), 36-82.

96. Thornberry T, Huizinga D, Loeber R. The prevention of serious delinquency and violence: Implications from the Program of Research on the Causes and Correlates of Delinquency. In: Howell J, Krisberg B, Hawking D, Wilson JD, editors. Sourcebook on serious, violent and chronic juvenile offenders Sage; 1995. p. 213-327.

97. Gartstein MA, Rothbart MK. Studying infant temperament via the Revised Infant Behavior Questionnaire. Infant Behavior & Development. 2003;26(1), 64–86. https://doi.org/10.1016/S0163-6383(02)00169-8

98. Rothbart MK. Measurement of temperament in infancy. Child Development. 1981; 52(2), 569-578. https://doi.org/10.2307/1129176

99. Gresham FM, Elliott SN. Social Skills Improvement System: Rating scales. Bloomington, MN: Pearson Assessments 2008.

100. van Buuren S, Groothuis-Oudshoorn K. mice: Multivariate imputation by chained equations in R. Journal of Statistical Software. 2011;45(3), 1-67. https://doi.org/10.18637/jss.v045.i03

101. Muthén B. Latent variable hybrids: Overview of old and new models. In: Hancock GR, Samuelsen KM, editors. Advances in latent variable mixture models. Charlotte, NC: Information Age Publishing, Inc; 2008. p. 1-24.

102. Tofighi D, Enders CK. Identifying the correct number of classes in growth mixture models. In: Hancock GR, Samuelsen KM, editors. Advances in latent variable mixture models. Information Age Publishing 2008, pp. 317-341.

103. Nagelkerke NJ. A note on a general definition of the coefficient of determination. Biometrika. 1991;78(3), 691-692.

104. Nylund K, Asparouhov T, Muthen BO. Deciding on the number of classes in latent class analysis and growth mixture modeling: A Monte Carlo simulation study. Structural Equation Modeling. 2007;14(4), 535–569. https://doi.org/10.1080/10705510701575396

105. Dziak JJ, Lanza ST, Tan X. Effect size, statistical power and sample size requirements for the bootstrap likelihood ratio test in latent class analysis. Structural Equation Modeling. 2014; 21(4): 534–552. doi:10.1080/10705511.2014.919819.

106. McLachlan G, Peel D. Finite Mixture Models. New York, NY: John Wiley and Sons Inc 2000.

107. Cox MJ, Paley B. Families as systems. Annual Review of Psychology. 1997;48*(1)*, 243–267. https://doi.org/10.1146/annurev.psych.48.1.243

108. Kerr ME, Bowen M. Family evaluation: An approach based on Bowen theory. WW Norton & Co 1988.

109. Olson S, Lopez-Duran N, Lunkenheimer E, Chang H, Sameroff A. Individual differences in the development of early peer aggression: Integrating contributions of self-regulation, theory of mind, and parenting. Development and Psychopathology. 2011;23(1), 253-266. doi:10.1017/S095457941000077

110. Spinrad TL, Eisenberg N, Gaertner B, Popp T, Smith CL, Kupfer A, et al. Relations of maternal socialization and toddlers' effortful control to children's adjustment and social competence. Developmental psychology. 2007; 43(5), 1170.

111. Sulik MJ, Blair C, Mills-Koonce R, Berry D, Greenberg M. Early parenting and the development of externalizing behavior problems: Longitudinal mediation through children's executive function. Child Development. 2015;86(5), 1588–1603. https://doi.org/10.1111/cdev.12386

112. Hoff E, Laursen B. Socioeconomic status and parenting. In: Bornstein MH, editor. Handbook of parenting. Volume 2: Biology and Ecology of Parenting. Routledge 2019, pp. 421-447.

113. Shaw DS, Shelleby EC. Early-onset conduct problems: Intersection of conduct problems and poverty. Annual Review of Clinical Psychology. 2014;10(503), 503-528. doi: 10.1146/annurev-clinpsy-032813-153650

114. Andersen E. Økt fruktbarhet for første gang på 12 år (Increased fertility for the first time in 12 years). Statistics Norway 2022, March 9. Available from: https://www.ssb.no/befolkning/fodte-og-dode/statistikk/fodte /artikler/okt-fruktbarhet-for-forste-gang-pa-12-ar

115. Bakermans-Kranenburg MJ, Lotz A, Alyousefi-van Dijk K, van IJzendoorn M. Birth of a father: Fathering in the first 1,000 days. Child Development Perspectives. 2019; 13(4), 247-253. https://doi.org/10.1111/cdep.12347

116. Leventhal T, Brooks-Gunn J. The neighborhoods they live in: The effects of neighborhood residence on child and adolescent outcomes. Psychological Bulletin. 2000;126(2), 309. https://doi.org/10.1037/0033-2909.126.2.309

1712 117. Poulain T, Vogel M, Kiess W. Review on the role of socioeconomic status in child
1713 health and development. Current Opinion in Pediatrics. 2020; 32(2), 308-314.
1714 https://doi.org/ 10.1097/mop.0000000000000876
1715 118. Karlberg P, Taranger J, Engström I, Klackenberg-Larsson I, Lichtenstein H, Stensson J
1716 et al. The development of children in a Swedish urban community: A prospective,
1717 longitudinal study. Acta Paediatrica. *1968*; 57(S187), 48-66. https://doi.org/10.1111
1718 /j.1651-2227.1968.tb06040.x
1719 119. Patterson GR, Chamberlain P, Reid JB. A comparative evaluation of a parent-training
1720 program. Behavior Therapy. 1982; 13(5), 638-650. https://doi.org/10.1016/S0005-
1721 7894(82)80021-X
1722 120. Singer JD, Willett JB. Applied longitudinal data analysis: Modeling change and event
1723 occurrence. Oxford University Press; 2003. https://doi.org/10.1093/ acprof:oso/
1724 9780195152968.001.0001
1725 121. Andrich D. Rasch models for measurement. Sage; 1988. https://dx.doi.org/ 10.4135/
1726 9781412985598.
1727 122. Olsen R, Garratt A, Iversen H, Bjertnaes O. Rasch analysis of the Psychiatric Out-
1728 Patient Experiences Questionnaire (POPEQ). BMC Health Services Research. 2010;
1729 10, Article 282. https://doi.org/10.1186/1472-6963-10-282
1730 123. Janson H, Bringedal, GE, Nærde A. A flowchart of participation and collected data
1731 across 15 years of The Behavior Outlook Norwegian Developmental Study (BONDS).
1732 Oslo, Norway: The Norwegian Center for Child Behavioral Development; 2023.
1733 Report. https://www.nubu.no/getfile.php /1323354-1687432522/ Filer/nubu.no
1734 /Intranett/Dokumenter/Rapporter/2023%20Report%20BONDS%20Participation.pdf
1735

Supporting information

Supporting information A: Physical aggression measure

SA1 Table. Physical aggression measure developed for the Behavior Outlook Norwegian Developmental Study.

SA2 Table. Physical aggression item inclusion in data collections at ages 1-5 years in the Behavior Outlook Norwegian Developmental Study.

Supporting information B: Scaling of the physical aggression measure

SB1 Table. Final Rasch-model item scale value estimates at target ages.

Supporting information C: Attrition

SC1 Table. Descriptive data for predictors in the original sample, for children with any physical aggression data, and for children with Grade 2 outcome data.

Supporting information D: Additional comparisons of 2-10 trajectory solutions.

SD1 Table. Initial and final starts, replications of best loglikelihood, and final starts replications and converged for 2 to 10 trajectory models.

SD2 Table. Bootstrap likelihood ratio tests for 2- to 9-trajectory solutions.

SD Fig. Trajectory Plots for 2-9 Trajectory Solutions.

Supporting information E: Details of the 9-trajectory solution.

SE1 Table. Class counts and proportions for latent classes based on estimated posterior probabilities.

SE2 Table. Estimated parameters with associated standard errors of trajectory parameters for the 9-trajectory solution.

SE3 Table. Average latent class probabilities for the most likely latent class membership (row) by latent class (column).

1759 **SE4 Table. Classification probabilities for the most likely latent class membership (column)**

1760 **by latent class (row).**

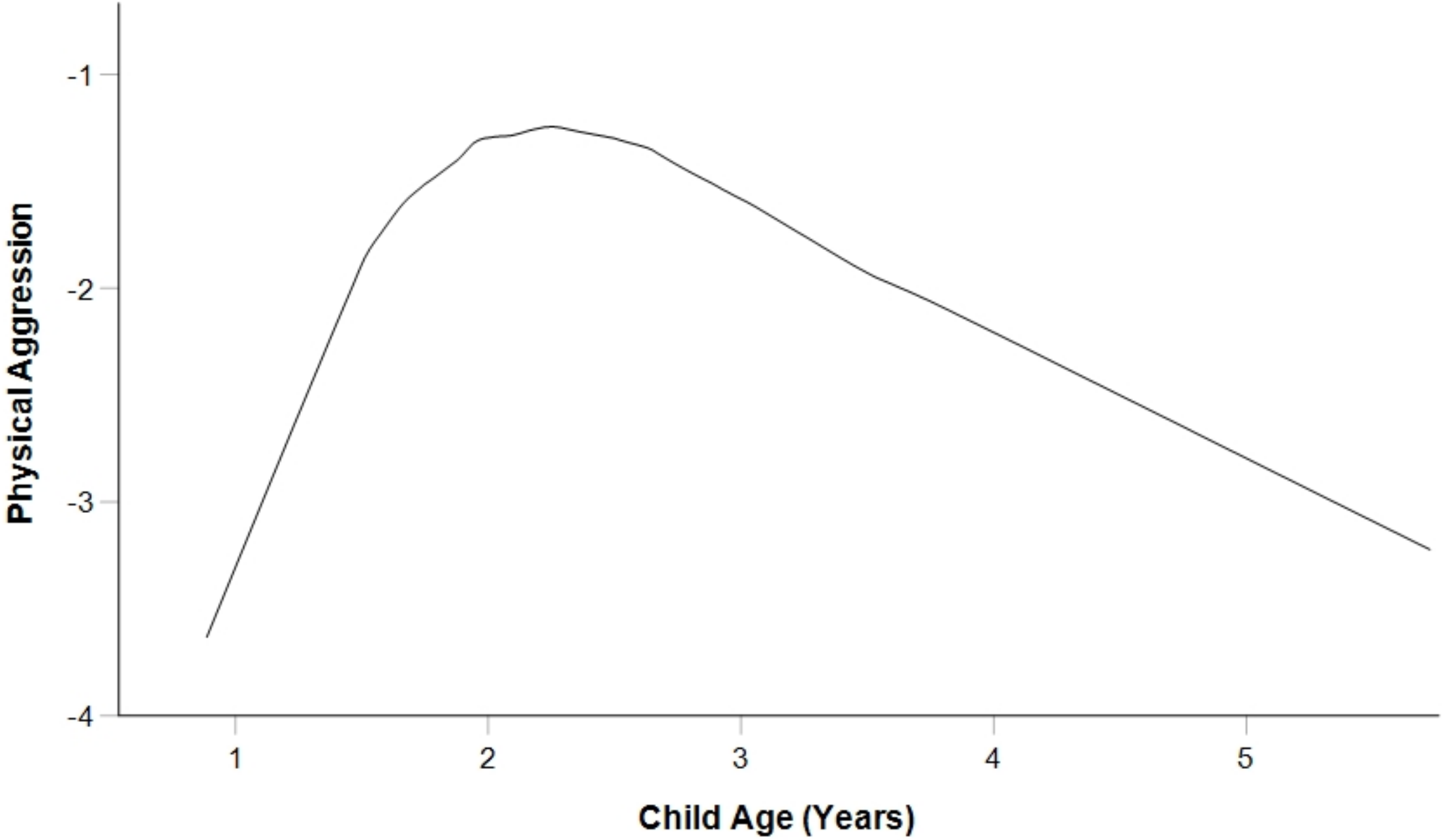

Figure 1

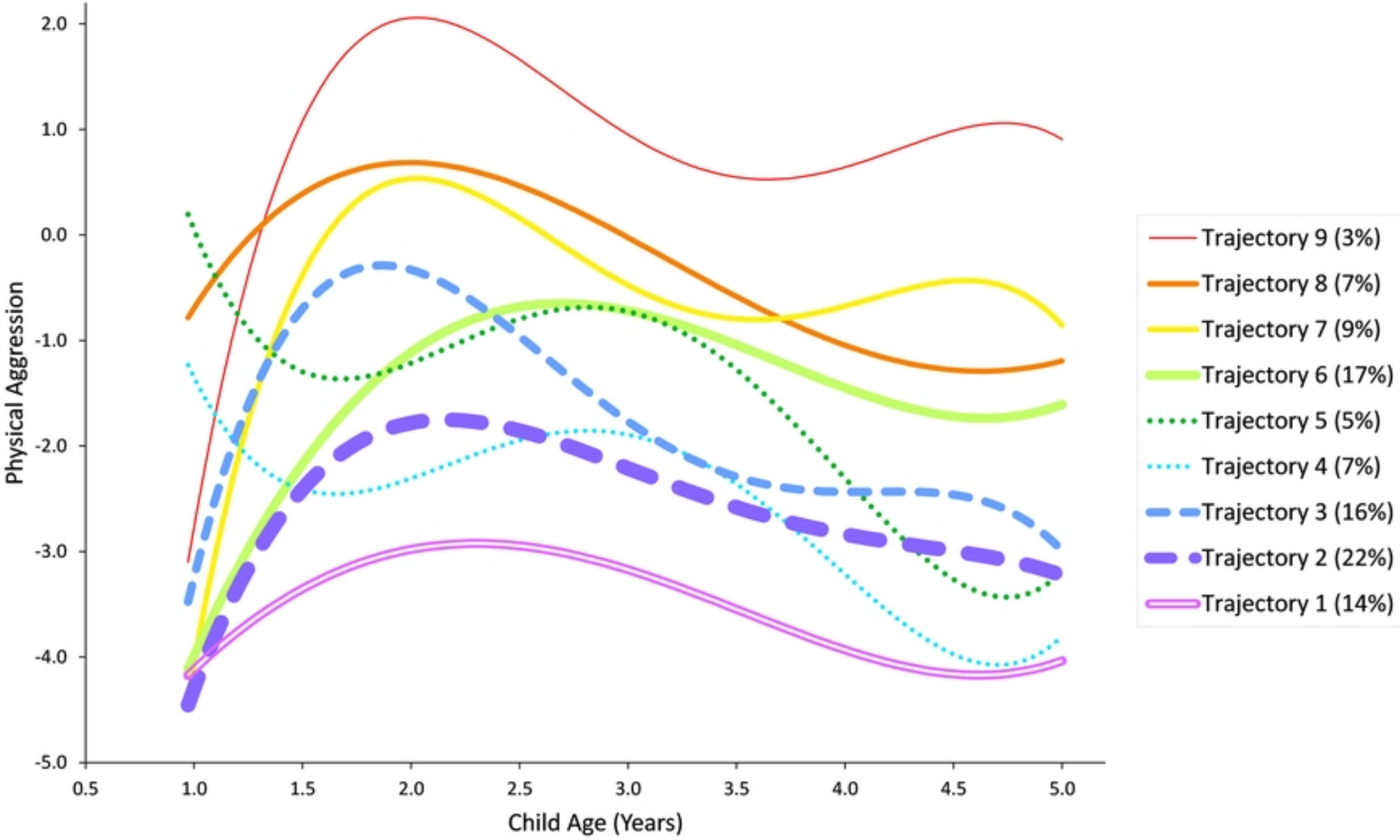
2.0
1.0
0.0
-1.0
-2.0
-3.0
-4.0
-5.0
Physical Aggression
0.5
1.0
1.5
2.0
2.5
3.0
3.5
4.0
4.5
5.0
Child Age (Years)
Trajectory 9 (3%)
Trajectory 8 (7%)
Trajectory 7 (9%)
Trajectory 6 (17%)
Trajectory 5 (5%)
Trajectory 4 (7%)
Trajectory 3 (16%)
Trajectory 2 (22%)
Trajectory 1 (14%)

Figure 2