UNIVERSITY OF AMSTERDAM

UvA-DARE (Digital Academic Repository)

Population-Level Transitions in Observed Difficulties Through Childhood and Adolescence

Benhamou, Elia; Akarca, Danyal; Bathelt, Joe; Fletcher-Watson, Sue; Astle, Duncan E.

DOI
10.1037/dev0001874

Publication date
2025

Document Version
Final published version

Published in
Developmental Psychology

License
CC BY

Link to publication

Citation for published version (APA):
Benhamou, E., Akarca, D., Bathelt, J., Fletcher-Watson, S., & Astle, D. E. (2025). Population-Level Transitions in Observed Difficulties Through Childhood and Adolescence. *Developmental Psychology*, *61*(8), 1495-1515. https://doi.org/10.1037/dev0001874

General rights
It is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), other than for strictly personal, individual use, unless the work is under an open content license (like Creative Commons).

Disclaimer/Complaints regulations
If you believe that digital publication of certain material infringes any of your rights or (privacy) interests, please let the Library know, stating your reasons. In case of a legitimate complaint, the Library will make the material inaccessible and/or remove it from the website. Please Ask the Library: https://uba.uva.nl/en/contact, or a letter to: Library of the University of Amsterdam, Secretariat, P.O. Box 19185, 1000 GD Amsterdam, The Netherlands. You will be contacted as soon as possible.

UvA-DARE is a service provided by the library of the University of Amsterdam (https://dare.uva.nl)

Download date:08 Nov 2026

AMERICAN PSYCHOLOGICAL ASSOCIATION

Developmental Psychology

© 2025 The Author(s)
ISSN: 0012-1649

2025, Vol. 61, No. 8, 1495–1515
https://doi.org/10.1037/dev0001874

Population-Level Transitions in Observed Difficulties Through Childhood and Adolescence

Elia Benhamou[1], Danyal Akarca[1], Joe Bathelt[2, 3], Sue Fletcher-Watson[4], and Duncan E. Astle[1, 5]

[1] MRC Cognition and Brain Sciences Unit, University of Cambridge
[2] Department of Psychology, Royal Holloway, University of London
[3] Department of Psychology, University of Amsterdam
[4] The Salvesen Mindroom Research Centre, University of Edinburgh
[5] Department of Psychiatry, University of Cambridge

In an attempt to better characterize the complexity of difficulties observed within developing populations, numerous data-driven techniques have been applied to large mixed data sets. However, many have failed to incorporate the core role of developmental time in these approaches, that is, the typical course of change in behavioral features that occurs over childhood to adolescence. In this study, we utilized manifold projections alongside a gradient-boosting model on data collected from the Millennium Cohort Study to unpack the central role of developmental time in how behavioral difficulties transition between the ages of 5, 11, and 17. Our analysis highlights numerous observations: (a) Girls develop relatively greater internalized behavioral problems during adolescence; (b) in the case of a chaotic home environment, co-occurring internalizing and externalizing difficulties tend to persist during childhood; (c) peer problems were the most likely to persist over the whole 12-year period (especially in the presence of early maternal depression and poor family relationships); and (d) there were two pathways with distinct risk factors leading to antisocial behaviors in adolescence—an early-childhood onset pathway and later adolescent onset pathway. Our findings provide evidence that investigations of child and adolescent difficulties must be open to the possibility of multiple subgroups and variability in trajectory over time. We further highlight the crucial role of family and social support and school experience-related factors in predicting children's outcomes.

Public Significance Statement

Using a large sample of data from the Millennium Cohort Study, followed from the age of 9 months to 17 years, we use data-driven methods to examine developmental change in children's observed behaviors as they grow up. We identify clusters of children over more than a decade and establish key factors that predict longitudinally particular profiles and transitions from childhood into adolescence.

Keywords: development, machine learning, Millennium Cohort study, childhood, adolescence

Supplemental materials: https://doi.org/10.1037/dev0001874.supp

This article was published Online First February 27, 2025.

Irma Arteaga served as action editor.

Danyal Akarca https://orcid.org/0000-0002-5931-0295

Duncan E. Astle https://orcid.org/0000-0002-7042-5392

The data and code that support the findings of this study are available from the corresponding author, Duncan E. Astle, upon reasonable request. All data exclusions, all data transformations, and all measures in the study are reported. All analysis code is available at https://github.com/eliacello/MCS_BehaviouralTransitions_2021/tree/main. This study's design and its analysis were not preregistered. The authors have no known conflicts of interest to disclose. The views expressed are those of the author(s) and do not necessarily reflect those of the National Institute for Health and Care Research (NIHR) or the Department of Health and Social Care.

The authors thank the James S. McDonnell Foundation Opportunity Award for funding awarded to Elia Benhamou, Danyal Akarca, Sue Fletcher-Watson, and Duncan E. Astle. Duncan E. Astle is supported by funding from the Medical Research Council United Kingdom (Grant MC-A0606-5PQ41). All research at the Department of Psychiatry in the University of Cambridge was supported by the NIHR Cambridge Biomedical Research Centre (NIHR203312) and the NIHR Applied Research Collaboration East of England.

Open Access funding provided by University of Cambridge: This work is licensed under a Creative Commons Attribution 4.0 International License (CC BY 4.0; https://creativecommons.org/licenses/by/4.0). This license permits copying and redistributing the work in any medium or format, as well as adapting the material for any purpose, even commercially.

Elia Benhamou played a lead role in formal analysis, investigation, visualization, and writing–original draft and an equal role in conceptualization, methodology, project administration, and writing–review and editing. Danyal Akarca played a supporting role in conceptualization, methodology, supervision, and writing–review and editing and an equal role in funding acquisition. Joe Bathelt played a supporting role in formal analysis and methodology. Sue Fletcher-Watson played a supporting role in methodology, supervision, and writing–review and editing and an equal role in conceptualization and funding acquisition. Duncan E. Astle played a lead role in funding acquisition, supervision, and writing–review and editing, a supporting role in writing–original draft, and an equal role in conceptualization and methodology.

Correspondence concerning this article should be addressed to Duncan E. Astle, Department of Psychiatry, University of Cambridge, Cambridge, CB2 2QQ, United Kingdom. Email: duncan.astle@mrc-cbu.cam.ac.uk

Existing classification systems that seek to categorize neurodevelopmental, psychiatric, and mental health conditions into discrete diagnostic entities have been widely critiqued (Astle et al., 2022; Cuthbert & Insel, 2013; Dalgleish et al., 2020; Etkin & Cuthbert, 2014). These systems struggle to accommodate the variability within, and shared features across, categories. They also struggle to accommodate profiles that do not neatly map onto diagnostic criteria but which are nonetheless impactful. As a result, people may be prevented to access support because their pattern of difficulties are not easily classified within existing structures (Hallett & Crompton, 2018). In an attempt to recharacterize the range and complexity of difficulties observed and experienced within a population, data-driven techniques such as latent class, factor analyses, or hierarchical clustering have been applied to large mixed data sets, which include clinical measures but also cognitive, sociodemographic, home–life, and mental health measurements. Cognitive or behavioral dimensions identified with latent class (Lacey et al., 2022; Swift et al., 2016), analyses of covariance (Dolan & Lennox, 2013), or exploratory factor analysis (Furlong et al., 2018; Holmes et al., 2021) can capture broad axes of variation and thereby provide a parsimonious account of the sample. In addition, clustering methods are increasingly being used to identify subgroups with similar characteristics, which may cut across established diagnostic categories (Archibald et al., 2013; Astle et al., 2019; Kushki et al., 2019).

However, these approaches—diverse samples combined with data-driven analysis—have largely ignored the role of *developmental time*. Where time has been considered, variability between individuals, between measures, and across time are each considered separately (Jacob et al., 2019; Wardenaar & de Jonge, 2013). This is problematic because behavioral problems seem to follow distinct developmental courses (Lewis & Rudolph, 2014). For instance, most children aged 3–5 with conduct and hyperactivity difficulties do not experience these difficulties by age 9–11 (Moffitt, 2006). Likewise, the prevalence of anxiety symptoms is high in late childhood, decreases in early adolescence, and then increases again from middle adolescence onward (Copeland et al., 2014; Van Oort et al., 2009). In short, behavioral difficulties need to be understood within their distinctive developmental context.

Studies that have examined developmental change have tended to focus on a single behavioral characteristic of interest. However, co-occurrence rates are high within neurodevelopmental and child psychiatric populations (Caspi & Moffitt, 2018; Hansen et al., 2018). For example, prevalence data for children with an attention-deficit/hyperactivity disorder (ADHD) diagnosis reveal that 40%–75% also meet criteria for oppositional defiant disorder, 14%–24% for conduct disorder, and approximately one third for an anxiety disorder (Mayes et al., 2009). To complicate matters further, the co-occurrence of problems fluctuates over time (Franke et al., 2018). In ADHD, new co-occurring difficulties tend to accumulate during adolescence and adulthood, for example, anxiety disorders, obsessive–compulsive disorder, and conduct disorders (Taurines et al., 2010). In contrast, some difficulties that often co-occur with the original presentation dissipate, for example, ADHD often co-occurs with a tic disorder, but the presentation of the tics usually improves substantially over developmental time (Black et al., 2021).

In one example, Bathelt et al. (2021) used a data-driven approach to capture the range, complexity, and co-occurrence of behavioral difficulties at two points in development. Using a population-level birth cohort, they followed 6,744 individuals as they transitioned from childhood to adolescence. They first identified latent dimensions representing parental ratings of behavioral difficulties and then used these to hierarchically cluster children at each time point. This allowed them to chart developmental changes in the frequency, presentation, and co-occurrence of behavioral difficulties. More importantly, it enabled the mapping of longitudinal transitions between clusters and the identification of environmental factors predicting particular transitions.

The purpose of the present study was to follow the behavioral profiles of children longitudinally from the ages of 5–17. Our aim was to chart not only changes in the overall range and complexity of difficulties but to map longitudinal transitions in profile. We used the Millennium Cohort Study (MCS), a nationally representative U.K. birth cohort of approximately 19,000 individuals, and focused on three pivotal developmental points: early childhood (age 5), late childhood (age 11), and adolescence (age 17). Behavioral difficulties likely change in nature between these ages, with known pronounced changes in physical features (e.g., puberty), brain structure, function (e.g., maturation of frontal lobe with gray matter volumes peaking at 11 years old followed by synaptic pruning and myelination; Blakemore & Choudhury, 2006), cognitive abilities (e.g., language development, problem solving, attention span, risk taking; Best & Miller, 2010; Schoon et al., 2010), social context (e.g., group identity, peer pressure; Blakemore, 2012), and environmental context (e.g., parenting, family responsibility, school setting; Furman & Buhrmester, 1992).

We focused our analysis on the validated derived subscales of the parent-reported Strength and Difficulties (SDQ) questionnaire, that is, conduct, hyperactivity, peer, emotional, and antisocial problems. After reducing the dimensional space using a uniform manifold approximation and projection (UMAP), we used *k*-means clustering to then identify subgroups of children and adolescents with different behavioral profiles within this reduced space. We then characterized the most common cluster transitions from early to late childhood and from late childhood to adolescence. Finally, using a regularized form of gradient-boosting algorithm for classification (XGBoost), we subsequently identified longitudinal and cross-sectional risk factors of these transitions using the preceding data sets from when the children were 9 months old and 3, 5, 7, and 11 years old.

Method

Participants

Data were taken from the MCS, a longitudinal cohort study of young people in the United Kingdom born between September 2000 and January 2002. MCS participants were randomly selected from U.K. child benefit records, and sampling was stratified to ensure adequate representation of the four U.K. countries, ethnic minorities, and highly deprived areas. Data from cohort members (CMs) and their families, including health, mental health, behavioral, socioemotional, education, and employment have been collected since participants were 9 months old. Seven sweeps have been collected at the time of writing, with another sweep planned for 2022. The present study draws on data from three sweeps: Sweep 3 (5 years old; n = 14,059; 50% female), Sweep 5 (11 years old; n = 12,751; 50% female), and Sweep 7 (17 years old; n = 9,191; 50% female; see Table 1 for demographic characteristics of each sweep). We also used data from the 9-month sweep (Sweep 1) and the 3-year-old sweep (Sweep 2) to

Table 1
Demographic Characteristics of MCS3 (5 Years Old), MCS5 (11 Years Old), and MCS7 (17 Years Old) Samples

Characteristic	Sweep 3 (age 5)	Sweep 5 (age 11)	Sweep 7 (age 17)
n (included)	14,059	12,751	9,191
Age, years (*SD*)	5.22 (0.25)	10.67 (0.48)	17.17 (0.33)
Female (%)	7,053 (50)	6,329 (50)	4,605 (50)
Country of birth (%)			
England	8,856 (63)	8,248 (65)	6,164 (67)
Scotland	1,728 (12)	1,442 (11)	957 (10)
Wales	2035 (15)	1,786 (14)	1,249 (14)
Northern Ireland	1,440 (10)	1,274 (10)	821 (9)
Ethnicity (%)			
White	12,072 (86)	10,882 (85)	5,593 (83)
Mixed	397 (3)	110 (0.9)	50 (0.7)
Pakistani and Bangladeshi	669 (5)	820 (6)	310 (5)
Indian	335 (2)	301 (2)	162 (2)
Black African, Caribbean or other	421 (3)	408 (3)	174 (3)
Other Asian	94 (0.7)	123 (1)	72 (1)
Other ethnic group	66 (0.5)	89 (0.7)	44 (0.6)
OECD below 60% median poverty (%)	4,452 (32)	3,040 (24)	NaN

Note. Refer to Supplemental Tables 1 and 2 for analysis of attrition between sweeps. MCS = Millennium Cohort Study; MCS3 = MCS (Sweep 3); MCS5 = MCS5 (Sweep 5); MCS7 = MCS7 (Sweep 7); OECD = Organisation for Economic Co-operation and Development; NaN = not a number.

assess risk factors for behavioral transitions related to circumstances and outcomes of birth and early childhood. All participants included in our analysis had complete responses to the SDQ. We excluded cases with missing information about age and gender. In our risk factor analysis, there were slight variations in the sample size due to small differences in missingness when combining all preceding and concurrent sweeps together. Actual sample sizes will be reported alongside the results.

Measures

Table 2 lists all tests administered at each sweep.

SDQ Behavioral Profiles

Behavioral profiles were obtained from Goodman's SDQ (Goodman, 1997) completed by the main caregiver at ages 5, 11, and 17 years old. The SDQ contains five subscales: Emotional Symptoms, Conduct Problems, Hyperactivity, Peer Problems, and Prosocial Behaviors (reversed to Antisocial Sehavior in this study). Example items from each subscale include: "Often unhappy, downhearted or tearful" (Emotional Symptoms); "Often has temper tantrums or hot tempers" (Conduct Problems); "Restless, overactive, cannot stay still for long" (Hyperactivity); "Generally liked by other children" (Peer Problems); and "Shares readily with other children (treats, toys, pencils etc.)" (Prosocial Behaviors). Scores ranging from 0 to 10 for each of the five subscales are obtained by adding up parental responses from each of its five items, which are rated on a 3-point Likert-type scale (0 = *not true*, 1 = *somewhat true*, or 2 = *certainly true*). After reverse scoring the original Prosocial subscale, higher scores reflect greater difficulties. The validated three-band categorization was used to define two groups of children: (a) presenting with difficulties in any of the five SDQ subscales (part of the "borderline" or "abnormal" bandings), referred to as the

Table 2
List of Tests Administered at Each Sweep

Behavior	Sweep 3 (age 5) SDQ	Sweep 5 (age 11) SDQ	Sweep 7 (age 17) SDQ
Cognition	BAS Naming vocabulary/BAS pattern construction and picture similarity	BAS verbal similarities/CANTAB Cambridge Gambling Task	Number Analogies Test—Cognitive Abilities Test 3
Mental health	Child Social Behavior Questionnaire (independence and self-regulating behavior/emotional dysregulation)	Five-item Rosenberg self-esteem scale/"How do you feel about your life as a whole?"	Five-item Rosenberg self-esteem scale/OCEAN (Neuroticism subscale)/six-item Kessler scale/ seven-item young person Warwick–Edinburgh Mental Wellbeing Scale/"Have you ever been diagnosed with depression and/or serious anxiety?"/"Have you ever attempted to end your life?"

Note. SDQ = Strength and Difficulties Questionnaire; BAS = British Ability Scale; OCEAN = Openness, Conscientiousness, Extraversion, Agreeableness, and Neuroticism.

"elevated" group, having relatively high difficulty scores, and (b) those who did not have any parent-reported difficulties across any scale, referred to as the "non-elevated" group.

Numerous studies have demonstrated the convergent and discriminant validity of the SDQ (van Widenfelt et al., 2003; Vogels et al., 2011). Specifically, when compared to the gold standard Child Behavior Checklist or Youth Self-Report scale(s), the SDQ scales are more strongly associated with its conceptually similar Child Behavior Checklist/Youth Self-Report scale(s) than with conceptually different Child Behavior Checklist/Youth Self-Report scales (for more details, see Vugteveen et al., 2021). The SDQ has also been validated across multiple clinical groups. For instance, for ADHD, the Hyperactivity/Inattention scale has a sensitivity of 91% and a specificity of 90%, and for autism, three scales combine to provide a sensitivity of 79% and specificity of 93% (Russell et al., 2013). However, it is important to recognize that it is not valid for *all* conditions. In a community sample of almost 8,000 5- to 15-year-olds, the SDQ identified individuals with a psychiatric diagnosis with a specificity of 95% and a sensitivity of 63%. However, the questionnaire was markedly better at detecting individuals with conduct, hyperactivity, depressive, and some anxiety disorders, identifying over 70% of them. But this detection rate drops to below 50% for those with specific phobias, separation anxiety, and eating disorders (Goodman et al., 2000).

Cognitive Performance

A detailed guide on the cognitive measures used within the MCS study is available (Moulton et al., 2020). Specifically, direct assessment of cognitive performance and mental health were obtained at the three time points. The earlier sweeps of the MCS consisted predominantly of the British Ability Scale II (BAS II) measures (Elliott et al., 1996).

The cognitive assessment for MCS3 (Sweep 3; 5 years old) consisted of:

1. BAS II, Naming Vocabulary. This is a verbal test which assessed the spoken vocabulary of the child. Test items consist of colored pictures of objects shown one at a time, and the cohort member responds verbally as to what the object is. The test comprises of 36 items (pictures of objectives) in total and has been used to predict delayed language development from 3 years (Law et al., 2012).

2. BAS II, Pattern Construction. This is a nonverbal, spatial, problem-solving test. For each item, a pattern was presented to the cohort member, and the cohort member was asked to replicate the pattern using flat form squats or solid plastic cubes with black and yellow patterns on each side. The test comprises of 23 items and has been used in modeling early-years school attainment (Sullivan et al., 2013).

3. BAS II, Picture Similarity. This is a nonverbal pictorial reasoning test. For each item, the cohort member was shown a row of four pictures or designs, and the cohort member placed a fifth card below the stimulus picture it best matched. The test comprises of 33 items and has been investigated in its moderating role on family and neighborhood risk (Flouri et al., 2015).

The cognitive assessment for MCS5 (Sweep 5; 11 years old) consisted of:

1. BAS II, Verbal Similarities (Elliott et al., 1996). This is a verbal test that assessed the child's verbal reasoning using verbal concepts. The child was given three stimulus words and asked to name the class to which all the examples belong. The test comprises of 35 items.

2. CANTAB Cambridge Gambling Task (Atkinson, 2015). This assessed the child's decision-making and risk-taking behavior. The child would be asked to bet on the number of points to bet on before the decision is made. Intelligence has been associated with risk adjustment and quality of decision making derived from this task (Flouri et al., 2019).

The cognitive assessment for MCS7 (Sweep 7; 17 years old) consisted of:

1. Granada Learning Assessment Number Analogies (Granada Learning Assessment, 2024). This assessed the young person's arithmetic knowledge and reasoning with numbers.

Mental Health

Mental health metrics were also different across age sweeps. At age 5, we used two subscales of the Child Social Behavior Questionnaire completed by the main caregiver measuring the child's independence and self-regulating behavior and the child's emotional dysregulation (Hartman et al., 2006). Items included in these subscales include "Likes to work things out for self" and "Easily frustrated," respectively.

At ages 11 and 17, a shortened five-item Rosenberg Self-Esteem scale was also asked of cohort members (Rosenberg, 1965). Eleven-year-old children's answer to the question "How do you feel about your life as a whole?" was used as a proxy for overall happiness level (1 = *completely happy* to 7 = *not happy at all*).

A range of different mental health measures were added at age 17, including the self-completed Neuroticism subscale of the Big Five personality traits (often listed under the acronym OCEAN: Openness, Conscientiousness, Extraversion, Agreeableness, and Neuroticism), assessing the young person's emotional reactiveness and vulnerability to stress; the self-completed six-item Kessler scale, which measures non-specific psychological distress (Kessler et al., 2003), including items such as "During the last 30 days, about how often did you feel so depressed that nothing could cheer you up?"; the self-completed, short, seven-item Young Person Warwick–Edinburgh Mental Well-being Scale (Stewart-Brown et al., 2009) which provides a single summary score indicating overall well-being, including items such as "I've been feeling optimistic about the future"; the answer to whether cohort members have ever been diagnosed with depression and/or serious anxiety; and the answer to whether they have ever attempted to end their life.

Longitudinal and Cross-Sectional Risk Analysis

Having used the measures above to compare groups after clustering, our second analysis drew on a large number of factors,

looking at longitudinal and cross-sectional risk factors for behavioral transitions between the ages of 5, 11, and 17 years old. We selected these on the basis of the extant literature. We included factors from all MCS sweeps, including MCS1 (Sweep 1; 9 months old) and MCS2 (Sweep 2; 3 years old), preceding or concurrent to the transition "source" sweep. In other words, these factors could originate from any time up and until the "take off" point for the behavioral transition.

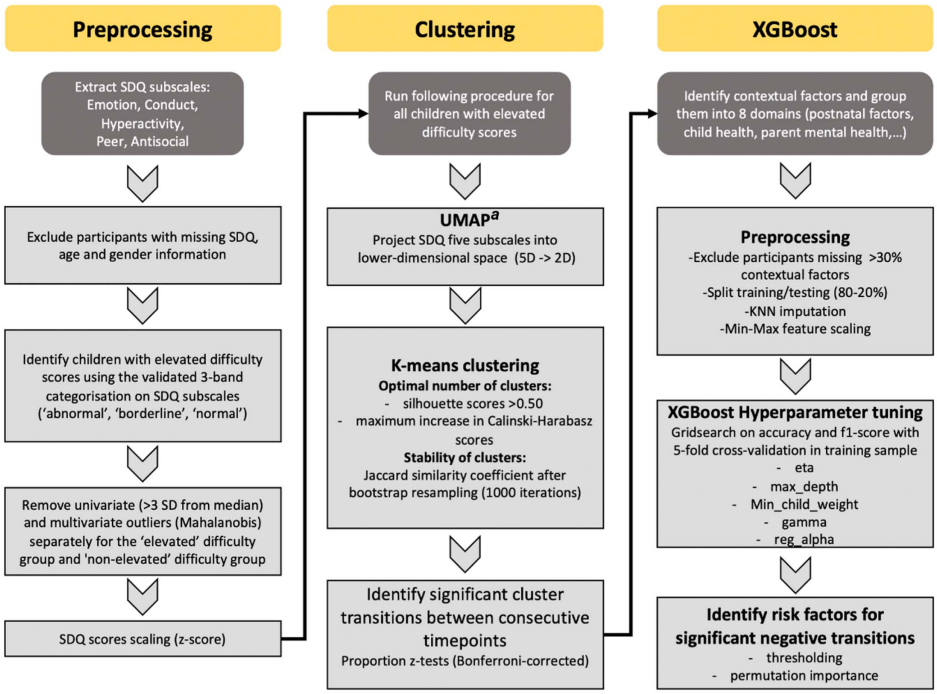
Factors were grouped into nine domains: peri- and postnatal factors (birth weight, type of delivery, feeding problems; Bhutta et al., 2002; McCormick et al., 1990; Serati et al., 2017; Wiles et al., 2006), household factors (e.g., housing tenure, neighborhood safety, average family income; Adjei et al., 2022; Clair, 2019, p. 20; Miller et al., 2021; Minh et al., 2017), child physical health (e.g., BMI, sleep; Griffiths et al., 2011; Heron et al., 2013; Holley et al., 2011; Turnbull et al., 2013), child well-being (e.g., reluctance to go to school, self-esteem; Donnellan et al., 2005; Egger et al., 2003), child cognition (e.g., language development, mathematical abilities; Baird et al., 2022; Quistberg & Mueller, 2020; Thompson et al., 2019), caregiver health and mental health (e.g., Kessler scale, use of recreational drugs, alcohol consumption; Adjei et al., 2022; Maselko et al., 2015; Meadows et al., 2007; Schepman et al., 2011), caregiver relations (e.g., marital status, satisfaction with partner; Giannakopoulos et al., 2009), child relations (e.g., time spent outside with friends, time spent on social media; Brunborg & Burdzovic Andreas, 2019; Van Roy et al., 2010), and caregiver–child relationship (e.g., style of parenting, closeness to child; Bloomfield & Kendall, 2012; Coldwell et al., 2006). See Supplemental Table 3 for the full list of factors.

Analysis Procedure

Our analysis pipeline was designed first to clean the data and partition it into those with elevated and non-elevated difficulty scores on the SDQ (note that elevated here includes children with borderline or elevated scores on at least one scale). Then, with the elevated group, we conducted a dimension reduction approach called UMAP, before clustering, identifying significant transitions across development, and then establishing significant risk factors for these transitions (Figure 1; Table 3). Readers may wonder why we partitioned the data to focus on those with elevated scores. This is necessary because otherwise the dimensional space is dominated by individuals with scores within the normal range and not optimized to capture behavioral difficulties. In turn, the data present lower

Figure 1

Flow Diagram Summarizing the Analysis Pipeline



Note. XGBoost = gradient-boosting algorithm for classification; SDQ = Strength and Difficulties Questionnaire; UMAP = uniform manifold approximation and projection; KNN = *k*-nearest neighbor; max_depth = maximum depth (of our tree); min_child_weight = minimum child weight (minimum sum of instance weight needed for a child of our tree); reg_alpha = regularisation alpha (L1 regularisation term on weights). See the online article for the color version of this figure.

[a] In the UMAP projection of SDQ, data from five dimensions (5D) to two dimensions (2D) reduce the size of the data but in a way that we do not lose out an essential information underlying the data. Subsequent figures present the 2D plot of each subject's reduced 2D data.

Table 3
Summarized Information of Key Steps Within the Analysis Pipeline

Step	Further detail
1. Clean and partition SDQ data	Create two groups of participants, identifying those with elevated parent-reported difficulties. Ensure a complete and high-quality data set for subsequent analysis.
2. UMAP	Reduce the dimensional space for subsequent clustering to maximize clustering effectiveness while retaining data complexity.
3. *k*-means clustering	Identify data-driven subgroups within the elevated sample, at each time point independently.
4. Proportion *z* tests	Identify the numbers of children who move between cluster over time, indicating a shift in their specific profile of difficulties with age. Identify significant behavioral transitions (higher proportion than predicted by chance).
5. Identify and clean contextual factors	Identify longitudinal and cross-sectional risk factors for behavioral transitions based on the extant literature and group them into domains. Ensure a high-quality data set.
6. XGBoost	Use a model that allows the use of multicollinear variables. Allow generalizability of model and avoid overfitting. Identify longitudinal and concurrent predictors of significant behavioral transitions.

Note. SDQ = Strength and Difficulties Questionnaire; UMAP = uniform manifold approximation and projection; XGBoost = Gradient-boosting algorithm for classification.

dimensionality because it is dominated by the extreme differences in overall severity across the whole sample and is then very difficult to cluster meaningfully.

Clustering

The first step of our analysis was to identify subgroups of children within the elevated sample at each time point by using a clustering method on *z*-scored SDQ derived values. Before clustering, we first excluded any participants with missing SDQ data or missing age or gender information. We then partitioned our children into the elevated and non-elevated difficulty groups. We then removed univariate (>3 *SD*s from the median) and multivariate outliers (Mahalanobis distance >α quantile of the chi-square distribution with 5 *df*) separately for each group. The total number of children excluded at each time point is indicated in the consort diagrams in Supplemental Figures 1–3 (MCS 3: n = 702 or 4.7%, MCS 5: n = 60 or 0.5%, MCS 7: n = 61 or 0.7%, for univariate or multivariate outliers or because of missing information).

To avoid clustering in a sparse dimensional space and to improve clustering performance (Dalmaijer et al., 2021), we then projected the five SDQ subscales into a lower dimensional space using UMAP. This method is based on Riemannian geometry and algebraic topology and has been show to outperform Principle Component Analysis, t-distributed Stochastic Neighbor Embedding, or Multidimensional Scaling methods for (a) its ability to retain higher order, nonlinear interaction of variables inherent in the type of data we are studying; (b) its preservation of both local and global structure of the original data; (c) its greater flexibility with several distance metrics to choose from; (d) its computational efficiency (O[N] vs. O[Nlog(N)]); and (e) more reproducible results (Yang et al., 2021, p. 20).

Briefly, UMAP first constructs a topological representation of the original data with local manifold approximations, and then it optimizes the lower dimension embedding by minimizing the cross-entropy between the new lower dimensional space and the original higher dimensional space. We set UMAP parameters with a relatively high number of neighbors (*n*_neighbors = 50) to decrease the probability of producing fine-grained cluster structure that is more prone to be a result of patterns of noise. We set a low minimum distance (min_dist = 0.01) to make denser clusters and cleaner separations. We also used a correlation distance metric. UMAP was implemented using the UMAP V.0.3.2 Python implementation (McInnes et al., 2020).

We then used *k*-means clustering on the UMAP data in the lower dimensional space using the scikit-learn implementation in Python. Optimal number of clusters were chosen based on silhouette coefficients exceeding 0.5 and a steep increase in the Calinski–Harabasz index, both indicating a good separation between clusters scores (Bathelt et al., 2021; Dalmaijer et al., 2021), and Jaccard similarity coefficients between clusters obtained after bootstrap resampling (1,000 iterations) and the original clusters (Zumel & Mount, 2014) exceeding 0.85, indicating a good stability of clusters.

The non-elevated portion of the data was included as a comparison group for each age point. Comparisons between clusters for gender, cognition, mental health, and neurodivergence diagnosis were conducted with chi-square tests and one-way analysis of variance with Bonferroni post hoc *t* tests. All clusters (including the comparison non-elevated group) were compared to the whole sample at each age point.

Assessment of Transitions

Proportion *z* tests (Bonferroni-corrected for multiple comparisons) were used to compare the proportion of participants transitioning from one cluster at one time point to another cluster at the next time point with an equal-split transition (e.g., in the case of five clusters, an equal split would be 20% of the sample transitioning into each cluster). Only significant transitions above the equal-split proportion, that is, transitions that occurred more than we would expect by chance, will be reported here. To make sure transitions were not driven by higher Euclidean distance from the center of each identified cluster, we compared the distribution of silhouette scores for each participant who was part of a significant transition to the distribution of silhouette scores for the entire cluster (see Supplemental Figures 4 and 5). In other words, borderline cases could switch between clusters over time not because of any meaningful change in behavior but because of imprecision in the

clustering solution. Comparing the silhouette scores in this way controls for this.

Identification of Longitudinal Risk Factors for Significant Transitions

Longitudinal and concurrent predictors of significant behavioral transitions between the ages of 5–11 years old and between the ages of 11–17 years old were identified looking at the feature importance gain after performing XGBoost on our train samples (80% of our data set) using the XGBoost Python package. For each significant transition, we compared the subgroup of children with a particularly elevated difficulty profile transitioning to another elevated difficulty profile cluster at the following time point ("transition of interest") to the subgroup of children with the same elevated difficulty profile transitioning to the non-elevated difficulty scores group at the following time point ("control transition"). In essence, we were looking to identify factors that distinguish those who transition to a different profile of behavioral difficulty from those who transition to no apparent difficulties, given the same starting profile. XGBoost is an ensemble algorithm based on gradient-boosted trees. It integrates the predictions of "weak" classifiers to achieve a "strong" classifier (tree model) via a serial training process. It incorporates an L1 (Least Absolute Shrinkage and Selection Operator) regression regularization, which prevents the model from overfitting and is robust to multicollinearity of features.

We first removed all participants missing more than 30% of the features (Uh et al., 2021; 282 and 178 participants were excluded for the "5–11" cohort and the "11–17" cohort, respectively), and after splitting our data set into training and testing sets (80%–20% as we prioritized low variance of our parameter estimates rather than our performance statistics), we imputed missing data using *k*-nearest neighbor imputation and scaled features using Min–Max method (3.35% and 3.30% of data were imputed for the 5–11 cohort and the 11–17 cohort, respectively). We then tuned the following hyperparameters optimizing both accuracy and *f*1-score using a GridSearch procedure with 5-fold cross validation on our training sample: nrounds (the number of sequentially built trees in the model), η (the learning rate), max_depth (the maximum levels deep that each tree can be grown), min_child_weight (the minimum degree of impurity needed in a node before a split is attempted), γ (the minimum amount of splitting by which a node must improve the predictions), and reg_alpha (the L1 regularization term). We also set the scale_pos_weight hyperparameter with the inverse of the class distribution to tune the behavior of the algorithm for imbalanced classification problems. We then identified the optimal feature importance score threshold and number of features, which led to the best model performance accuracy and *f*1-score for predicting the transition of interest in our test sample. Permutation importance was finally implemented on the test data by sequentially replacing each feature with its random shuffling and measuring at which extent permutation of a particular feature affected the accuracy of predictions. Only the features that survived both thresholding and permutation importance were considered genuine risk factors for a significant behavioral transition. We chose to report feature importance gain in this article, which corresponds to the improvement in accuracy brought by a feature to the branches it is on.

The data and code that support the findings of this study are available from the corresponding author upon reasonable request.

Transparency and Openness

We report all data exclusions, all data transformations, and all measures in the study, and we follow the Journal Article Reporting Standards (Appelbaum et al., 2018). All analysis code is available at https://github.com/eliacello/MCS_BehaviouralTransitions_2021/tree/main. The code in this repository was used to clean, preprocess, cluster, and model data from the Millennium Cohort Study, known as "Child of the New Century," and freely accessible through the UK Data Service (University College London, UCL Social Research Institute, Centre for Longitudinal Studies, 2024). The Millennium Cohort Study data are not included in the repository. Data were analyzed using R, Version 4.0.0 (R Core Team, 2020), and the packages EFA.dimensions Version 0.1.8.4 (https://CRAN.R-project.org/package=EFA.dimensions), ggplot Version 3.5.1 (Wickham, 2016), polycor Version 0.8-1 (https://CRAN.R-project.org/package=polycor), and ez Version 4.4-0 (https://CRAN.R-project.org/package=ez). Data were also analysed using Python, Version 3.10, and the libraries numpy, pandas, matplotlib, seabird, sklearn (preprocessing, metrics, cluster), scipy, umap, and statistics. This study's design and its analysis were not preregistered.

Results

Our dimensionality reduction and clustering analysis identified three clusters at age 5, five clusters at age 11, and three clusters at age 17. For these cluster solutions, silhouette coefficients all exceeded 0.5, we observed a steep increase in the Calinski–Harabasz index, and the Jaccard similarity coefficients all exceeded 0.85 after bootstrapping. See Figures 2–4 for an illustration of the behavioral profiles and descriptive statistics for each age group, respectively.

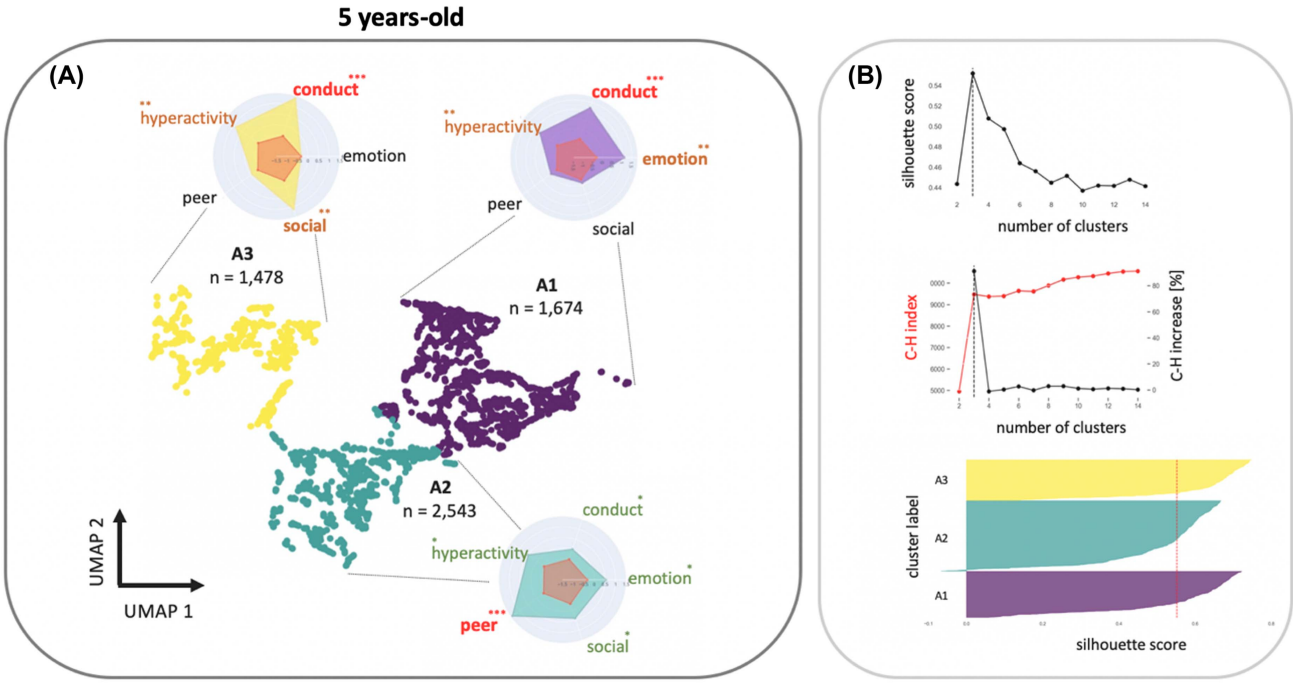
Cluster Comparisons

Early Childhood

For the early childhood sweep (age 5), the first cluster (A1, n = 1,674) scored highly on both the Conduct Problems and Emotional Problems subscales; the second cluster (A2, n = 2,543) was principally defined by high antisocial behavior scores alone, and the third cluster (A3, n = 1,478) scored highly for conduct problems and antisocial behaviors. There were more girls in the non-elevated group (identified before clustering; A4, n = 8,364) and more boys in A2 and A3 compared to the whole sample gender distribution. The clusters showed differences in cognition and emotion regulation abilities, one-way analysis of variance: BAS naming: $F(3, 14055) = 121.8$, $p < .001$; BAS similarities: $F(3, 14028) = 48.98$, $p < .001$; BAS pattern: $F(3, 14029) = 92.02$, $p < .001$; Children's Social Behavior Questionnaire (CSBQ) independence: $F(3, 14055) = 356.3$, $p < .001$; CSBQ emotional dysregulation: $F(3, 14054) = 1{,}652$, $p < .001$). Post hoc tests revealed that all groups significantly differed from the whole sample ($p < .05$, Bonferroni corrected): Children with non-elevated difficulty scores had higher cognitive and CSBQ scores than the whole sample whereas children with behavioral problems had significantly reduced cognitive and

Figure 2

Results of the k-Means Clustering of the Early Childhood Data



		A1 conduct-emotion	A2 peer	A3 conduct-antisocial	A4 no difficulties
Demographics	Size	1674	2543	1478	8364
	% (total sweep sample)	12	18	11	59
Cognition	% Female	50	46 *	38 ***	53 ***
	BAS naming vocabulary	51.79 (11.76) ***	51.53 (12.14) ***	52.64 (11.12) ***	55.56 (10.98) ***
	BAS Picture Similarity	54.02 (10.89) ***	54.13 (11.51) ***	54.01 (10.71) ***	56.30 (10.43) ***
	BAS Pattern Construction	48.57 (10.71) ***	48.17 (11.38) ***	48.63 (10.95) ***	51.34 (9.75) ***
Mental health	Independence Self-regulation	2.47 (0.34) *	2.43 (0.36) ***	2.40 (0.35) ***	2.61 (0.30) ***
	Emotional dysregulation	2.06 (0.41) ***	1.86 (0.42) ***	2.07 (0.41) ***	1.53 (0.36) ***
Neurodivergent Diagnosis	ADHD	23 (1.4%)**	21 (0.8%)	19 (1.3%)***	11 (0.1%)***
	Autism/Asperger's	4 (0.2%)	39 (1.5%)***	6 (0.4%)	15 (0.2%)

Note. (A) The reduced 2D UMAP space on which each participant Strength and Difficulties Questionnaire-derived variable were projected. Clusters are coded with colors with radar plots indicating the profiles of behavioral ratings for each cluster. (B) The average silhouette for different number of clusters (top), the Calinski–Harabasz index for different cluster numbers (middle), and the silhouette coefficient for each group (bottom). The table gives an overview of size, gender ratio, cognitive assessments, incidence of mental health problems, and neurodivergent diagnosis for each identified cluster. Green indicates the values above the whole sample average (or below for "negative" domains such as emotional dysregulation). Red indicates the values below the whole sample average (or above for "negative" domains such as emotional dysregulation). UMAP = uniform manifold approximation and projection; BAS = British Ability Scale; ADHD = attention-deficit/hyperactivity disorder. See the online article for the color version of this figure.
The asterisks in (A) indicate Cohen's d relative to $\mu = 0$ and the standard deviation across all groups: $^{*}d > 0.5$ (green). $^{**}d > 1$ (amber). $^{***}d > 1.5$ (red). Asterisks in the table indicate the statistical comparison to the whole Bonferroni-corrected sample: $^{*}p < .05$. $^{**}p < .01$. $^{***}p < .001$.

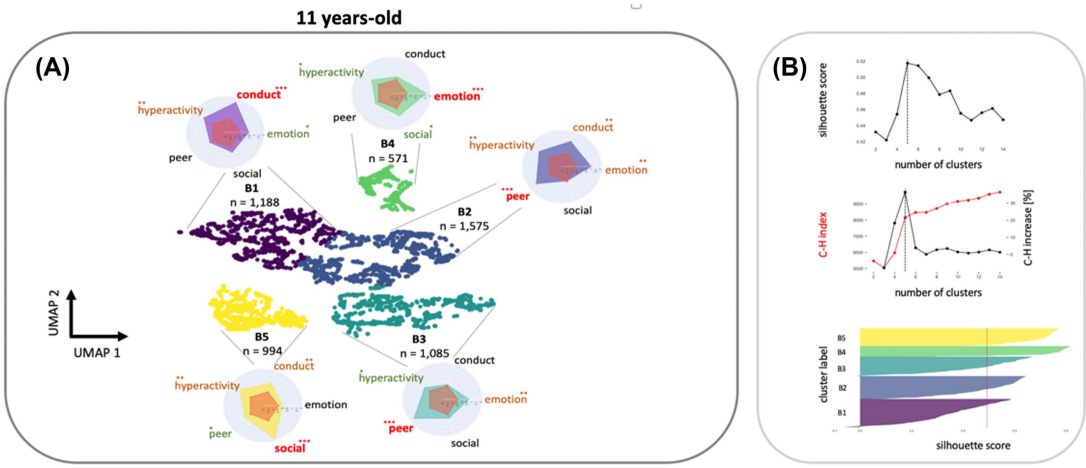
CSBQ scores. The prevalence of neurodevelopmental diagnoses was also different across groups with significantly more ADHD children diagnosed in clusters A1 and A3 presenting with conduct problems, fewer children diagnosed with ADHD in A4 presenting with non-elevated difficulty scores (A1: $\chi^2 = 16.2$, $p < .001$; A2: $\chi^2 = 11.7$, $p < .001$; A4: $\chi^2 = 20.6$, $p < .001$), and more children diagnosed with autism in cluster A2 ($\chi^2 = 38.8$, $p < .001$) presenting with peer relationship difficulties.

Late Childhood

For the late childhood sweep (age 11), the first cluster (B1, n = 1,330) scored highly for conduct problems mainly; the second cluster (B2, $n = 1{,}463$) was defined by high scores on a whole range of subscales mainly involving peer relationships problems, and at a lesser extent, hyperactivity, conduct, and emotional problems; the third cluster (B3, $n = 1{,}085$) scored highly for internalized problems

Figure 3

Results of the k-Means Clustering of the Late Childhood Data



		B1 conduct	B2 complex peer	B3 peer	B4 emotion	B5 antisocial	B6 no difficulties
Demographics	Size	1188	1575	1085	571	994	7339
	% (total sweep sample)	9	12	9	4	8	58
	% Female	47	50	49	51	31 ***	52 ***
Cognition	BAS verbal similarities standardised	56.37 (10.40) ***	55.87 (10.80) ***	58.15 (10.71)	57.14 (10.29) *	57.21 (10.44) ***	60.16 (9.19) ***
	CANTAB – Cambridge Gambling Task	0.55 (0.17) ***	0.54 (0.17)	0.53 (0.17)	0.53 (0.18)	0.56 (0.16) ***	0.52 (0.16) *
Mental health	Life satisfaction	5.77 (1.55) ***	5.76 (1.60) ***	5.88 (1.44) ***	5.90 (1.48) *	5.87 (1.58) ***	6.24 (1.23) ***
	Self-esteem	2.26 (0.60) *	2.22 (0.62) ***	2.24 (0.64) ***	2.25 (0.57) *	2.27 (0.60)	2.37 (0.56) ***
Neurodivergent Diagnosis	ADHD	31 (2.6%)***	46 (2.9%)***	20 (1.8%)	7 (1.2%)	23 (2.3%)***	12 (0.2%)***
	Autism/Asperger's	20 (1.7%)	53 (3.4%) ***	46 (4.2%) ***	21 (3.7%) ***	24 (2.4%)	16 (0.2%)***
	Dyslexia	49 (4.1%)*	69 (4.4%)***	37 (3.4%)	29 (5.1%) ***	19 (1.9%)	118 (1.6%) ***
	Dyspraxia and/or dyscalculia	32 (2.7%)	76 (4.8%) ***	37 (3.4%) *	22 (3.9%) *	37 (3.7%) ***	48 (0.7%) ***

Note. (A) The reduced 2D UMAP space on which each participant Strength and Difficulties Questionnaire-derived variable were projected. Clusters are coded with colors with radar plots indicating the profiles of behavioral ratings for each cluster. (B) The average silhouette for different number of clusters (top), the Calinski–Harabasz index for different cluster numbers (middle), and the silhouette coefficient for each group (bottom). The table gives an overview of size, gender ratio, cognitive assessments, incidence of mental health problems, and neurodivergent diagnosis for each identified cluster. Green indicates the values above the whole sample average (or below for "negative" domains such as emotional dysregulation). Red indicates the values below the whole sample average (or above for "negative" domains such as emotional dysregulation). UMAP = uniform manifold approximation and projection; BAS = British Ability Scale; ADHD = attention-deficit/hyperactivity disorder. See the online article for the color version of this figure.
The asterisks in (A) indicate Cohen's d relative to $\mu = 0$ and the standard deviation across all groups: $^{*}d > 0.5$ (green). $^{**}d > 1$ (amber). $^{***}d > 1.5$ (red). Asterisks in the table indicate the statistical comparison to the whole Bonferroni-corrected sample: $^{*}p < .05$. $^{**}p < .01$. $^{***}p < .001$.

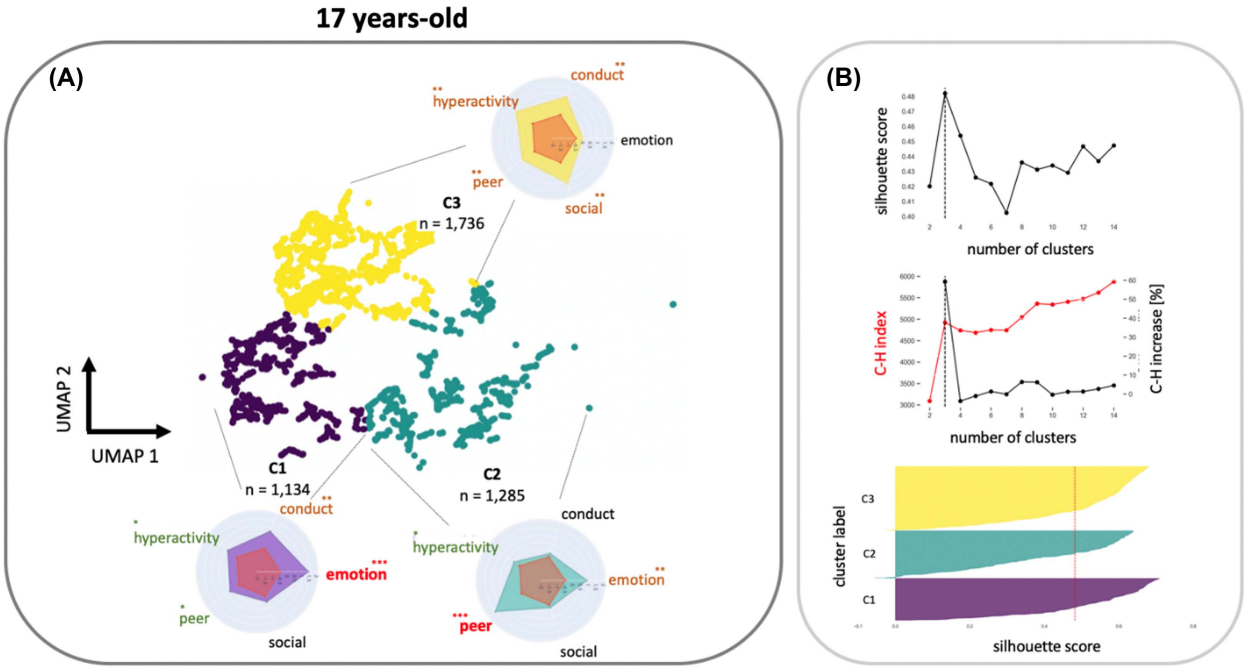
mainly weighted toward peer problems (and less toward emotional dysregulation); the fourth cluster (B4, $n = 539$) had high emotional problems scores; and the last cluster (B5, $n = 996$) scored highly for antisocial behaviors mainly. There were more girls in the non-elevated group without behavioral problems (B6, $n = 7{,}339$) and more boys in the group with antisocial behavior compared to the whole sample gender distribution. The clusters also showed differences in cognition, verbal similarities: $F(5, 12491) = 78.53$, $p < .001$; Cambridge Gambling Task risk taking: $F(5, 12050) = 13.5$, $p < .001$, and Bonferroni-corrected post hoc t tests indicated that children with conduct problems (cluster B1) scored lower at the verbal task and were more prone for risk-taking behaviors compared to the whole sample ($p < .001$). Children with complex peer problems (cluster B2) had also significantly reduced verbal scores ($p < .05$). Mental health also differed between groups, satisfaction with life: $F(5, 12328) = 53.08$, $p < .001$; self-esteem: $F(5, 12123) = 30.26$, $p < .001$, with all groups with behavioral problems (clusters B1–B4), except the one displaying antisocial behaviors, showing reduced overall satisfaction with life and self-esteem; children with non-elevated difficulty scores (B6) increased satisfaction with life and self-esteem compared to the whole sample ($p_{\text{Bonferroni}} < 0.05$). Finally, the proportion of children diagnosed with neurodevelopmental disorders differed between clusters (ADHD: $\chi^2 = 20.6$, $p < .001$; autism: $\chi^2 = 20.6$, $p < .001$; dyslexia: $\chi^2 = 20.6$, $p < .001$; dyspraxia: $\chi^2 = 20.6$, $p < .001$).

Adolescence

For the adolescence sweep (age 17), the first cluster (C1, $n = 1{,}134$) scored highly on the Emotional subscale and, to a lesser extent, on the Conduct subscale; the second cluster (C2, $n = 1{,}285$) indicated high scores for peer problems; and the third cluster (C3, $n = 1{,}736$) scored highly for antisocial problems and, to a lesser extent, peer relationship, hyperactivity, and conduct problems. There were more girls in the group with internalizing problems (C1 and C2) and more boys in the group with complex antisocial

Figure 4

Results of the k-Means Clustering of the Adolescence Data



		C1 emotion-conduct	C2 internalised	C3 antisocial	C4 no difficulties
Demographics	Size	1134	1285	1736	5036
	% (total sweep sample)	12	14	19	55
	% Female	70 ***	56 ***	34 ***	49
Cognition	Number Analogies Test /10	4.59 (2.62) ***	4.81 (2.66) ***	4.87 (2.73) ***	5.60 (2.66) ***
Mental health	WEMBS /35	20.97 (4.02) ***	21.41 (4.03) ***	22.25 (4.22)	23.21 (3.90) ***
	Self-esteem	1.63 (0.82) ***	1.70 (0.83) ***	1.91 (0.78)	2.00 (0.75) ***
	OCEAN neuroticism /21	14.19 (5.08) ***	13.56 (4.88) ***	10.86 (4.64) ***	11.15 (4.68) ***
	Kessler /24	9.66 (5.41) ***	8.87 (5.49) ***	7.04 (4.72)	6.30 (4.29) ***
	Diagnosis depression/anxiety	24.85% ***	21.14% ***	7.49% *	4.58% ***
	At least one suicide attempt	15.91% ***	11.53% ***	6.58%	3.96% ***

Note. (A) The reduced 2D UMAP space on which each participant Strength and Difficulties Questionnaire-derived variable were projected. Clusters are coded with colors with radar plots indicating the profiles of behavioral ratings for each cluster. (B) The average silhouette for different number of clusters (top), the Calinski–Harabasz index for different cluster numbers (middle), and the silhouette coefficient for each group (bottom). The table gives an overview of size, gender ratio, cognitive assessments, incidence of mental health problems, and neurodivergent diagnosis for each identified cluster. Green indicates the values above the whole sample average (or below for "negative" domains such as emotional dysregulation). Red indicates the values below the whole sample average (or above for "negative" domains such as emotional dysregulation). UMAP = uniform manifold approximation and projection; BAS = British Ability Scale; ADHD = attention-deficit/hyperactivity disorder. See the online article for the color version of this figure.
The asterisks in (A) indicate Cohen's d relative to $\mu = 0$ and the standard deviation across all groups: $^{*}d > 0.5$ (green). $^{**}d > 1$ (amber). $^{***}d > 1.5$ (red). Asterisks in the table indicate the statistical comparison to the whole Bonferroni-corrected sample: $^{*}p < .05$. $^{**}p < .01$. $^{***}p < .001$.

behavior (C3) compared to the whole sample gender distribution. The clusters showed differences in arithmetic abilities, $F(3, 8423) = 68.27$, $p_{\text{Bonferroni}} < 0.001$, and Bonferroni-corrected post hoc t tests indicated that adolescents with non-elevated difficulty scores (C4) had higher cognitive scores than the whole sample whereas children with behavioral problems had significantly reduced cognitive scores ($p < .001$).The groups showed overall difference for several measures of mental health, Warwick–Edinburgh Mental Wellbeing Scale: $F(3, 8645) = 134.4$, $p < .001$; Kessler scale: $F(3, 8686) = 209.6$, $p < .001$; Openness, Conscientiousness, Extraversion, Agreeableness, and Neuroticism scale: $F(3, 8687) = 195.6$, $p < .001$; self-esteem: $F(3, 8681) = 94.33$, $p < .001$; diagnosis of depression and/or anxiety: $\chi^2 = 635.2$, $p < .001$; proportion of adolescents having attempted suicide: $\chi^2 = 250.66$, $p < .001$. As for the preceding sweep, clusters with behavioral problems (clusters C1 and C2), except the one displaying complex antisocial behaviors, showed lower mental health on all

scales while children in the non-elevated group (C4) showed higher mental health compared to the whole sample ($p_{\text{Bonferroni}} < 0.05$).

Transitions

The validity of our clusters was also reflected in the patterns seen in the behavioral transitions across childhood and adolescence. Figure 5 describes the proportion of participants transitioning from one cluster to another across time. For all elevated clusters found at ages 5 and 11, a significant proportion transitioned to the non-elevated group. At age 5, a larger-than-expected proportion of children (23%) transitioned from having conduct and emotional problems (cluster A1) to having problems forming peer relationships co-occurring with hyperactivity, conduct, and emotional problems (cluster B2) at age 11. A significant proportion of children aged 5 (21%) also transitioned from having conduct problems combined with antisocial behaviors (cluster A3) to having conduct problems alone (cluster B1).

At age 11, a larger-than-expected 46% of children displaying antisocial behaviors (cluster B4) and 32% of children with conduct problems (cluster B1) transitioned to a group of adolescents having antisocial problems co-occurring with peer relationship, hyperactivity, and conduct problems (cluster C3). A larger-than-expected 33% of children having difficulties forming peer relationships transitioned to a group of adolescents with the same peer relationship problems co-occurring with emotional problems (cluster C2).

The comparison of silhouette scores distribution between children from a "significant" transition (e.g., 23% going from A1 to B2) and the rest of the cluster (e.g., A1) can be found in the Supplemental Figures 4 and 5.

Longitudinal and Cross-Sectional Risk Factors

We next applied the classification model XGBoost to determine the predictors for the significant behavioral transitions listed in the previous section, contrasting transition of interest (elevated difficulty cluster to elevated difficulty cluster) to its corresponding control transition (elevated difficulty cluster to non-elevated cluster). The strength of features surviving thresholding and permutation in cross-validated data sets corresponds to the importance gain (g), which corresponds to the improvement in accuracy brought by a feature to the branches of the decision tree it is on.

We summarize our findings in Table 4 and Figure 6. At age 5, the strongest predictors for transitioning from having conduct and emotional problems (cluster A1) to having problems forming peer relationships co-occurring with hyperactivity, conduct, and emotional problems (cluster B2) at age 11 were having long-standing illness ($g = 9.00$), a low Organisation for Economic Co-operation and Development family income ($g = 4.11$), an agitated atmosphere at home ($g = 3.08$), and a lower cognitive score for pattern construction ($g = 2.78$). Conversely, the persistence of conduct problems between 5 and 11 years old was best predicted by a set of factors measured at age 3: having caregivers reporting a difficult relationship with their child ($g = 17.64$), irregular bedtimes ($g = 6.30$), a high BMI ($g = 6.60$), caregiver's high psychological distress ($g = 7.78$), caregiver's low satisfaction with home ($g = 5.61$), and caregiver considering the neighborhood a disadvantaged area to raise kids ($g = 4.20$). At age 11, the persistence of antisocial behaviors with the addition of new difficulties with peer relationships, hyperactivity, and conduct problems was best predicted by caregiver's mental health measured when the child was 3, 7, and 11 years old: caregivers self-reporting lower extraversion ($g = 13.61$), a higher emotional reactiveness and vulnerability to stress ($g = 10.68$), a lower overall life satisfaction ($g = 10.25$; $g = 9.98$), and a lower satisfaction with work/family balance ($g = 9.87$). The transition from having conduct problems at age 11 to having a combination of conduct, hyperactivity, peer relationship difficulties, and antisocial behaviors at age 17 was best predicted by the child's mental health mainly linked with school at ages 7 and 11 and reading abilities at age 7: self-reported feeling unhappy about school work ($g = 10.61$), often getting fed up at school ($g = 6.54$), often feeling scared ($g = 8.54$), and lower age-standardized word reading scores ($g = 9.19$). Finally, the persistence of internalized problems (difficulties with peer relationships and emotional dysregulation) across adolescence was best predicted by the main caregiver's diagnosis of depression at least once in their lifetime ($g = 1.29$), the main caregiver feeling unhappy with their relationship with partner ($g = 0.28$; reported when the child was 9 months old), the caregiver reporting not being close to their child at age 7 ($g = 0.19$), and the child considering school a waste of time ($g = 0.56$) at age 11.

Discussion

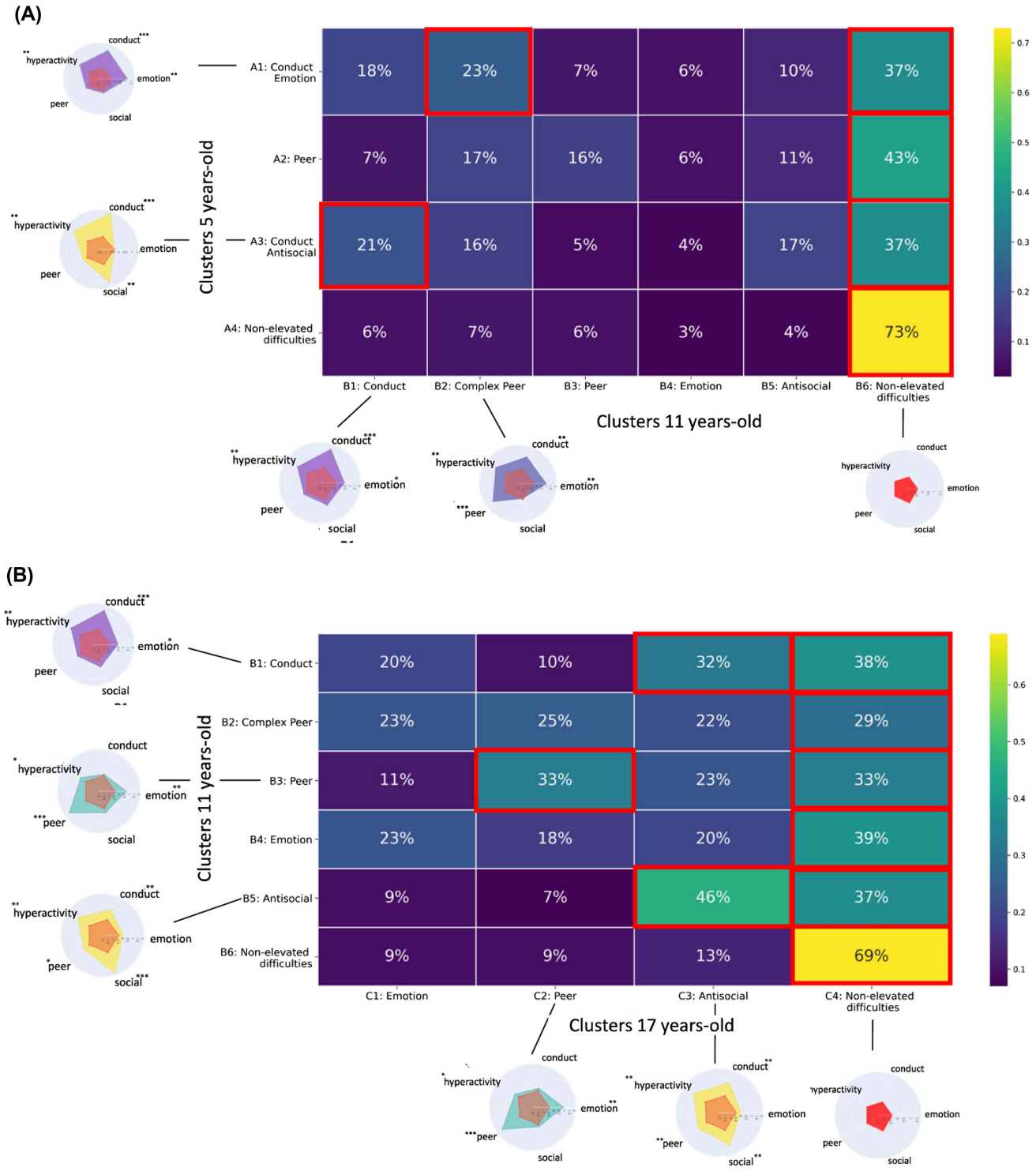
Behavioral difficulties change radically in their range, scope, and complexity across development, something poorly captured by rigid diagnostic frameworks. By integrating a dimension reduction and clustering paradigm with a subset of the population-representative MCS, we found distinct subgroups of children based on parental reports of emotional, conduct, hyperactivity, peer, and social difficulties. The number of clusters and their respective profiles varied across developmental time. We identified several common cluster-to-cluster transitions between early childhood and late childhood and between late childhood and adolescence. Most children transitioned to the non-elevated difficulty scores group, but a significant minority of children were at elevated risk of persistent difficulties or of developing new difficulties over time. Our subsequent analysis of longitudinal risk revealed that a diverse range of socioeconomic, health problems, poor caregiver mental health, and challenging household factors predicted transitions during childhood. Poor mental health of both caregiver and adolescents accompanied with difficulty in child–parent and parent–parent relationships explained the most common behavioral transitions during adolescence.

Gender Gap

There is a shift in gender distribution across our clusters between childhood and adolescence: While there were systematically more girls in the portion of the sample with non-elevated behavioral problems at ages 5 and 11, there was an equal gender representation by 17 years old. Furthermore, there was an overrepresentation of girls in clusters with internalized problems at age 17 (70% female in the group with emotional problems and 56% female in the group with peer problems). Conversely, there were systematically more boys in the groups displaying conduct and antisocial behaviors throughout childhood and adolescence. In other words, the proportion of females being members of the "elevated scores"

Figure 5

Overview of Transitions Showing the Percentage of Children Transitioning Between Groups



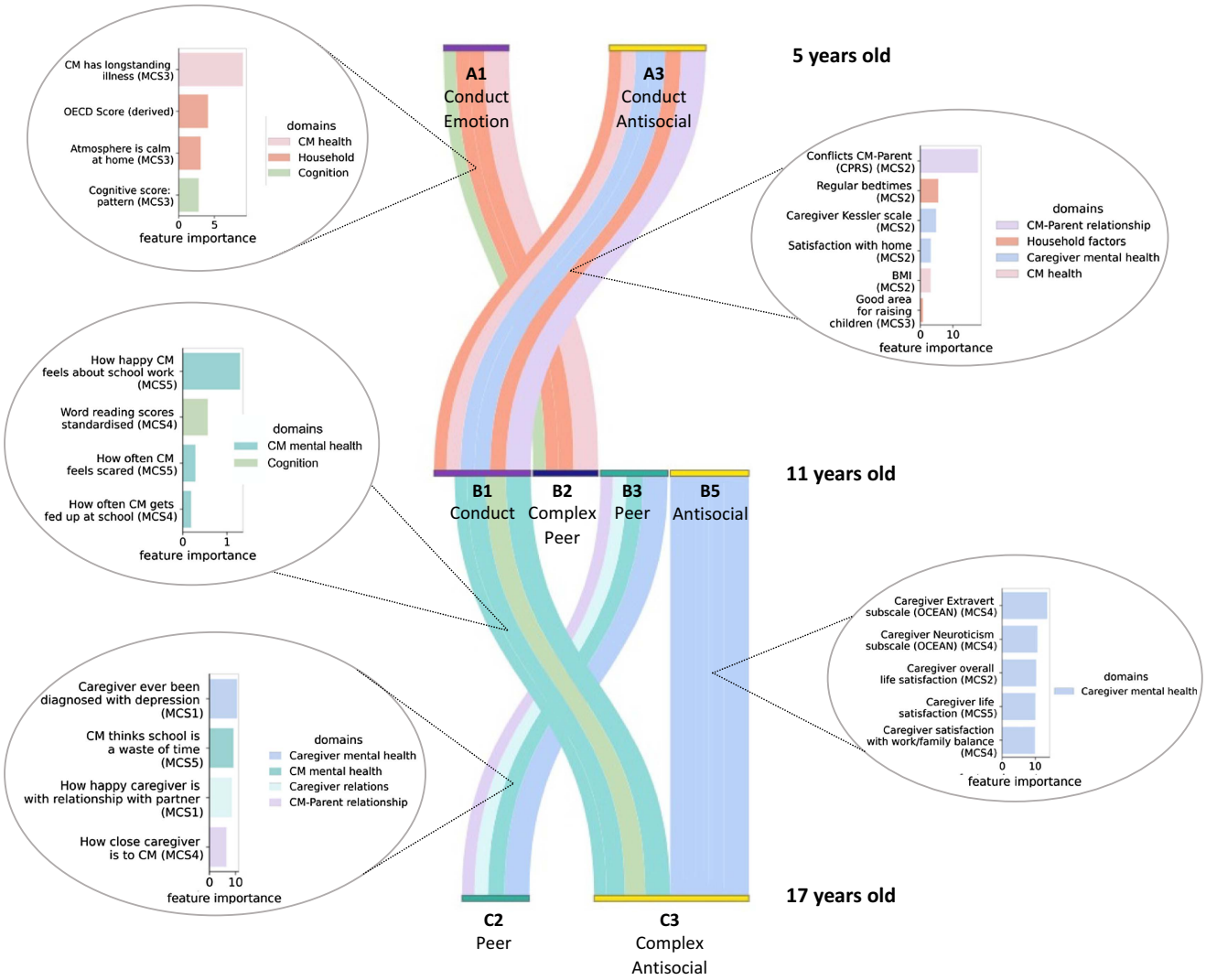
Note. These groups were identified in (A) early childhood (5 years old; rows) to groups identified in late childhood (11 years old; columns) and from groups identified in (B) late childhood (rows) to groups in adolescence (17 years old; columns). Cells with red rectangles indicate proportions that were significantly above what was expected by chance (Bonferroni-corrected). See the online article for the color version of this figure.

Table 4
Summary of Longitudinal and Cross-Sectional Risk Factors Linking to Transitions Over Time

Age transition	Transition of interest (%)	Control transition (%)	Predictor (age of child when reported; importance gain)
5–11	Conduct–emotion (A1) to complex peer (B2; 23%)	Conduct–emotion (A1) to nonelevated (B6; 37%)	Cohort member has longstanding illness (5; $g = 9.00$) Low Organisation for Economic Co-operation and Development family income (5; $g = 4.11$) Agitated atmosphere at home (5; $g = 3.08$) Lower cognitive score for pattern construction (5; $g = 2.78$)
	Conduct–antisocial (A2) to conduct (B1; 21%)	Conduct–antisocial (A2) to nonelevated (B6; 37%)	Difficulty caregiver–child relationship (3; $g = 17.64$) Irregular bedtimes (3; $g = 6.30$) High child BMI (3; $g = 6.60$) High Caregiver Kessler (K6) score (3; $g = 7.78$) Caregiver's low satisfaction with home (3; $g = 5.61$) Disadvantaged area to raise kids (5; $g = 4.20$)
11–17	Conduct (B1) to antisocial (C3; 32%)	Conduct (B1) to nonelevated (C4; 38%)	Feeling unhappy about school work (11; $g = 10.61$) Lower age-standardized word reading scores (7; $g = 9.19$) Often feeling scared (11; $g = 8{,}54$) Often getting fed up at school (7; $g = 6.54$)
	Peer (B3) to peer (C2; 33%)	Peer (B3) to nonelevated (C4; 33%)	Caregiver ever been diagnosed with depression (9 months; $g = 1.29$) Caregiver feeling unhappy with relationship with partner (9 months; $g = 0.28$) Caregiver not close to their child (7; $g = 0.19$) Child considering school a waste of time (11; $g = 0.56$)
	Antisocial (B5) to antisocial (C3; 46%)	Antisocial (B5) to nonelevated (C4; 37%)	Lower caregiver extraversion score (7; $g = 13.61$) High caregiver Neuroticism score (7; $g = 10.68$) Lower caregiver overall life satisfaction (3, 11; $g = 10.25$, $g = 9.98$) Lower caregiver work–family balance (7; $g = 9.87$)

Figure 6

Sankey Diagram Showing Risk Factors That Predict Significant Transitions



Note. Colors in both Sankey diagram and bar plots correspond to risk factors domains, and the thickness of color lines between clusters indicates the importance gain associated to the risk factor. Transitions to the nonelevated difficulty scores group are omitted from the figure. MCS = Millennium Cohort Study; CPRS = Child-Parent Relationship Scale; CM = cohort member; OECD = Organisation for Economic Co-Operation and Development. See the online article for the color version of this figure.

portion of our data set increases with age, driven mostly by increasing internalizing difficulties. Given the reliance on parent-reported data in this study and the fact that internalized difficulties are by their nature less detectable by others, it is unclear whether the difficulties themselves increase in this time period, or parents become more aware of them.

The gradual shift in gender distribution replicates the well-known "gender gap" in behavioral problems during adolescence: Girls in late adolescence are at least twice as likely as boys to become depressed (Wade et al., 2002) or anxious (Zahn-Waxler et al., 2000) while boys generally have greater odds of developing externalizing problems than internalizing problems (Kovess-Masfety et al., 2016, 2021).

In addition to highlighting the gendered differences in the prevalence and presentation of parent-rated behavioral difficulties, our data-driven approach also confirms the particular importance of adolescence as a crucial transition for girls. This confirms a finding also present in other large-scale U.K.-based population-level studies that instead rely on a diagnostic framework (e.g., Ford et al., 2003). But *why* does adolescence present such a crucial developmental stage for the emergence of internalizing disorders, most especially in girls and young women? We can conceive of two broad categories of (nonmutually exclusive) possible explanations. First, there are substantial social changes that happen around this time, which may change the nature of interpersonal relationships and social structures. The most obvious example of this is school transitions

that commonly occur around early and midadolescence in the United Kingdom. Contemporaneous data that look at the role of school-level factors show that characteristics of that secondary school environment (e.g., location, "school climate") are predictive of well-being (Ford et al., 2021), something observed in other countries also (Nielsen et al., 2017). Second, there could be significant biological changes, such as hormonal changes, that play a role in driving the increase in the prevalence of internalizing difficulties particularly in young women (see Patton & Viner, 2007, for a review). Our study is not in a position to disentangle these potential influences, but it is nonetheless important to note that the gendered pattern and developmental timing that emerges from our novel data-driven approach aligns well with other large-scale studies covering this age range.

Co-Occurrence of Externalizing and Internalizing Problems in Early Childhood Predicts Complex Peer Problems in Late Childhood

In our sample, almost a quarter of children with co-occurring externalizing (conduct) and internalizing (emotional problems) problems at age 5 continued displaying a constellation of both externalizing and internalizing problems at 11 years old. Co-occurrence of externalizing and internalizing symptoms is most often found in early childhood (Fanti & Henrich, 2010; Lilienfeld, 2003; Oland & Shaw, 2005). However, little is known about their long-term prospective associations with developmental outcomes, as the literature predominantly focuses on internalizing or externalizing problems only and not their interaction over time. A few studies have examined the consequences of co-occurrence of externalizing and internalizing problems on general functioning in emerging adulthood (Arslan et al., 2021; Sourander et al., 2007). However, no studies have carefully followed up children capturing such broad areas of difficulty and assessing their trajectories in later stages of childhood or adolescence.

While most children entering preschool have developed a set of strategies to cope with their emotions and manage their behavior in demanding situations, our study suggests that children displaying broad behavioral problems at this age might find it harder to establish good coping strategies, which in turn could result in additional co-occurring difficulties over time, with a high toll on peer relationships. It has been shown that difficulties attaining adaptive strategies for emotional self-regulation during preschool age have various consequences such as notably diminished perceived social competence, lower peer status, and more externalizing problems (Blandon et al., 2010; Denham et al., 2003; Kim & Cicchetti, 2010). However, most studies look at behavioral problems in isolation while our findings suggest a link between the breadth of difficulties co-occurring in preschool age and the experience of isolation and rejection in peer contexts 6 years later.

Across this same period, from 5 to 11 years old, we found that the persistence of undifferentiated problems during childhood is related to a diverse range of socioeconomic and child health factors. The common denominator to these factors was a relatively high environmental unpredictability: a chaotic home environment, low income, and child's chronic illness—which usually requires lifestyle changes and continuous behavioral adaptation to the unpredictable course of the illness. Lower executive function was also a significant predictor of the persistence of problems over childhood, but this could be a direct consequence of the child's exposure to an unpredictable environment as repeatedly shown in previous studies (Campbell et al., 2009; Compas et al., 2012; Ferguson et al., 2013; LeBlanc et al., 2003; Skinner & Zimmer-Gembeck, 2007). Investigating the links between neurocognitive delays associated with childhood illness should be a high priority as altered trajectories of brain development may delay the deployment of the effective coping skills necessary to alleviate the occurrence of emotional and behavioral problems. But it is important to note that it is particularly difficult to establish causality, despite the temporal sequencing of predictors and their associated transitions. There could be crucial shared genetic predictors between children and parents, which themselves may significantly covary with (Trzaskowski et al., 2014) or interact with (Price & Jaffee, 2008) early environmental stressors.

Two Pathways Leading to Complex Antisocial Behaviors in Late Adolescence

We identified two pathways leading to antisocial behavior at 17 years old: The first pathway tracked by children who started developing antisocial behaviors for the first time at age 11 while the second pathway comprised children with persistent conduct problems since 5 years of age. At first glance, this mirrors Moffit's developmental taxonomy of antisocial behavior, which posits an "adolescent-limited" pathway and a "life course persistent" pathway (Moffitt, 1993). This is an important demonstration of variability of trajectory because the approach to studying antisocial behavior has hitherto largely assumed a single grouping. For example, our understanding of the role of relationship dynamics, parenting practices, cultural and community context (Dishion & Patterson, 2006), but also environmental and genetic factors (Gard et al., 2019; Wesseldijk et al., 2018), are based upon the assumption of a single etiological pathway. To our knowledge, the present study is the first to assess the risk factors for more than one possible trajectory. However, our findings do not fully align with Moffit's developmental taxonomy, according to which young people on the adolescent-limited pathway engage in antisocial behaviors as a rebellion against adult "privilege" as they become aware of those privileges but are yet unable to access them (Moffitt et al., 1996). By contrast, in our findings, this trajectory was best predicted by a caregiver's poor mental health and not by any socioeconomic factors or mental or physical health of adolescents themselves. One possibility is that this is because a parent's own mental health shapes their ability to rate, or perception of, their offspring's antisocial behavior or vice versa. Rather than situating the cause of adolescent antisocial behavior in the individual, more attention should be paid to the challenges they face in their relationships and environment.

Our childhood-onset antisocial group at 17 years old displayed lower cognitive scores in all domains and higher risk-taking scores on the gambling task (Tieskens et al., 2018). The risk factors predicting the persistence of antisocial behaviors during adolescence were related to general anxiety and antischool feelings as well as specific difficulties in reading. This latter predictor replicates several studies showing how language and reading difficulties increase the risk for behavioral problems notably by affecting social competence and moral development (Ketelaars et al., 2010; Piotrowska et al., 2019; Yew & O'Kearney, 2013). The antischool

risk factor for antisocial behavior is also in accordance with recent studies showing the importance of a positive school climate for promoting a young person's sense of community, cooperation, supportive ties among students, and more generally, spontaneous prosocial behavior (Thapa et al., 2013). It has notably been shown that adolescents reporting higher levels of perceived positive school climate at age 12 showed higher levels of prosocial behaviors over time (Luengo Kanacri et al., 2017). From this, it might be tempting to conclude that interventions should focus on improving that school environment and the young person's perception of it. However, given that this group also experiences educational difficulties that can themselves be isolating, often incurring social exclusion and stigma (e.g., Haft et al., 2023), it is likely that effective interventions also need to consider a more multisystemic approach that also supports academic development across both school and home settings.

Finally, membership of the "childhood-onset" antisocial pathway at age 11 was related to a more diverse range of risk factors, including parent–child conflicts, inconsistent discipline, high child's BMI score, poor caregiver's mental health, and caregivers' negative view of the neighborhood. These results closely match the childhood risk factors predicting future antisocial behavior found in another U.K. population-representative cohort (Avon Longitudinal Study of Parents and Children), that is, maternal hostility toward the child, maternal depression, mothers' negative view of their neighborhood, and difficulty paying the rent (Stevens, 2018). In our case, all factors were identified when the child was only 3 years old, which confirms the need for early family support interventions to prevent long-lasting antisocial behaviors that are likely to persist during adulthood and develop into antisocial personality disorders.

Persistence of Peer Problems Over Childhood and Adolescence

Peer problems were more likely to persist over childhood and adolescence than other behavioral difficulties we measured. At age 11, children had either peer problems in isolation or in combination with high scores on subscales measuring both internalized and externalized problems (emotional, conduct, and hyperactivity). Both groups principally originated from the subgroup with peer problems at age 5, meaning that some children maintain peer problems over time while others accumulate evidence of a broader range of difficulties. A significant proportion of children with isolated peer problems at age 11 then transitioned to a group with the exact same problems at age 17, which was predicted by a combination of poor caregiver's mental health and a difficult relationship with both partner and child, preceding the transition by 10 years (when the child was only 9 months old) and 4 years. Perinatal depression has repetitively been linked to poorer child development outcomes, and our study confirms the need for maternal depression interventions to prevent the persistence of peer problems in their child and adolescent (Maselko et al., 2015). Programs targeting parenting are also the leading early intervention strategy for the prevention of child behavior problems. These are typically rooted in social learning theory, which suggests that children's behavior is shaped by the behavior they observe in their caregivers (Bandura & Walters, 1963; Ryan et al., 2017). In our study, it is possible that both caregivers' relationship and child–parent relationship (closeness) act as maladaptive models for the early adolescent to initiate and foster healthy relationships with their peers. However, it is also possible, because of our reliance on parent-reported data to capture child difficulties, that parents with poor mental health and a difficult relationship with their child systematically perceive their child more negatively than parents and caregivers with better mental health and relationships. Another possibility is that shared genetic factors between children and parents may increase the risk of developing psychosocial difficulties in the presence of an unstable early environment.

Limitations

There are several limitations to this study. Although SDQ subscales have been repeatedly validated, they combine separate constructs in the same scale, such as inattention and impulsivity in the Hyperactivity scale or anxiety and depressive states in the Emotional Problems scale. Identifying dimensions based on more granular questionnaire items could capture more variability between individuals presenting with graduated levels of functioning and could potentially bring us closer to revealing underlying mechanisms.

Differential attrition between sweeps of data collection may have introduced a bias as children with fewer behavioral problems were more likely to provide data in adolescence. Consequently, the prevalence of behavioral difficulties in adolescence might be underestimated in the present study. Furthermore, this analysis was based on SDQ parent ratings as these were the only ratings consistently available across time points. Self-reported or teacher ratings might help disambiguate some of our results, especially where child behavior is related to parent mental health and parent–child relationships. Similar potential for bias applies when reporting mental health problems, which were provided either by parents or adolescents themselves, using a different set of questionnaires at each time point. Direct assessments from specialist professionals, structured interviews, and a systematic use of the same questionnaires at all time points will be needed to get a more accurate account of mental health problems incidence in childhood and adolescence. Moreover, as a U.K.-representative cohort, our sample was over 90% White: examining the impact of ethnicity on these trajectories in the United Kingdom would require oversampling of non-White participants. Although ethnicity was included as a potential risk factor in our transition analysis, we still need to replicate our study with international samples to investigate potential interactions between ethnicity and incidence of behavioral difficulties. Finally, soft clustering algorithms or latent variable models may provide more accurate allocation of participants to clusters incorporating probabilistic assignment instead of hard categorization. Although this approach might be more representative of reality, it could also dramatically increase the complexity of the analysis.

Conclusion

Establishing how children and young people's patterns of difficulties shift over time, and what predicts those shifts, is a crucial step to identify individuals in need of support and to move from a reactive to a preventive model of mental health and well-being. Using data-driven clustering and a regularizing gradient-boosting model, we identified key transitions between the age of 5–17 years old. We showed that there was an increasing number of girls developing internalized behavioral problems during adolescence,

that co-occurring internalizing and externalizing difficulties persisted during childhood in the case of a chaotic home environment, that peer problems were most likely to persist over this 12 year-period (especially in the presence of early maternal depression and poor family relationships), and finally that there were two pathways with distinct risk factors leading to antisocial behaviors in adolescence—childhood onset and adolescent onset. These findings provide evidence that investigations of child and adolescent difficulties must be open to the possibility of multiple subgroups and variability in trajectory over time. They also highlight the crucial role of family and social support and school experience factors in predicting children's outcomes, rather than situating the cause of observed difficulties solely within the young person themselves.

References

Adjei, N. K., Schlüter, D. K., Straatmann, V. S., Melis, G., Fleming, K. M., McGovern, R., Howard, L. M., Kaner, E., Wolfe, I., Taylor-Robinson, D. C., & the ORACLE consortium. (2022). Impact of poverty and family adversity on adolescent health: A multi-trajectory analysis using the UK Millennium Cohort Study. *The Lancet Regional Health Europe*, *13*, Article 100279. https://doi.org/10.1016/j.lanepe.2021.100279

Appelbaum, M., Cooper, H., Kline, R. B., Mayo-Wilson, E., Nezu, A. M., & Rao, S. M. (2018). Journal article reporting standards for quantitative research in psychology: The APA Publications and Communications Board task force report. *American Psychologist*, *73*(1), 3–25. https://doi.org/10.1037/amp0000191

Archibald, L. M. D., Oram Cardy, J., Joanisse, M. F., & Ansari, D. (2013). Language, reading, and math learning profiles in an epidemiological sample of school age children. *PLOS ONE*, *8*(10), Article e77463. https://doi.org/10.1371/journal.pone.0077463

Arslan, İ. B., Lucassen, N., van Lier, P. A. C., de Haan, A. D., & Prinzie, P. (2021). Early childhood internalizing problems, externalizing problems and their co-occurrence and (mal)adaptive functioning in emerging adulthood: A 16-year follow-up study. *Social Psychiatry and Psychiatric Epidemiology*, *56*(2), 193–206. https://doi.org/10.1007/s00127-020-01959-w

Astle, D. E., Bathelt, J., Holmes, J., & the CALM Team. (2019). Remapping the cognitive and neural profiles of children who struggle at school. *Developmental Science*, *22*(1), Article e12747. https://doi.org/10.1111/desc.12747

Astle, D. E., Holmes, J., Kievit, R., & Gathercole, S. E. (2022). Annual research review: The transdiagnostic revolution in neurodevelopmental disorders. *Journal of Child Psychology and Psychiatry*, *63*(4), 397–417. https://doi.org/10.1111/jcpp.13481

Atkinson, M. (2015). *Millennium Cohort Study—Interpreting the CANTAB cognitive measures*. Centre for Longitudinal Studies, Institute of Education, University of London.

Baird, A., Papachristou, E., Hassiotis, A., & Flouri, E. (2022). The role of physical environmental characteristics and intellectual disability in conduct problem trajectories across childhood: A population-based Cohort study. *Environmental Research*, *209*, Article 112837. https://doi.org/10.1016/j.envres.2022.112837

Bandura, A., & Walters, R. H. (1963). *Social learning and personality development*. Holt Rinehart and Winston.

Bathelt, J., Vignoles, A., & Astle, D. E. (2021). Just a phase? Mapping the transition of behavioural problems from childhood to adolescence. *Social Psychiatry and Psychiatric Epidemiology*, *56*(5), 821–836. https://doi.org/10.1007/s00127-020-02014-4

Best, J. R., & Miller, P. H. (2010). A developmental perspective on executive function. *Child Development*, *81*(6), 1641–1660. https://doi.org/10.1111/j.1467-8624.2010.01499.x

Bhutta, A. T., Cleves, M. A., Casey, P. H., Cradock, M. M., & Anand, K. J. S. (2002). Cognitive and behavioral outcomes of school-aged children who were born preterm: A meta-analysis. *JAMA*, *288*(6), 728–737. https://doi.org/10.1001/jama.288.6.728

Black, K. J., Kim, S., Yang, N. Y., & Greene, D. J. (2021). Course of tic disorders over the lifespan. *Current Developmental Disorders Reports*, *8*(2), 121–132. https://doi.org/10.1007/s40474-021-00231-3

Blakemore, S.-J. (2012). Development of the social brain in adolescence. *Journal of the Royal Society of Medicine*, *105*(3), 111–116. https://doi.org/10.1258/jrsm.2011.110221

Blakemore, S.-J., & Choudhury, S. (2006). Development of the adolescent brain: Implications for executive function and social cognition. *Journal of Child Psychology and Psychiatry*, *47*(3–4), 296–312. https://doi.org/10.1111/j.1469-7610.2006.01611.x

Blandon, A. Y., Calkins, S. D., & Keane, S. P. (2010). Predicting emotional and social competence during early childhood from toddler risk and maternal behavior. *Development and Psychopathology*, *22*(1), 119–132. https://doi.org/10.1017/S0954579409990307

Bloomfield, L., & Kendall, S. (2012). Parenting self-efficacy, parenting stress and child behaviour before and after a parenting programme. *Primary Health Care Research and Development*, *13*(4), 364–372. https://doi.org/10.1017/S1463423612000060

Brunborg, G. S., & Burdzovic Andreas, J. (2019). Increase in time spent on social media is associated with modest increase in depression, conduct problems, and episodic heavy drinking. *Journal of Adolescence*, *74*(1), 201–209. https://doi.org/10.1016/j.adolescence.2019.06.013

Campbell, L. K., Scaduto, M., Van Slyke, D., Niarhos, F., Whitlock, J. A., & Compas, B. E. (2009). Executive function, coping, and behavior in survivors of childhood acute lymphocytic leukemia. *Journal of Pediatric Psychology*, *34*(3), 317–327. https://doi.org/10.1093/jpepsy/jsn080

Caspi, A., & Moffitt, T. E. (2018). All for one and one for all: Mental disorders in one dimension. *The American Journal of Psychiatry*, *175*(9), 831–844. https://doi.org/10.1176/appi.ajp.2018.17121383

Clair, A. (2019). Housing: An under-explored influence on children's well-being and becoming. *Child Indicators Research*, *12*(2), 609–626. https://doi.org/10.1007/s12187-018-9550-7

Coldwell, J., Pike, A., & Dunn, J. (2006). Household chaos—Links with parenting and child behaviour. *Journal of Child Psychology and Psychiatry*, *47*(11), 1116–1122. https://doi.org/10.1111/j.1469-7610.2006.01655.x

Compas, B. E., Jaser, S. S., Dunn, M. J., & Rodriguez, E. M. (2012). Coping with chronic illness in childhood and adolescence. *Annual Review of Clinical Psychology*, *8*(1), 455–480. https://doi.org/10.1146/annurev-clinpsy-032511-143108

Copeland, W. E., Angold, A., Shanahan, L., & Costello, E. J. (2014). Longitudinal patterns of anxiety from childhood to adulthood: The Great Smoky Mountains Study. *Journal of the American Academy of Child and Adolescent Psychiatry*, *53*(1), 21–33. https://doi.org/10.1016/j.jaac.2013.09.017

Cuthbert, B. N., & Insel, T. R. (2013). Toward the future of psychiatric diagnosis: The seven pillars of RDoC. *BMC Medicine*, *11*(1), Article 126. https://doi.org/10.1186/1741-7015-11-126

Dalgleish, T., Black, M., Johnston, D., & Bevan, A. (2020). Transdiagnostic approaches to mental health problems: Current status and future directions. *Journal of Consulting and Clinical Psychology*, *88*(3), 179–195. https://doi.org/10.1037/ccp0000482

Dalmaijer, E. S., Nord, C. L., & Astle, D. E. (2021). *Statistical power for cluster analysis*. arXiv. https://doi.org/10.48550/arXiv.2003.00381

Denham, S. A., Blair, K. A., DeMulder, E., Levitas, J., Sawyer, K., Auerbach-Major, S., & Queenan, P. (2003). Preschool emotional competence: Pathway to social competence? *Child Development*, *74*(1), 238–256. https://doi.org/10.1111/1467-8624.00533

Dishion, T. J., & Patterson, G. R. (2006). The development and ecology of antisocial behavior in children and adolescents. In D. Cicchetti & D.

Cohen (Eds.), *Developmental psychopathology: Vol. 3. Risk, disorder, and adaptation* (2nd ed., pp. 503–541). Wiley.

Dolan, M., & Lennox, C. (2013). Cool and hot executive function in conduct-disordered adolescents with and without co-morbid attention deficit hyperactivity disorder: Relationships with externalizing behaviours. *Psychological Medicine*, *43*(11), 2427–2436. https://doi.org/10.1017/S0033291712003078

Donnellan, M. B., Trzesniewski, K. H., Robins, R. W., Moffitt, T. E., & Caspi, A. (2005). Low self-esteem is related to aggression, antisocial behavior, and delinquency. *Psychological Science*, *16*(4), 328–335. https://doi.org/10.1111/j.0956-7976.2005.01535.x

Egger, H. L., Costello, E. J., & Angold, A. (2003). School refusal and psychiatric disorders: A community study. *Journal of the American Academy of Child and Adolescent Psychiatry*, *42*(7), 797–807. https://doi.org/10.1097/01.CHI.0000046865.56865.79

Elliott, C. D., Smith, P., & McCulloch, K. (1996). *British ability scales second edition (BAS II)*. Nelson.

Etkin, A., & Cuthbert, B. (2014). Beyond the DSM: Development of a transdiagnostic psychiatric neuroscience course. *Academic Psychiatry*, *38*(2), 145–150. https://doi.org/10.1007/s40596-013-0032-4

Fanti, K. A., & Henrich, C. C. (2010). Trajectories of pure and co-occurring internalizing and externalizing problems from age 2 to age 12: Findings from the National Institute of Child Health and Human Development Study of early child care. *Developmental Psychology*, *46*(5), 1159–1175. https://doi.org/10.1037/a0020659

Ferguson, K. T., Cassells, R. C., MacAllister, J. W., & Evans, G. W. (2013). The physical environment and child development: An international review. *International Journal of Psychology*, *48*(4), 437–468. https://doi.org/10.1080/00207594.2013.804190

Flouri, E., Midouhas, E., & Joshi, H. (2015). Family and neighbourhood risk and children's problem behaviour: The moderating role of intelligence. *Intelligence*, *53*, 33–42. https://doi.org/10.1016/j.intell.2015.08.003

Flouri, E., Moulton, V., & Ploubidis, G. B. (2019). The role of intelligence in decision-making in early adolescence. *British Journal of Developmental Psychology*, *37*(1), 101–111. https://doi.org/10.1111/bjdp.12261

Ford, T., Degli Esposti, M., Crane, C., Taylor, L., Montero-Marín, J., Blakemore, S.-J., Bowes, L., Byford, S., Dalgleish, T., Greenberg, M. T., Nuthall, E., Phillips, A., Raja, A., Ukoumunne, O. C., Viner, R. M., Williams, J. M. G., Allwood, M., Aukland, L., Casey, T., … Kuyken, W. (2021). The role of schools in early adolescents' mental health: Findings from the MYRIAD Study. *Journal of the American Academy of Child and Adolescent Psychiatry*, *60*(12), 1467–1478. https://doi.org/10.1016/j.jaac.2021.02.016

Ford, T., Goodman, R., & Meltzer, H. (2003). The British Child and Adolescent Mental Health Survey 1999: The prevalence of DSM-IV disorders. *Journal of the American Academy of Child & Adolescent Psychiatry*, *42*(10), 1203–1211. https://doi.org/10.1097/00004583-200310000-00011

Franke, B., Michelini, G., Asherson, P., Banaschewski, T., Bilbow, A., Buitelaar, J. K., Cormand, B., Faraone, S. V., Ginsberg, Y., Haavik, J., Kuntsi, J., Larsson, H., Lesch, K.-P., Ramos-Quiroga, J. A., Réthelyi, J. M., Ribases, M., & Reif, A. (2018). Live fast, die young? A review on the developmental trajectories of ADHD across the lifespan. *European Neuropsychopharmacology*, *28*(10), 1059–1088. https://doi.org/10.1016/j.euroneuro.2018.08.001

Furlong, M., Herring, A. H., Goldman, B. D., Daniels, J. L., Wolff, M. S., Engel, L. S., & Engel, S. M. (2018). Early life characteristics and neurodevelopmental phenotypes in the Mount Sinai Children's Environmental Health Center. *Child Psychiatry and Human Development*, *49*(4), 534–550. https://doi.org/10.1007/s10578-017-0773-5

Furman, W., & Buhrmester, D. (1992). Age and sex differences in perceptions of networks of personal relationships. *Child Development*, *63*(1), 103–115. https://doi.org/10.2307/1130905

Gard, A. M., Dotterer, H. L., & Hyde, L. W. (2019). Genetic influences on antisocial behavior: Recent advances and future directions. *Current Opinion in Psychology*, *27*, 46–55. https://doi.org/10.1016/j.copsyc.2018.07.013

Giannakopoulos, G., Mihas, C., Dimitrakaki, C., & Tountas, Y. (2009). Family correlates of adolescents' emotional/behavioural problems: Evidence from a Greek school-based sample. *Acta Paediatrica*, *98*(8), 1319–1323. https://doi.org/10.1111/j.1651-2227.2009.01314.x

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

Goodman, R., Ford, T., Simmons, H., Gatward, R., & Meltzer, H. (2000). Using the Strengths and Difficulties Questionnaire (SDQ) to screen for child psychiatric disorders in a community sample. *The British Journal of Psychiatry*, *177*(6), 534–539. https://doi.org/10.1192/bjp.177.6.534

Granada Learning Assessment. (2024). Number analogies: Quantitative reasoning battery. *Millennium Cohort Study—Age 17*. https://closer.ac.uk/cross-study-data-guides/cognitive-measures-guide/mcs-cognition/mcs-age-17-number-analogies

Griffiths, L. J., Dezateux, C., & Hill, A. (2011). Is obesity associated with emotional and behavioural problems in children? Findings from the Millennium Cohort Study. *International Journal of Pediatric Obesity*, *6*(2), e423–e432. https://doi.org/10.3109/17477166.2010.526221

Haft, S. L., Greiner de Magalhães, C., & Hoeft, F. (2023). A systematic review of the consequences of stigma and stereotype threat for individuals with specific learning disabilities. *Journal of Learning Disabilities*, *56*(3), 193–209. https://doi.org/10.1177/00222194221087383

Hallett, S., & Crompton, C. J. (2018). *Too complicated to treat? Austic people seeking mental health support in Scotland*. Autistic Mutual Aid Society Ediburgh (AMASE). https://amase.org.uk/wp-content/uploads/2018/11/AMASE-autism-and-mh-report-full-ver.pdf

Hansen, B. H., Oerbeck, B., Skirbekk, B., Petrovski, B. É., & Kristensen, H. (2018). Neurodevelopmental disorders: Prevalence and comorbidity in children referred to mental health services. *Nordic Journal of Psychiatry*, *72*(4), 285–291. https://doi.org/10.1080/08039488.2018.1444087

Hartman, C. A., Luteijn, E., Serra, M., & Minderaa, R. (2006). Refinement of the Children's Social Behavior Questionnaire (CSBQ): An instrument that describes the diverse problems seen in milder forms of PDD. *Journal of Autism and Developmental Disorders*, *36*(3), 325–342. https://doi.org/10.1007/s10803-005-0072-z

Heron, J., Maughan, B., Dick, D. M., Kendler, K. S., Lewis, G., Macleod, J., Munafò, M., & Hickman, M. (2013). Conduct problem trajectories and alcohol use and misuse in mid to late adolescence. *Drug and Alcohol Dependence*, *133*(1), 100–107. https://doi.org/10.1016/j.drugalcdep.2013.05.025

Holley, S., Hill, C. M., & Stevenson, J. (2011). An hour less sleep is a risk factor for childhood conduct problems. *Child: Care, Health and Development*, *37*(4), 563–570. https://doi.org/10.1111/j.1365-2214.2010.01203.x

Holmes, J., Mareva, S., Bennett, M. P., Black, M. J., & Guy, J. (2021). Higher-order dimensions of psychopathology in a neurodevelopmental transdiagnostic sample. *Journal of Abnormal Psychology*, *130*(8), 909–922. https://doi.org/10.1037/abn0000710

Jacob, S., Wolff, J. J., Steinbach, M. S., Doyle, C. B., Kumar, V., & Elison, J. T. (2019). Neurodevelopmental heterogeneity and computational approaches for understanding autism. *Translational Psychiatry*, *9*(1), Article 63. https://doi.org/10.1038/s41398-019-0390-0

Kessler, R. C., Barker, P. R., Colpe, L. J., Epstein, J. F., Gfroerer, J. C., Hiripi, E., Howes, M. J., Normand, S.-L. T., Manderscheid, R. W., Walters, E. E., & Zaslavsky, A. M. (2003). Screening for serious mental illness in the general population. *Archives of General Psychiatry*, *60*(2), 184–189. https://doi.org/10.1001/archpsyc.60.2.184

Ketelaars, M. P., Cuperus, J., Jansonius, K., & Verhoeven, L. (2010). Pragmatic language impairment and associated behavioural problems.

International Journal of Language & Communication Disorders, *45*(2), 204–214. https://doi.org/10.3109/13682820902863090

Kim, J., & Cicchetti, D. (2010). Longitudinal pathways linking child maltreatment, emotion regulation, peer relations, and psychopathology. *Journal of Child Psychology and Psychiatry*, *51*(6), 706–716. https://doi.org/10.1111/j.1469-7610.2009.02202.x

Kovess-Masfety, V., Husky, M. M., Keyes, K., Hamilton, A., Pez, O., Bitfoi, A., Carta, M. G., Goelitz, D., Kuijpers, R., Otten, R., Koç, C., Lesinskiene, S., & Mihova, Z. (2016). Comparing the prevalence of mental health problems in children 6–11 across Europe. *Social Psychiatry and Psychiatric Epidemiology*, *51*(8), 1093–1103. https://doi.org/10.1007/s00127-016-1253-0

Kovess-Masfety, V., Woodward, M. J., Keyes, K., Bitfoi, A., Carta, M. G., Koç, C., Lesinskiene, S., Mihova, Z., Otten, R., & Husky, M. (2021). Gender, the gender gap, and their interaction; analysis of relationships with children's mental health problems. *Social Psychiatry and Psychiatric Epidemiology*, *56*(6), 1049–1057. https://doi.org/10.1007/s00127-020-01950-5

Kushki, A., Anagnostou, E., Hammill, C., Duez, P., Brian, J., Iaboni, A., Schachar, R., Crosbie, J., Arnold, P., & Lerch, J. P. (2019). Examining overlap and homogeneity in ASD, ADHD, and OCD: A data-driven, diagnosis-agnostic approach. *Translational Psychiatry*, *9*(1), Article 318. https://doi.org/10.1038/s41398-019-0631-2

Lacey, R. E., Howe, L. D., Kelly-Irving, M., Bartley, M., & Kelly, Y. (2022). The clustering of adverse childhood experiences in the Avon longitudinal study of parents and children: Are gender and poverty important? *Journal of Interpersonal Violence*, *37*(5–6), 2218–2241. https://doi.org/10.1177/0886260520935096

Law, J., Rush, R., Anandan, C., Cox, M., & Wood, R. (2012). Predicting language change between 3 and 5 years and its implications for early identification. *Pediatrics*, *130*(1), e132–e137. https://doi.org/10.1542/peds.2011-1673

LeBlanc, L. A., Goldsmith, T., & Patel, D. R. (2003). Behavioral aspects of chronic illness in children and adolescents. *Pediatric Clinics of North America*, *50*(4), 859–878. https://doi.org/10.1016/S0031-3955(03)00072-5

Lewis, M., & Rudolph, K. D. (2014). *Handbook of developmental psychopathology*. Springer Science & Business Media. https://doi.org/10.1007/978-1-4614-9608-3

Lilienfeld, S. O. (2003). Comorbidity between and within childhood externalizing and internalizing disorders: Reflections and directions. *Journal of Abnormal Child Psychology*, *31*(3), 285–291. https://doi.org/10.1023/A:1023229529866

Luengo Kanacri, B. P., Eisenberg, N., Thartori, E., Pastorelli, C., Uribe Tirado, L. M., Gerbino, M., & Caprara, G. V. (2017). Longitudinal relations among positivity, perceived positive school climate, and prosocial behavior in Colombian adolescents. *Child Development*, *88*(4), 1100–1114. https://doi.org/10.1111/cdev.12863

Maselko, J., Sikander, S., Bhalotra, S., Bangash, O., Ganga, N., Mukherjee, S., Egger, H., Franz, L., Bibi, A., Liaqat, R., Kanwal, M., Abbasi, T., Noor, M., Ameen, N., & Rahman, A. (2015). Effect of an early perinatal depression intervention on long-term child development outcomes: Follow-up of the thinking healthy programme randomised controlled trial. *The Lancet Psychiatry*, *2*(7), 609–617. https://doi.org/10.1016/S2215-0366(15)00109-1

Mayes, S. D., Calhoun, S. L., Chase, G. A., Mink, D. M., & Stagg, R. E. (2009). ADHD subtypes and co-occurring anxiety, depression, and oppositional-defiant disorder: Differences in Gordon diagnostic system and Wechsler working memory and processing speed index scores. *Journal of Attention Disorders*, *12*(6), 540–550. https://doi.org/10.1177/1087054708320402

McCormick, M. C., Gortmaker, S. L., & Sobol, A. M. (1990). Very low birth weight children: Behavior problems and school difficulty in a national sample. *The Journal of Pediatrics*, *117*(5), 687–693. https://doi.org/10.1016/S0022-3476(05)83322-0

McInnes, L., Healy, J., & Melville, J. (2020). *UMAP: Uniform manifold approximation and projection for dimension reduction*. arXiv. https://doi.org/10.48550/arXiv.1802.03426

Meadows, S. O., McLanahan, S. S., & Brooks-Gunn, J. (2007). Parental depression and anxiety and early childhood behavior problems across family types. *Journal of Marriage and Family*, *69*(5), 1162–1177. https://doi.org/10.1111/j.1741-3737.2007.00439.x

Miller, P., Whitfield, K., Betancur, L., & Votruba-Drzal, E. (2021). Income dynamics and behavior problems in early childhood, middle childhood, and the transition to adolescence. *Journal of Applied Developmental Psychology*, *77*, Article 101345. https://doi.org/10.1016/j.appdev.2021.101345

Minh, A., Muhajarine, N., Janus, M., Brownell, M., & Guhn, M. (2017). A review of neighborhood effects and early child development: How, where, and for whom, do neighborhoods matter? *Health & Place*, *46*, 155–174. https://doi.org/10.1016/j.healthplace.2017.04.012

Moffitt, T. E. (1993). Adolescence-limited and life-course-persistent antisocial behavior: A developmental taxonomy. *Psychological Review*, *100*(4), 674–701. https://doi.org/10.1037/0033-295X.100.4.674

Moffitt, T. E. (2006). Life-course-persistent versus adolescence-limited antisocial behavior. In D. Cicchetti & D. J. Cohen (Eds.), *Developmental psychopathology: Vol. 3. Risk, disorder, and adaptation* (2nd ed., pp. 570–598). Wiley.

Moffitt, T. E., Caspi, A., Dickson, N., Silva, P., & Stanton, W. (1996). Childhood-onset versus adolescent-onset antisocial conduct problems in males: Natural history from ages 3 to 18 years. *Development and Psychopathology*, *8*(2), 399–424. https://doi.org/10.1017/S0954579400007161

Moulton, V., McElroy, E., Richards, M., Fitzsimons, E., Northstone, K., Conti, G., Ploubidis, G. B., Sullivan, A., & O'Neill, D. (2020). *A guide to the cognitive measures in five British birth cohort studies* (Report). Closer. https://www.closer.ac.uk/cognitive-measures-guide/

Nielsen, L., Shaw, T., Meilstrup, C., Koushede, V., Bendtsen, P., Rasmussen, M., Lester, L., Due, P., & Cross, D. (2017). School transition and mental health among adolescents: A comparative study of school systems in Denmark and Australia. *International Journal of Educational Research*, *83*, 65–74. https://doi.org/10.1016/j.ijer.2017.01.011

Oland, A. A., & Shaw, D. S. (2005). Pure versus co-occurring externalizing and internalizing symptoms in children: The potential role of socio-developmental milestones. *Clinical Child and Family Psychology Review*, *8*(4), 247–270. https://doi.org/10.1007/s10567-005-8808-z

Patton, G. C., & Viner, R. (2007). Pubertal transitions in health. *Lancet*, *369*(9567), 1130–1139. https://doi.org/10.1016/S0140-6736(07)60366-3

Piotrowska, P. J., Stride, C. B., Maughan, B., & Rowe, R. (2019). Mechanisms underlying social gradients in child and adolescent antisocial behaviour. *SSM—Population Health*, *7*, Article 100353. https://doi.org/10.1016/j.ssmph.2019.100353

Price, T. S., & Jaffee, S. R. (2008). Effects of the family environment: Gene–environment interaction and passive gene–environment correlation. *Developmental Psychology*, *44*(2), 305–315. https://doi.org/10.1037/0012-1649.44.2.305

Quistberg, K. A., & Mueller, U. (2020). Prospective relations between kindergarteners' executive function skills and their externalizing and internalizing behaviors. *The Clinical Neuropsychologist*, *34*(4), 845–862. https://doi.org/10.1080/13854046.2019.1591510

R Core Team. (2020). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing. https://www.R-project.org/

Rosenberg, M. (1965). *Rosenberg Self-Esteem Scale (RSES)*. APA PsycTests. https://doi.org/10.1037/t01038-000

Russell, G., Rodgers, L. R., & Ford, T. (2013). The strengths and difficulties questionnaire as a predictor of parent-reported diagnosis of autism spectrum disorder and attention deficit hyperactivity disorder. *PLOS ONE*, *8*(12), Article e80247. https://doi.org/10.1371/journal.pone.0080247

Ryan, R., O'Farrelly, C., & Ramchandani, P. (2017). Parenting and child mental health. *London Journal of Primary Care*, *9*(6), 86–94. https://doi.org/10.1080/17571472.2017.1361630

Schepman, K., Collishaw, S., Gardner, F., Maughan, B., Scott, J., & Pickles, A. (2011). Do changes in parent mental health explain trends in youth emotional problems? *Social Science & Medicine*, *73*(2), 293–300. https://doi.org/10.1016/j.socscimed.2011.05.015

Schoon, I., Parsons, S., Rush, R., & Law, J. (2010). Childhood language skills and adult literacy: A 29-year follow-up study. *Pediatrics*, *125*(3), e459–e466. https://doi.org/10.1542/peds.2008-2111

Serati, M., Barkin, J. L., Orsenigo, G., Altamura, A. C., & Buoli, M. (2017). Research review: The role of obstetric and neonatal complications in childhood attention deficit and hyperactivity disorder—A systematic review. *Journal of Child Psychology and Psychiatry*, *58*(12), 1290–1300. https://doi.org/10.1111/jcpp.12779

Skinner, E. A., & Zimmer-Gembeck, M. J. (2007). The development of coping. *Annual Review of Psychology*, *58*(1), 119–144. https://doi.org/10.1146/annurev.psych.58.110405.085705

Sourander, A., Jensen, P., Davies, M., Niemelä, S., Elonheimo, H., Ristkari, T., Helenius, H., Sillanmäki, L., Piha, J., Kumpulainen, K., Tamminen, T., Moilanen, I., & Almqvist, F. (2007). Who is at greatest risk of adverse long-term outcomes? The Finnish from a Boy to a Man study. *Journal of the American Academy of Child and& Adolescent Psychiatry*, *46*(9), 1148–1161. https://doi.org/10.1097/chi.0b013e31809861e9

Stevens, M. (2018). Preventing at-risk children from developing antisocial and criminal behaviour: A longitudinal study examining the role of parenting, community and societal factors in middle childhood. *BMC Psychology*, *6*(1), Article 40. https://doi.org/10.1186/s40359-018-0254-z

Stewart-Brown, S., Tennant, A., Tennant, R., Platt, S., Parkinson, J., & Weich, S. (2009). Internal construct validity of the Warwick-Edinburgh Mental Well-being Scale (WEMWBS): A Rasch analysis using data from the Scottish Health Education Population Survey. *Health and Quality of Life Outcomes*, *7*(1), Article 15. https://doi.org/10.1186/1477-7525-7-15

Sullivan, A., Ketende, S., & Joshi, H. (2013). Social class and inequalities in early cognitive scores. *Sociology*, *47*(6), 1187–1206. https://doi.org/10.1177/0038038512461861

Swift, W., Slade, T., Carragher, N., Coffey, C., Degenhardt, L., Hall, W., & Patton, G. (2016). Adolescent predictors of a typology of DSM-5 alcohol use disorder symptoms in young adults derived by latent class analysis using data from an Australian Cohort Study. *Journal of Studies on Alcohol and Drugs*, *77*(5), 757–765. https://doi.org/10.15288/jsad.2016.77.757

Taurines, R., Schmitt, J., Renner, T., Conner, A. C., Warnke, A., & Romanos, M. (2010). Developmental comorbidity in attention-deficit/hyperactivity disorder. *Attention Deficit and Hyperactivity Disorders*, *2*(4), 267–289. https://doi.org/10.1007/s12402-010-0040-0

Thapa, A., Cohen, J., Guffey, S., & Higgins-D'Alessandro, A. (2013). A review of school climate research. *Review of Educational Research*, *83*(3), 357–385. https://doi.org/10.3102/0034654313483907

Thompson, W. K., Barch, D. M., Bjork, J. M., Gonzalez, R., Nagel, B. J., Nixon, S. J., & Luciana, M. (2019). The structure of cognition in 9 and 10 year-old children and associations with problem behaviors: Findings from the ABCD study's baseline neurocognitive battery. *Developmental Cognitive Neuroscience*, *36*, Article 100606. https://doi.org/10.1016/j.dcn.2018.12.004

Tieskens, J. M., Buil, J. M., Koot, S., Krabbendam, L., & van Lier, P. A. C. (2018). Elementary school children's associations of antisocial behaviour with risk-taking across 7–11 years. *Journal of Child Psychology and Psychiatry*, *59*(10), 1052–1060. https://doi.org/10.1111/jcpp.12943

Trzaskowski, M., Harlaar, N., Arden, R., Krapohl, E., Rimfeld, K., McMillan, A., Dale, P. S., & Plomin, R. (2014). Genetic influence on family socioeconomic status and children's intelligence. *Intelligence*, *42*(100), 83–88. https://doi.org/10.1016/j.intell.2013.11.002

Turnbull, K., Reid, G. J., & Morton, J. B. (2013). Behavioral sleep problems and their potential impact on developing executive function in children. *Sleep*, *36*(7), 1077–1084. https://doi.org/10.5665/sleep.2814

University College London, UCL Social Research Institute, Centre for Longitudinal Studies. (2024). *Millennium cohort study* [data series]. 16th Release. UK Data Service. SN: 2000031. https://doi.org/10.5255/UKDA-Series-2000031

Uh, S., Dalmaijer, E. S., Siugzdaite, R., Ford, T. J., & Astle, D. E. (2021). Two pathways to self-harm in adolescence. *Journal of the American Academy of Child & Adolescent Psychiatry*, *60*(12), 1491–1500. https://doi.org/10.1016/j.jaac.2021.03.010

Van Oort, F. V., Greaves-Lord, K., Verhulst, F. C., Ormel, J., & Huizink, A. C. (2009). The developmental course of anxiety symptoms during adolescence: The TRAILS study. *Journal of Child Psychology and Psychiatry*, *50*(10), 1209–1217. https://doi.org/10.1111/j.1469-7610.2009.02092.x

Van Roy, B., Groholt, B., Heyerdahl, S., & Clench-Aas, J. (2010). Understanding discrepancies in parent–child reporting of emotional and behavioural problems: Effects of relational and socio-demographic factors. *BMC Psychiatry*, *10*(1), Article 56. https://doi.org/10.1186/1471-244X-10-56

van Widenfelt, B. M., Goedhart, A. W., Treffers, P. D. A., & Goodman, R. (2003). Dutch version of the Strengths and Difficulties Questionnaire (SDQ). *European Child & Adolescent Psychiatry*, *12*(6), 281–289. https://doi.org/10.1007/s00787-003-0341-3

Vogels, A. G. C., Siebelink, B. M., Theunissen, M. H. C., Wolff, M. S. D., & Reijneveld, S. A. (2011). *Vergelijking van de KIVPA en de SDQ als signaleringsinstrument voor problemen bij adolescenten in de jeugdgezondheidszorg* [Comparison of the KIVPA and SDQ as used for screening among adolescents in youth healthcare]. TNO.

Vugteveen, J., de Bildt, A., Theunissen, M., Reijneveld, S. A., & Timmerman, M. (2021). Validity aspects of the Strengths and Difficulties Questionnaire (SDQ) adolescent self-report and parent-report versions among Dutch adolescents. *Assessment*, *28*(2), 601–616. https://doi.org/10.1177/1073191119858416

Wade, T. J., Cairney, J., & Pevalin, D. J. (2002). Emergence of gender differences in depression during adolescence: National panel results from three countries. *Journal of the American Academy of Child and Adolescent Psychiatry*, *41*(2), 190–198. https://doi.org/10.1097/00004583-200202000-00013

Wardenaar, K. J., & de Jonge, P. (2013). Diagnostic heterogeneity in psychiatry: Towards an empirical solution. *BMC Medicine*, *11*(1), Article 201. https://doi.org/10.1186/1741-7015-11-201

Wesseldijk, L. W., Bartels, M., Vink, J. M., van Beijsterveldt, C. E. M., Ligthart, L., Boomsma, D. I., & Middeldorp, C. M. (2018). Genetic and environmental influences on conduct and antisocial personality problems in childhood, adolescence, and adulthood. *European Child & Adolescent Psychiatry*, *27*(9), 1123–1132. https://doi.org/10.1007/s00787-017-1014-y

Wickham, H. (2016). *ggplot2: Elegant Graphics for Data Analysis*. Springer-Verlag. https://ggplot2.tidyverse.org

Wiles, N. J., Peters, T. J., Heron, J., Gunnell, D., Emond, A., Lewis, G., & the ALSPAC Study Team. (2006). Fetal growth and childhood behavioral problems: Results from the ALSPAC cohort. *American Journal of Epidemiology*, *163*(9), 829–837. https://doi.org/10.1093/aje/kwj108

Yang, Y., Sun, H., Zhang, Y., Zhang, T., Gong, J., Wei, Y., Duan, Y.-G., Shu, M., Yang, Y., Wu, D., & Yu, D. (2021). Dimensionality reduction by UMAP reinforces sample heterogeneity analysis in bulk transcriptomic data. *Cell Reports*, *36*(4), Article 109442. https://doi.org/10.1016/j.celrep.2021.109442

Yew, S. G. K., & O'Kearney, R. (2013). Emotional and behavioural outcomes later in childhood and adolescence for children with specific language impairments: Meta-analyses of controlled prospective studies.

Journal of Child Psychology and Psychiatry, *54*(5), 516–524. https://doi.org/10.1111/jcpp.12009

Zahn-Waxler, C., Klimes-Dougan, B., & Slattery, M. J. (2000). Internalizing problems of childhood and adolescence: Prospects, pitfalls, and progress in understanding the development of anxiety and depression. *Development and Psychopathology*, *12*(3), 443–466. https://doi.org/10.1017/S0954579400003102

Zumel, N., & Mount, J. (2014). *Practical data science with R* (1st ed.). Manning Publications.

Received April 19, 2023
Revision received August 17, 2024
Accepted September 6, 2024 ■

E-Mail Notification of Your Latest Issue Online!

Would you like to know when the next issue of your favorite APA journal will be available online? This service is now available to you. Sign up at https://my.apa.org/portal/alerts/ and you will be notified by e-mail when issues of interest to you become available!