nature mental health

Article

https://doi.org/10.1038/s44220-026-00727-0

Covariation among higher-order psychopathology dimensions is driven by nomorbidity and not comorbidity

Received: 6 August 2025

Accepted: 14 August 2026

Published online: 1 October 2026

Check for updates

Ashley L. Watts[1] ✉, Adon F. G. Rosen[1], Ashley L. Greene[2,3,4], Kevin M. King[5], Francisco A. C. Meyer[6], Timothy J. Trull[7], Kenneth J. Sher[7] & Keanan J. Joyner[8]

Emerging classification frameworks claim that covariation among psychopathology dimensions reflects comorbidity. We scrutinized this claim by applying Bayesian change point models to four waves of the Adolescent Brain Cognitive Development Study data (N = 11,868 youths; 48% female) and examined associations among higher-order psychopathology dimensions. Covariation between psychopathology dimensions was driven by nomorbidity, not comorbidity: we observed strong positive associations among psychopathology dimensions at relatively extreme low (asymptomatic) levels (for example, when three to five symptoms were endorsed) and weak positive associations at higher levels. For instance, at wave 1, a 1-point increase in internalizing was associated with a 32% increase in externalizing at nearly asymptomatic levels (B = 0.28), compared with only a 7% increase at higher levels (B = 0.07). These findings challenge existing assumptions about comorbidity in nonclinical samples, suggesting that covariation is more accurately explained by the absence of symptomatology rather than shared psychiatric conditions.

Psychiatric comorbidity, the co-occurrence of two or more mental disorders, is often regarded as the 'rule rather than the exception'[1,2]. But epidemiologic data suggest that it is statistically uncommon to be diagnosed with two or more conditions, let alone three or more, at the same time[3]. At a given time, most people in the general population do not exhibit appreciable levels of symptomatology to warrant a psychiatric diagnosis. Around 70% of the population exhibits what we term 'nomorbidity'—wherein an individual is not diagnosed with a single mental disorder—when past-year diagnoses are assessed, and 50% when lifetime diagnoses are assessed[3]. Further, when followed over long stretches of time, many people (~17–40%) are not diagnosed with a single mental disorder at any time[4,5]. Of course, conditional upon receipt of one diagnosis, comorbidity is far more common. If a person receives one diagnosis, their odds of receiving two diagnoses increases, and so on.

Modern classification systems, such as the Hierarchical Taxonomy of Psychopathology[6], argue that this 'rampant comorbidity problem'[7,8] results from the fact that the Diagnostic and Statistical Manual of Mental Disorders and related systems 'do not carve nature at its joints'[9–11]. According to this criticism, comorbidity is largely conceivably artefactual, arising because we draw arbitrary borders around phenomena that overlap phenotypically and etiologically[6,12]. In response, these systems harness the supposed overlap among putatively distinct conditions using covariance-based modeling and reorganize

[1]Department of Psychological Sciences, Vanderbilt University, Nashville, TN, USA. [2]Long Island University-Brooklyn, Brooklyn, NY, USA. [3]VISN 2 Mental Illness Research, Education and Clinical Center, James J. Peters VA Medical Center, Bronx, NY, USA. [4]Department of Psychiatry, Icahn School of Medicine at Mount Sinai, New York, NY, USA. [5]Department of Psychology, University of Washington, Seattle, WA, USA. [6]Department of Psychiatry, Yale University School of Medicine, New Haven, CT, USA. [7]Department of Psychological Sciences, University of Missouri, Columbia, MO, USA. [8]Department of Psychology, University of California, Berkeley, Berkeley, CA, USA. ✉e-mail: ashley.watts@vanderbilt.edu

psychopathology into a more parsimonious set of dimensions. Products of covariance-based models have, in turn, been widely promoted as capturing sources of psychiatric comorbidity[13,14].

Across nearly half of a century and thousands of studies, these efforts have converged upon two broad psychopathology dimensions: externalizing and internalizing[6,15–17], which were first studied in youth[15] nearly thirty or so years before they were studied in adults[17]. Externalizing encompasses conditions of poor behavioral and emotional control[18] (for example, conduct disorder, oppositional defiant disorder and substance use disorder), whereas internalizing encompasses conditions that are characterized by intense negative emotions[19] (for example, generalized anxiety disorder, major depressive disorder and specific phobia). The idea behind these transdiagnostic dimensions is that, say, conduct disorder and substance use disorder exhibit high degrees of comorbidity because they are phenotypically and etiologically overlapping, with disinhibition explaining their overlap[18,20].

On the basis of the correlation between externalizing and internalizing ($r \approx 0.5–0.6$), various researchers further proposed that a single dimension[21–23]—the *p*-factor—is the etiologic source of psychiatric comorbidity: 'a single dimension is able to measure and maybe even explain a person's liability to mental disorder, comorbidity among disorders, persistence of disorders over time, and severity of symptoms' (pp. 831[21]; see also refs. 24–29). This claim positively asserts that the covariation between externalizing and internalizing reflects some shared substantive mechanism (for example, poor emotional control[30,31] and negative emotionality[32,33]).

We propose an alternative explanation, that evidence of so-called comorbidity arises owing to the considerable extent of asymptomatic cases in the population. That is, covariation among forms of psychopathology is driven by the inclusion of people who report little to no psychopathology, a typical presentation in large-scale epidemiologic and otherwise community-based samples that are most often used to model hierarchical structures of psychopathology[18,19].

As the basis of this claim, we borrow from a larger literature that suggests that the inclusion of asymptomatic (or undiagnosed) cases in covariance-based analyses of the structure of psychological constructs encourages their unidimensionality[34], or inflates the overlap among the phenomena of interest. Including asymptomatic cases produces dense network[35] structures and coherent (unidimensional) factor[34,36] structures that—although people assume they capture comorbidity—are driven by people who endorse few, if any, symptoms[37]. That is, what might glue together symptoms is a lack of comorbidity, or nomorbidity (if you find this claim counterintuitive, we expand on our argument in Appendix A).

We extend these findings to examine associations among higher-order psychopathology dimensions, whose covariation is interpreted as being synonymous with psychiatric comorbidity, using innovative statistical modeling techniques. Here, we focus on youth who took part in the Adolescent Brain Cognitive Development Study (ABCD)[38] given that externalizing and internalizing were first observed among youth[15] and because epidemiologic data are often used to characterize the structure of psychopathology[6]. The large size of the ABCD Study affords one of the best powered examinations of the interface between higher-order psychopathology dimensions in youth.

Results

Univariate and bivariate normality

At wave 1, 81% of the sample had nonclinical levels of internalizing and externalizing and only 5.3% of the sample exhibited some combination of at-risk or clinical levels of both internalizing and externalizing (Fig. 1a). All psychopathology dimensions were highly skewed, with a greater weight in the left tail than is expected in a normal distribution (Table 1; see Supplementary Fig. B8 and Supplementary Tables B4–6 for all other waves), kurtotic and not bivariate normal.

Associations between externalizing and internalizing

When we regressed externalizing onto internalizing, the change point occurred when internalizing was equal to a value of 4.85, corresponding to endorsement of approximately two to three internalizing symptoms. The pre-change point slope was strong ($B_1 = 0.28$), with a 1-point increase in internalizing indicating a 32% increase in externalizing (exp[0.28] of 1.32) (Fig. 2a,b). By contrast, the post-change point slope was relatively weak ($B_2 = 0.07$) (Fig. 2b), with a 1-point increase in internalizing only indicating a 7% increase in externalizing. The relative rate of the pre-change point slope was 4.6-times larger than that of the post-change point slope (32/7 = 4.6), suggesting far stronger slopes among largely asymptomatic cases who endorse fewer than three symptoms. Thus, nomorbidity, rather than comorbidity, drove the covariance commonly reported in literature.

In contrast with the one-change point model, the zero-change point model (that is, the ordinary least squares (OLS) estimate with a negative binomial link function) was associated with considerable prediction error (Fig. 3). The OLS estimate resulted in substantial prediction error, such that it overpredicted the association between externalizing and internalizing at nonclinical levels of internalizing (per the T scores), and particularly at symptomatic and clinical levels of internalizing.

We also considered two robustness checks: (1) the ordering of the predictor and criterion variables, and (2) replication across development. First, these patterns held when we regressed internalizing onto externalizing as opposed to the other way around (Fig. 2c). The change point occurred when externalizing was equal to a value of 2.5, corresponding to endorsement of one to two symptoms. Again, the pre-change point slope ($B_1 = 0.36$) was far stronger than the post-change point slope ($B_2 = 0.06$), with the former slope indicating that a 1-point change in externalizing was associated with a 43% increase in internalizing, and the latter indicating a 6% increase; the pre-change point slope was seven-times stronger than the post-change point slope.

Second, these patterns replicated across four waves (Fig. 2d). When we regressed externalizing onto internalizing across each wave, the location of the change points observed shifted slightly but not substantially across waves, with all change points being located between internalizing scores of 3 and 5 (corresponding to endorsement of one to five symptoms) and decreasing over time ($\tau_1 = 4.85$, $\tau_2 = 4.61$, $\tau_3 = 3.57$ and $\tau_4 = 3.53$). The pre- and post-change point slopes were consistent across waves ($B_1 = 0.28$ (wave 1 and 2), 0.38 (wave 3 and 4) and $B_2 = 0.07$ (wave 1–4)), with all waves suggesting much stronger slopes at low levels of internalizing. At waves 3 and 4, the pre-change point slopes were even stronger than previous waves, with a 1-point change in internalizing associated with a 46% increase in externalizing. These findings suggest that covariance among psychopathology dimensions may be increasingly driven by asymptomatic cases across development.

Associations among other higher-order psychopathology dimensions

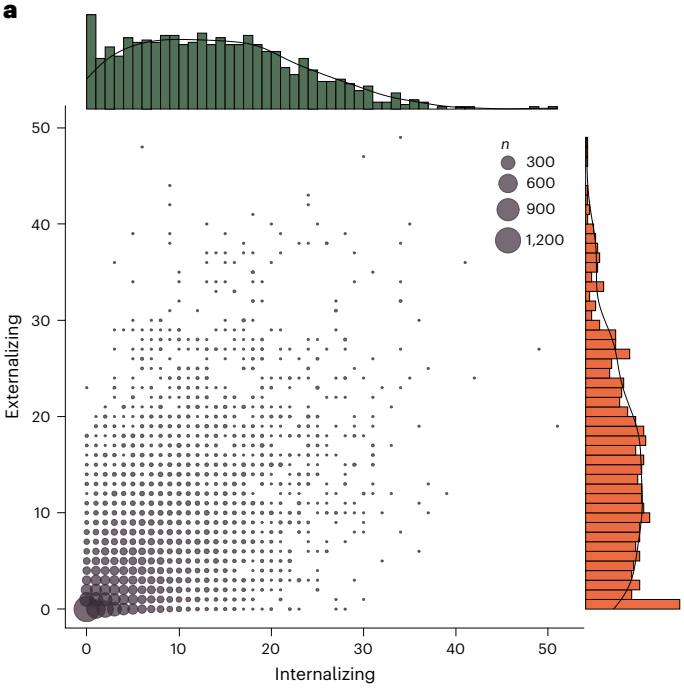
We also considered externalizing and internalizing's associations with other higher-order psychopathology dimensions: thought problems and attention problems (for wave 1 see Table 2 and Fig. 4; see Supplementary Tables B7–B11 for findings from all other waves). We observed the same trend across all models, whereby pre-change point slopes were far stronger than post-change point slopes, including when we regressed externalizing onto attention problems ($B_1 = 0.76$, $B_2 = 0.15$), externalizing onto thought problems ($B_1 = 0.54$, $B_2 = 0.15$), internalizing onto attention problems ($B_1 = 0.25$, $B_2 = 0.11$), internalizing onto thought problems ($B_1 = 0.47$, $B_2 = 0.13$) and attention problems onto thought problems ($B_1 = 0.52$, $B_2 = 0.12$).

Discussion

Researchers commonly make the inference that the correlation between two higher-order psychopathology dimensions captures

https://doi.org/10.1038/s44220-026-00727-0

a



b

	Classification combinations	Sample (%)
Nomorbidity	Nonclinical internalizing and externalizing	81.4
Internalizing	At-risk internalizing	4.8
	Clinical internalizing	3.3
Externalizing	At-risk externalizing	3.2
	Clinical externalizing	1.9
Comorbidity	At-risk internalizing and at-risk externalizing	0.9
	At-risk internalizing and clinical externalizing	0.9
	At-risk externalizing and clinical internalizing	1.3
	Clinical internalizing and externalizing	2.2

Fig. 1 | Distributions of externalizing and internalizing across levels of clinical severity. a, The bivariate distribution between externalizing and internalizing dimensions, with the marginal distributions for each on their respective axes. The size of the individual points corresponds to the number of cases at each point: the larger the point, the more cases. **b**, The percentage of cases that make up different categories of clinical severity for externalizing and internalizing based on the T-score categorizations for the Child Behavior Checklist. T scores below 60 correspond to nonclinical levels, between 60 and 63 correspond to 'at-risk' or borderline levels, and greater than 64 correspond to clinical levels. Supplementary Table C1 (Appendix C) contains a translation between T scores and raw scores for all scales across all waves.

comorbidity[16,21,39]. Nevertheless, we provide empirical evidence here that covariation among psychopathology dimensions is largely driven by nomorbidity, or a lack of comorbidity. Associations between psychopathology dimensions were strong and positive among asymptomatic people and substantially weaker among people who experience more than a handful of symptoms with relative frequency. That is, psychopathology dimensions were more differentiated (or less overlapping) at higher severity levels, to the extent that associations were negligible at moderate to clinically severe levels of psychopathology. Thus, covariance among psychopathology dimensions does not necessarily reflect or capture comorbidity[16,21,40]. Our findings underscore important limitations of covariance-based analyses for studying comorbidity, the *p*-factor[41–44] and the latent structure of psychopathology more generally.

Covariance-based similarity measures are dramatically influenced by the distributional qualities (for example, base rates and skewness) of the variables of interest[45], an issue that has been better appreciated in other fields (for example, ecology, computer science and engineering). Ecology, for instance, relies on association measures that exclude the mutual absences of rare species' locations[46] for the very reason we demonstrate here: traditional association measures (such as the correlation or factor analysis[34]) erroneously indicate that two species, or forms of psychopathology, are highly related simply because they do not co-occur often. The use of association measures that are driven by the absence of relatedness is particularly relevant to textbook definitions of comorbidity, which specifically emphasize a shared presence—rather than shared absence—of two or more disorders[47]. Covariance-based analyses, which include factor analyses of psychopathology data[6,16,48], in the general population may largely characterize the absence of different forms of psychopathology rather than their co-occurrence (or comorbidity).

A construct whose statistical and ontological status is particularly challenged by our findings is the *p*-factor, a dimension that is claimed to give rise to all forms of psychopathology and explains why different forms of psychopathology are associated with each other[21,22]. So it goes, the higher one's *p*-factor score, the higher propensity to develop several forms of psychopathology[21,22]. Critically, our findings suggest that the *p*-factor appears largely irrelevant to clinical levels of psychopathology and the subpopulation of people who experience psychopathology in some appreciable form, which is inconsistent with the notion that it is a unitary dimension on which all humans can be placed[21,22].

Common interpretations of the *p*-factor as the source of psychiatric comorbidity thus appear to rest on a faulty inference because the *p*-factor may not reflect comorbidity, but rather, an absence of clinically significant psychopathology. In this way, the *p*-factor appears more relevant to low as opposed to moderate or high levels of psychopathology. Although this conclusion may appear overly conjectural, related literature supports it. According to Spearman's law of diminishing returns (also termed 'cognitive ability differentiation'), the general factor of intelligence (or *g*)[49], the supposed theoretical predecessor to the *p*-factor[21,50], contains more explanatory power at lower as opposed to higher levels of intelligence[51–54].

Our findings also have implications for structural models of psychopathology that form the bedrock of modern classification efforts[6]. Such models, which typically rely on the general linear model, assume rather than test linear (or additive) relationships among psychopathology variables. Factor analysis, which is most commonly used to construct taxonomic frameworks[6,55], actually makes even stronger assumptions about linearity—that a latent normal distribution underlies the data[56], that the statistical relationship between a given item and the latent factor is linear, and that the correlations among latent factors are linear[6,57–59]. Even though factor analysis assumes an underlying normal distribution, estimated factor scores are not required to be normally distributed. In fact, in these data, factor scores of psychopathology dimensions were highly skewed and kurtotic, often even more so than the sum scores (Supplementary Fig. B9 and Supplementary Tables B12–15), and assumptions of linearity between psychopathology dimensions were violated in the data considered here (Supplementary Figs. B9 and B10).

Given apparent violations of linearity, when researchers implicitly or explicitly adopt a model with one slope in the population (that is, a zero-change point model), the slopes simply capture the approximate average of the two slopes estimated in a one-change point model (Appendix D). Although it may be tempting to think that the 'average' is a sufficient guess for the typical person in the population, that average has no psychological meaning when there are substantially different slopes in the population. Such an estimate is not only an oversimplification of the true population relationship, but a biased estimate of it[56,60–64]. Specifically, we showed that these estimates are upwardly biased, over characterizing the statistical association between higher-order psychopathology dimensions at nonclinical levels of psychopathology, and particularly at clinical levels. Because the

Table 1 | Univariate and bivariate normality for higher-order psychopathology dimensions at wave one

Predictor	Univariate normality		Bivariate normality		
	Skewness	Kurtosis	Criterion	Skewness	Kurtosis
Externalizing	2.30	7.02	Internalizing	18,326	203
			Attention problems	15,085	152
			Thought problems	19,115	211
Internalizing	1.94	5.12	Externalizing	18,326	203
			Attention problems	11,309	116
			Thought problems	16,804	195
Attention problems	1.44	1.83	Externalizing	15,085	152
			Internalizing	11,309	116
			Thought problems	13,436	137
Thought problems	2.26	6.88	Externalizing	19,115	211
			Internalizing	16,804	195
			Attention problems	13,436	137

Univariate skewness values >1 or <−1 indicate a high degree of skewness, and univariate kurtosis values >2 or <−2 indicate substantial nonnormality. Higher estimates of Mardia bivariate skewness and kurtosis indicate deviations from bivariate normality, with 0 indicating true multivariate normality. Bonferroni-corrected two-tailed *P* values associated with the Mardia statistics were significant at $P \approx 2.220446 \times 10^{-16}$, indicating that the data were significantly deviant from bivariate normality.

assumptions of the linear model were extremely violated in these data (Supplementary Figs. A9 and A10), we conducted a simulation study (Appendix D) to show that the linear model produces a biased, and often extremely misleading, estimate of the population association between two psychopathology dimensions in the presence of nonlinearity under many conditions.

Ultimately, there may be no single, 'true' structure of psychopathology in the population[65,66]. Instead, the assumption of population homogeneity may be inappropriate[66], and there may be discrete subpopulations of people that are differentially susceptible to psychopathology, have differential comorbidity patterns and, in turn, have dramatically different structures of psychopathology. Related literature supports such a possibility: self-esteem appears to have different factor structures—one unidimensional and one multidimensional—depending on the subpopulation of interest[67]. Further, the structure of intelligence appears moderated by ability level, with the strength of correlations among intelligence tasks decreasing as intelligence levels increase[68–70].

With the considerable degree of asymptomatic cases in most large-scale epidemiologic samples in mind, it may be the case that there is a large class with a relative absence of diagnosed psychopathology for whom psychopathology appears to conform to a unidimensional solution[35,36,71], and one or more relatively symptomatic classes for which psychopathology appears multidimensional. We suspect that there is a subpopulation that comprises people with an exceptionally high threshold to experience and/or accurately report psychopathology, whose detection is obscured—even precluded—in most modern classification efforts[72] owing to the widespread assumption of population homogeneity[73].

As the framers of the Diagnostic and Statistical Manual of Mental Disorders (DSM-6) are charged with capturing the continuum (or dimensionality) of mental health, and to explicitly consider severity and transdiagnostic features[74], our findings are highly relevant to current efforts to incorporate dimensionality into future psychiatric nosology[9,14]. In particular, our findings suggest that dimensional structures of psychopathology should be evaluated in relation to their target population and intended clinical use[75]. A structure estimated from covariance in a community sample may be useful for population screening or describing overall symptom burden, but it should not automatically be interpreted as evidence of clinical comorbidity, shared etiology or treatment-relevant co-occurrence among symptomatic patients. In community samples overrepresented by asymptomatic cases, global covariance may primarily reflect shared absence of symptoms. Thus, studies used to support dimensional nosology should report the distributional qualities of the data (for example, skewness and kurtosis), confirm that assumptions of the general linear model are met in their data, examine nonlinear associations among forms of psychopathology and evaluate whether broad structures remain informative in the clinical range. Our data suggest that the general linear model results in overprediction of covariance (or so-called comorbidity) among psychopathology dimensions at clinically severe levels of psychopathology.

Given the strong implications of our findings, it is worth discussing several potential limitations or criticisms of them. First, because we examined nonlinear relationships among higher-order psychopathology dimensions, skeptics of our findings might argue that we have invoked Berkson's bias[76] or the like[77,78] by restricting the sample post-change points to reflect higher degrees of severity (that is, conditioning on severity), thereby attenuating the associations between psychopathology dimensions at higher severity levels. That argument holds only if there is a single linear relationship between two psychopathology dimensions of interest in the population, or an absence of nonlinearity and population heterogeneity[73] (Appendix D). It is also worth noting that our change points were consistently located at extremely mild levels of psychopathology (for example, when two or three symptoms were endorsed), meaning that our results are not artifacts of restricting the sample to high severity levels.

Second, because we allowed for a change point in the associations between higher-order psychopathology dimensions, skeptics might argue that we artificially produced different slopes and, in turn, evidence of nonlinearity. Change point models do not inherently produce evidence of nonlinearity. Indeed, our simulation study (Appendix E) demonstrates that the change point model will generate estimates identical to the linear model when there is no evidence of nonlinearity. We also observed evidence of this fact in our analysis of psychopathology data. For instance, in the Gaussian one-change point model that regressed externalizing onto internalizing (Appendix C), the pre- and post-change point slopes were almost identical to each other, and to the slope estimated in the zero-change point model. Thus, it is not always the case that modeling a change point will engender evidence of nonlinearity.

Third, we relied on sum scores of youth psychopathology data. In sensitivity analyses, we examined caregivers' psychopathology data from the Adult Self-Report, the adult counterpart to the Child Behavior Checklist (CBCL), and our findings mirrored those reported here

https://doi.org/10.1038/s44220-026-00727-0

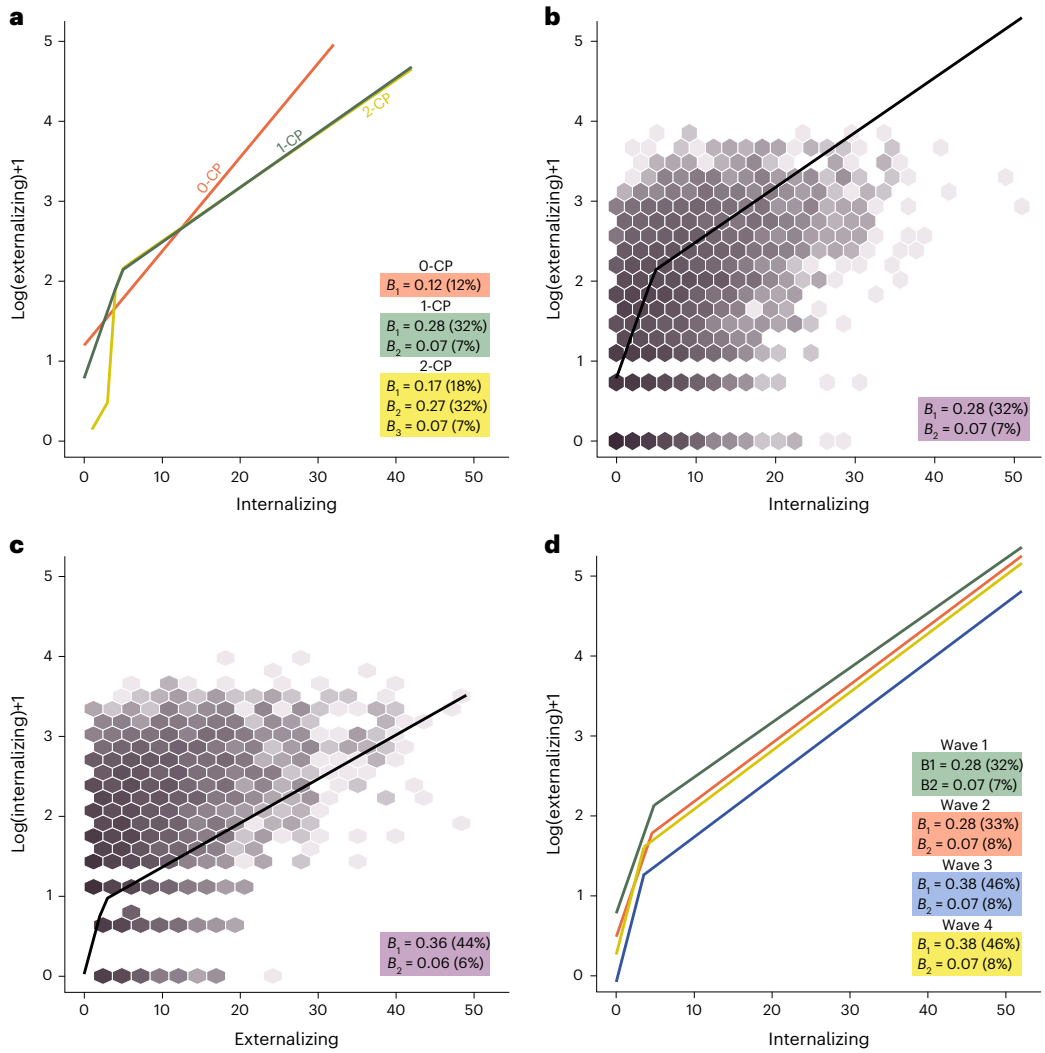


Fig. 2 | Associations between externalizing and internalizing dimensions from the negative binomial regression models at wave 1. a, The association across zero-, one- and two-change point models. **b**, The slopes of the associations from the one-change point model, with the bivariate density overlaid. **c**, The slopes from a one-change point model where internalizing was regressed onto externalizing. **d**, The associations across all waves. The *y* axis reflects log-transformed values of externalizing to facilitate interpretation of the slopes as linear and additive. **b** and **c** depict the density of the cases at each value of externalizing and internalizing in the bivariate distributions—the darker the color, the more cases at that pair of values. B_1 represents the unstandardized beta (slope) coefficient of the first slope estimated, B_2 represents the unstandardized beta (slope) coefficient of the second slope estimated, B_3 represents the unstandardized beta (slope) coefficient of the third slope estimated. The percentages next to the beta coefficients reflect the rate ratios, or the multiplicative factor (percentage) by which the expected count changes when the predictor increases by one point. For Internalizing, values lower than ~9 reflect nonclinical levels (per the T scores), between ~9 and ~14 reflect at-risk levels, and above ~14 reflect clinical levels. For externalizing, values lower than 12 reflect nonclinical levels, between ~12 and ~16 reflect at-risk levels, and above ~16 reflect clinical levels. The specific raw score that corresponds to a T score varies as a function of age and sex. Supplementary Table C1 contains a translation between raw scores and T scores for all scales across all waves. CP, change point.

(Supplementary Table C9). Change point models indicated a preference for one (as opposed to zero or two) change points, and the pre-change point slopes were considerably stronger than the post-change point slopes. It is perhaps unsurprising that our findings replicated among adults because the adult data in the ABCD Study were highly skewed and kurtotic (Supplementary Fig. B11 and Supplementary Tables B16 and B17). As demonstrated in our simulation study (Appendix E), any data are liable to the problems we demonstrated here: the associations among higher-order psychopathology dimensions—in youths and adults—are driven up with the inclusion of asymptomatic cases.

With that said, here, we focused only on the associations among higher-order psychopathology dimensions and did not conduct factor analyses of psychopathology, which is typical in classification efforts (see Appendix C for subscale-level analyses). As yet, there is no implementation of the Bayesian change point model for factor analysis. Still, the sum score is a special case of a latent variable[79], so our findings bear directly on factor models and the interpretation of correlations among higher-order psychopathology factors as evidence of comorbidity.

Fourth, given the highly positively skewed nature of our psychopathology dimensions, the ABCD data may be an inadequate representation of psychopathology at clinically severe levels. We offer several counterpoints to this potential conclusion. Although nonclinical cases reflected the plurality of data in the ABCD Study (Fig. 1), we had an appreciable amount of clinically severe levels of externalizing and internalizing, with 2.2% of youths (n = 257) falling into this category. Moreover, when we forced two-change point models, we did not find a change point at severe levels of psychopathology dimensions and did not find evidence of increasingly strong associations at severe levels (Appendix C). In one instance, when we regressed externalizing onto internalizing in the adult data, we detected a second change point at a

https://doi.org/10.1038/s44220-026-00727-0

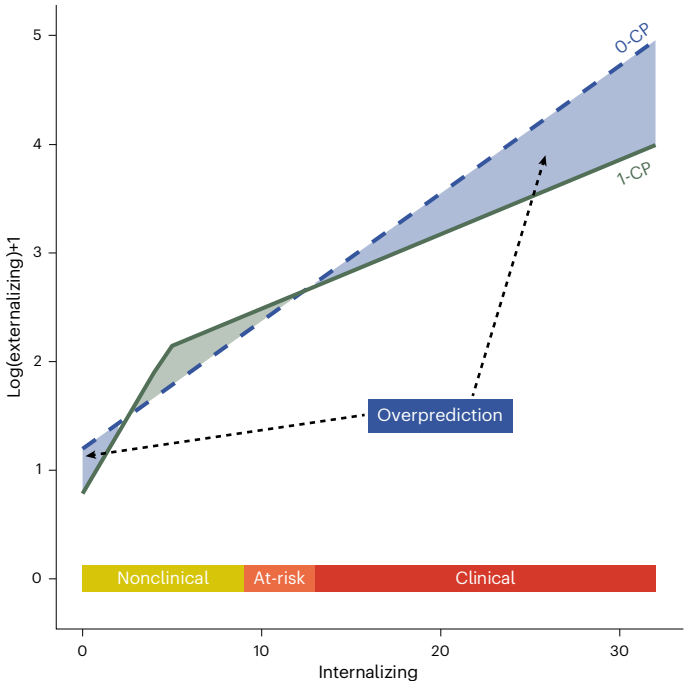


Fig. 3 | Prediction errors generated by the OLS regression model in the presence of nonlinearity. Here, we depict the same data as presented in Fig. 2a, with an emphasis on the negative binomial models with zero- and one-change points (the dashed blue and solid green lines, respectively), where we regressed externalizing onto internalizing. The zero-change point model corresponds to the OLS regression model most typically used in literature, albeit with a negative binomial link function as opposed to a Gaussian link function. By comparing the slopes from the OLS (zero-change point) and one-change point models, we see that the OLS estimate results in substantial prediction error (per the shaded regions of the plot) across levels of severity for internalizing, with much of the error occurring at clinical levels of severity. Shaded blue regions indicate where the OLS model is overestimating the association relative to what is predicted by the one-change point model, and shaded green regions indicate where the OLS model is underestimating the association. Severity levels are denoted by the bar that ranges from nonclinical levels (yellow) to at-risk levels (orange) to clinical levels (red). For Internalizing, values lower than -9 reflect nonclinical levels (per the T scores), between -9 and -14 reflect at-risk levels, and above -14 reflect clinical levels (see Supplementary Table C1 for a translation between raw scores and T scores). Thus, because of the considerable extent of nonclinical (and asymptomatic cases) in the data, the OLS estimate generally overstates the statistical association between externalizing and internalizing at nonclinical levels of internalizing, and especially at symptomatic and clinically severe levels of internalizing.

score of 30 (Supplementary Table C9). Still, the three slopes detected in this model suggested a decreasing association between externalizing and internalizing as a function of severity ($B_1 = 0.29$, $B_2 = 0.12$ and $B_3 = 0.04$).

Along those lines, we used an alternate estimation method, generalized additive multilevel models (Appendix C), that may be more sensitive to the upper tail of the distribution, and still found no support for this hypothesis. Thus, we are not convinced that our two-change point models were preordained to reject the notion that higher-order psychopathology dimensions become more strongly associated at severe levels[21]. Instead, there was no uptick for the third estimated slope, ruling out the possibility that comorbidity 'emerges' at severe levels of psychopathology.

We welcome investigations of nonlinearity among samples with more representation of severe levels of psychopathology, such as in clinical or mixed community–clinical samples, and differing assessments of psychopathology, including ones that assess finer-grained phenomena than we consider here. Although not of focus here, it is possible that the dimensionality of narrower phenomena is also driven by asymptomatic cases, and that higher-order psychopathology dimensions may appear more multidimensional at higher levels of psychopathology. For instance, perhaps depression and anxiety are most overlapping at low levels and less overlapping at higher levels (Appendix C).

Mitigating potential generalizability concerns, a recent meta-analysis[80] indicates that there is strong reason to believe that higher-order psychopathology dimensions become more weakly—as opposed to more strongly—associated at higher levels of severity. The correlation between externalizing and internalizing is dramatically stronger in nonenriched (or more nomorbid) samples ($\Phi = 0.48$) compared with psychopathology enriched samples (or samples with more comorbidity, $\Phi = 0.20$). This pattern holds for assessment strategies that manipulate the extent of positive skew or nomorbidity in the data[35,36,71]; the correlation between externalizing and internalizing was much stronger when current diagnoses ($\Phi = 0.37$), as opposed to lifetime diagnoses ($\Phi = 0.19$), were modeled, and it was also much stronger when studies did not rely on hierarchical exclusion (or 'skip-out') rules ($\Phi = 0.41$) compared with those that did ($\Phi = 0.24$)[81]. Compared with their counterparts, current diagnoses and diagnoses that did not rely on hierarchical exclusion rules are more strongly skewed[82,83] and hence contain more asymptomatic cases[36,84].

Conclusion

Our findings challenge the prevailing assumption that statistical covariance among forms of psychopathology should be taken as evidence of psychiatric comorbidity. Just as correlation does not imply causation, correlation does not imply comorbidity. The inclusion of asymptomatic, or nomorbid, cases in covariance-based analyses of psychopathology created the illusion of strong covariation among higher-order psychopathology dimensions. At higher levels of psychopathology, higher-order psychopathology dimensions were far more differentiated than is implied by the existing literature[16,17,85,86], providing weak evidence of a *p*-factor (or general factor of psychopathology). Ignoring the implications of the inclusion of largely asymptomatic cases in psychopathology research could jeopardize the integrity and utility of psychiatric classification systems[6,9,14,74].

More generally, we call for a reconsideration of the uncritical adoption of the general linear model in psychopathology research, which often produces comorbidity-inference artifacts in community-level data, where most individuals do not exhibit clinically severe psychopathology. We encourage researchers to embrace nonlinearity and population heterogeneity and adopt conceptual and analytic frameworks that prioritize symptomatology over statistical correlation. An exaggerated sense of comorbidity among distinct forms of psychopathology has, and will continue to, perpetuate the reification of artifactual transdiagnostic dimensions[42,87,88] and cease investigation of disorder-specific mechanisms[21], fundamentally impeding the field's detection of psychopathology's etiologic mechanisms, treatment targets and consequences.

Methods

Data

Data were drawn from the first four waves of ABCD (data release 5.0), a collaboration between 21 sites across the USA[38,89] that tracks a convenience sample of 11,868 nine- and ten-year-olds from the USA annually across development (wave 1: $M_{age} = 9.94$ years, s.d. 0.62 years; 48% female). The ABCD data are not nationally representative, but the study team endeavored to make the sample as representative of the US population as possible. Overall, 74% of youth participants identified as white (14% Hispanic), 21% as Black (2% Hispanic), 7% as Asian and 11% as 'other' (see Supplementary Table B1 in Appendix B for demographics of all other waves). In total, 20% identified as Hispanic. These data are

Table 2 | Results from the one-change point negative binomial regressions for all higher-order psychopathology dimensions at wave 1

Predictor variable	Criterion variable	Change point location			Pre-change point slope				Post-change point slope				Ratio of slopes
		τ	SE	$\hat{R}$	B_1	SE	(%) exp(B_1)	$\hat{R}$	B_2	SE	(%) exp(B_2)	$\hat{R}$	
Internalizing	Externalizing	4.85	0.07	1.04	0.28	0.00	32	1.05	0.07	0.00	7	1.03	4.5
	Attention problems	3.82	0.05	1.02	0.27	0.00	31	1.03	0.06	0.00	6	1.01	5.0
	Thought problems	5.26	0.06	1.04	0.27	0.00	31	1.03	0.07	0.00	7	1.03	4.3
Externalizing	Internalizing	2.52	0.01	1.00	0.36	0.00	43	1.00	0.06	0.00	6	1.00	7.0
	Attention problems	3.48	0.02	1.01	0.37	0.00	45	1.02	0.06	0.00	6	1.01	7.2
	Thought problems	5.15	0.09	1.07	0.24	0.00	27	1.06	0.06	0.00	6	1.04	4.4
Attention problems	Externalizing	1.50	0.00	1.00	0.76	0.00	114	1.00	0.15	0.00	16	1.00	7.0
	Internalizing	1.07	0.05	1.02	0.25	0.05	28	1.03	0.11	0.00	12	1.02	2.4
	Thought problems	1.51	0.01	1.00	0.54	0.00	72	1.00	0.15	0.00	16	1.00	4.4
Thought problems	Externalizing	2.51	0.01	1.00	0.52	0.00	68	1.00	0.15	0.00	16	1.00	4.2
	Internalizing	2.50	0.00	1.00	0.47	0.00	60	1.00	0.13	0.00	14	1.00	4.3
	Attention problems	2.51	0.01	1.00	0.52	0.00	68	1.01	0.12	0.00	13	1.01	5.3

B, unstandardized beta (slope) coefficients; exp(B_1), exponentiated beta coefficient; $\hat{R}$, scale reduction factor (convergence criterion); SE, standard error; τ, the change point location. The exponentiated beta coefficients reflect the percent increase in the criterion variable per point of the predictor variable; for instance, when we regressed externalizing onto internalizing, B_1=0.28, so exp(B_1)=exp(0.28)=1.32, or a 32% increase. The ratios of slopes are computed as follows: % exp(B_1) / % exp(B_2). Table C1 contains a translation between raw scores and T scores for all scales across all waves.

available to researchers upon receipt of a data use agreement from the National Institutes of Health (https://nda.nih.gov/).

Measures

We used the caregiver-reported CBCL[90], a 113-item assessment of psychiatric symptoms. The CBCL is among the most widely used broadband assessments of youth psychopathology, with over 10,000 publications reporting its use across over 50 different countries[91]. Symptoms are rated on a three-point scale ('not true', 'somewhat true/sometimes true' and 'very true/often true'). We used summary score composites for all psychopathology variables. For readers interested in T scores, Supplementary Table C1 (Appendix C) contains a translation between raw scores and manual-based T scores for all scales across all waves. Of primary focus here were the two broadband psychopathology scores, externalizing and internalizing (at wave 1, coefficient omegas were 0.89 and 0.87, respectively) (see Supplementary Table B2 for omegas for all other waves).

We also considered two subscales that contain items not included in externalizing and internalizing: attention problems and thought problems (coefficient omegas were 0.86 and 0.70, respectively), because they map onto broad transdiagnostic dimensions in modern classification frameworks[18,92]. In addition to externalizing and internalizing spectra, modern frameworks also include neurodevelopmental[93] and psychosis[92] (or thought disorder) spectra, which map broadly onto attention problems and thought problems, respectively. As the field incorporates lesser studied higher-order psychopathology dimensions into hierarchical frameworks of psychopathology, we felt that broadening our focus to include these dimensions would be of relevance to the field.

Readers interested in the interface among lower-order psychopathology dimensions within a higher-order dimension (that is, anxiety and depression within internalizing) are directed to Supplementary Table C8 in Appendix C, where we show the results for all change point models based on CBCL subscales. In brief, results were consistent with what we report here, with a few exceptions: across models, the one-change point model was superior to others, and pre-change point slopes were far stronger than post-change point slopes.

Analyses

Univariate and bivariate skew and kurtosis. First, we estimated univariate skewness and kurtosis for all psychopathology dimensions to inform deviations from a normal distribution. A normal distribution's skewness is 0, with positive values indicating positive skew, and its kurtosis is 3, with values greater than 3 suggesting leptokurtosis. Second, we examined bivariate skewness and kurtosis[84], which inform how bivariate relationships deviate from a theoretical multivariate normal distribution. For bivariate estimates, we report Mardia's chi-squared tests, whose *P* values indicate whether the data satisfy the assumptions of multivariate normality. We considered all unique pairwise comparisons (that is, six) among psychopathology dimensions across all four waves, and Bonferroni corrected all *P* values to penalize for the number of tests conducted (24, or 6 × 4).

Bayesian change point models. We used Bayesian change point models to examine whether the relationship between psychopathology dimensions changes at different points along their joint distribution, allowing for nonlinear relationships. Change point models have other common names, including broken-stick models[94], or turning point[95] or break point models[96]. They are used routinely in ecology to identify transitions in animal behavior[97] and environmental thresholds[98] in cross-sectional data, as well as in time-series analysis[99]. They are conceptually similar to regression discontinuity designs[100].

The Bayesian change point model identifies locations, or change points, along the predictor variable where its relationship with the criterion variable may change. For instance, the relationship between x and y might be strong and positive before ten units of x, at which point a change point occurs, and the relationship may become weak and negative after ten units of x. By contrast, an OLS regression assumes that the relationship is equivalent, or does not change, across the joint distribution (in Appendix D we show, via a simulation study, that the OLS estimate is biased in the presence of nonlinearity).

The number of change points can range from zero to infinity. A one-change point model adheres to the following structure:

$$y_i = \begin{cases} \alpha_1 + \beta_1 X, & x < \tau_r \\ \alpha_2 + \beta_2 X, & x \geq \tau_r \end{cases},$$

where τ_r represents the location of the change point (or the value of the predictor variable where the slope changes) and where the α and β terms are the intercept and slope terms, respectively. That is, the one-change point model is a combination of two generalized linear

https://doi.org/10.1038/s44220-026-00727-0

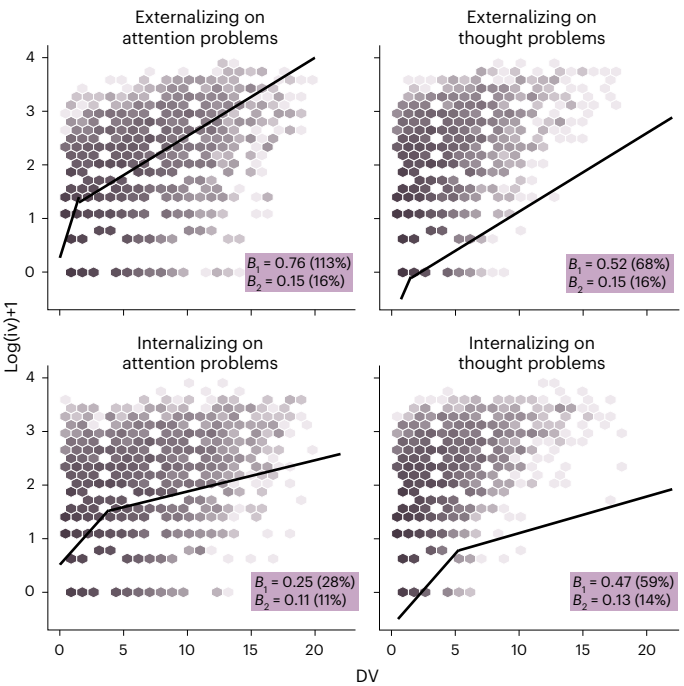


Fig. 4 | Externalizing and internalizing's associations with other higher-order psychopathology dimensions from the one-change point negative binomial regression models at wave 1. B_1 represents the unstandardized beta (slope) coefficient of the first slope estimated in change point models. B_2 represents the unstandardized beta (slope) coefficient of the second slope estimated. The percentages next to the beta coefficients reflect the rate ratios or the multiplicative factor (percentage) by which the expected count changes when the predictor increases by one point. In these models, the dependent variable corresponds to the first higher-order dimension listed in the plot title and the independent variable corresponds to the second (for example, 'attention problems' is the independent variable in the plot titled 'externalizing on attention problems'). For externalizing, values lower than 12 reflect nonclinical levels, between ~12 and ~16 reflect at-risk levels, and above ~16 reflect clinical levels. For internalizing, values lower than ~9 reflect nonclinical levels (per the T scores), between ~9 and ~14 reflect at-risk levels, and above ~14 reflect clinical levels. For attention problems, values lower than ~7 reflect nonclinical levels, between ~7 and ~8 reflect at-risk levels, and above ~9 reflect clinical levels. For thought problems, values lower than ~3 reflect nonclinical levels, between ~3 and ~4 reflect at-risk levels, and above ~4 reflect clinical levels. The specific raw score that corresponds to a T score varies as a function of age and sex. Supplementary Table C1 contains a translation between raw scores and T scores for all scales across all waves. DV, dependent variable; IV, independent variable.

models, with each model estimating an intercept and slope before and after the change point. Because the change point model is a generalized linear model, it can be extended to other extensions of the generalized linear model (for example, hierarchical and generalized linear models).

Model estimation. We allowed for up to two change points and compared two link functions, the Gaussian link and the negative binomial link. We used the Gaussian link given its overwhelming popularity, and we used the negative binomial link because it should be more appropriate given the positively skewed (or overdispersed) count data used here[101,102].

In contrast with the Gaussian link-based regression, negative binomial regression assumes a nonlinear (log–linear) relationship between the predictor variables and the expected count of the criterion variable, whereby change in the predictor is multiplicative. The beta coefficient (or slope, B) represents the change in the log of the expected count for a 1-point increase in the predictor variable. To translate the nonlinear estimates, one can exponentiate the slope (exp[B]) so that it reflects the rate ratio or the multiplicative factor by which the expected count changes when the predictor increases by one point.

Every model used Markov Chain Monte Carlo sampling with a burn-in length of 10,000, a total chain length of 20,000 and a thinning of every second sample to control for potential autocorrelation. All sampling was performed using the Stan language and the No-U-Turn Sampler algorithm[103,104]. We selected diffuse and naive priors. Estimation of the change point(s) used uniform priors that ranged from the minimum to the maximum value of the domain, and estimation of the slope(s) and intercept(s) used normal priors centered on 0 with a variance of 1. Random effect variance used uniform priors ranging from 0 to 10. In terms of other covariates, we included a fixed effect for age and nested random effects for complex survey design features (family within site).

Model convergence. We examined model convergence using the potential scale reduction factor $\hat{R}$ (ref. 105) to determine if the chains mixed. $\hat{R}$ measures the ratio of the average variance of samples within each chain to the variance of the pooled samples across chains. If all chains are at equilibrium, the average variance of samples within each chain will be equal to the variance of the pooled samples across chains, producing an $\hat{R}$ of 1.0. Values less than 1.2 generally indicate convergence across chains[105]. To confirm strong model convergence, we also visually inspected the trace, density and autocorrelation plots (Supplementary Figs. B1–B4; Supplementary Table 2 contains $\hat{R}$ values for the slope parameters of interest here; and Appendix C contains all model parameters, including $\hat{R}$ values, for all models).

Model comparison. We prioritized the Leave-One-Out Information Criterion (LOOIC), which is commonly used for Bayesian model selection[106]. LOOIC facilitates the comparison of models within and across link functions, with lower values indicating a better-fitting model. We also used model R^2 (or the proportion of variability explained).

Of the 288 models compared—Gaussian and negative binomial, each with either zero, one or two change points—LOOIC supported one-change point negative binomial models over Gaussian models (Supplementary Fig. B5 and Supplementary Table B3). Additional sources of evidence converged on the inappropriateness of the Gaussian models, including that the data were not univariate or bivariate normal, and that model diagnostics of linear regressions suggested substantial violations of linear regression assumptions, including linearity, homogeneity of variance and homogeneity in residuals (Supplementary Figs. B9 and B10). Thus, results from the Gaussian models are invalid.

Among the negative binomial models, the two-change point models were over extracted, with little to no increment in R^2 above and beyond the one-change point models (Supplementary Fig. A7). For instance, when we regressed externalizing onto internalizing at wave 1, model R^2 (or variance explained in externalizing) was 21% in the zero-change point model, 56% in the one-change point model and 55% in the two-change point model.

We anticipated that a two-change point model might reveal a change in the slope between internalizing and externalizing at high levels of internalizing (or high levels of severity), with the slope strengthening and reflecting increased comorbidity. Instead, the second change point in this model was located at $\tau_2 = 5.09$, corresponding to endorsement of around three or four internalizing symptoms. Further, the slope after the second change point was lower than the slope after the first change point ($B_1 = 0.17$, $B_2 = 0.27$ and $B_3 = 0.07$, respectively), which is inconsistent with the notion that their association would increase as a function of presumed comorbidity. There were no instances in the two-change point negative binomial models where the slope after the second change point (B_3) increased relative to the prior slope (B_2). Given the lack of incremental R^2 for the two-change point models, we carried forward all one-change point negative binomial models (Fig. 2a).

We considered an alternative estimation method, generalized additive multilevel models[107,108], to examine nonlinear associations between higher-order psychopathology dimensions, as in previous work[109]. We detected the same nonlinear associations using this method and did not find consistent evidence that there is a strong positive association between externalizing and internalizing at extreme high levels of internalizing (Appendix E).

Reporting summary

Further information on research design is available in the Nature Portfolio Reporting Summary linked to this article.

Data availability

Data used in the preparation of this article were obtained from the ABCD Study (https://abcdstudy.org), held in the NIMH Data Archive. This is a multisite, longitudinal study designed to recruit more than 10,000 children aged 9–10 and follow them over 10 years into early adulthood. The ABCD Study is supported by the National Institutes of Health and additional federal partners under award nos. U01DA041048, U01DA050989, U01DA051016, U01DA041022, U01DA051018, U01DA051037, U01DA050987, U01DA041174, U01DA041106, U01DA041117, U01DA041028, U01DA041134, U01DA050988, U01DA051039, U01DA041156, U01DA041025, U01DA041120, U01DA051038, U01DA041148, U01DA041093, U01DA041089, U24DA041123 and U24DA041147. A full list of supporters is available at https://abcdstudy.org/federal-partners.html. A listing of participating sites and a complete listing of the study investigators can be found at https://abcdstudy.org/consortium_members/. ABCD consortium investigators designed and implemented the study and/or provided data but did not necessarily participate in the analysis or writing of this report. This manuscript reflects the views of the authors and may not reflect the opinions or views of the NIH or ABCD consortium investigators. The ABCD data repository grows and changes over time. The ABCD data used in this report are available via the NDA at https://doi.org/10.15154/8873-zj65 (ref. 110). These data are available to researchers upon receipt of a data use agreement from the National Institutes of Health (https://nda.nih.gov/).

Code availability

Our results are fully reproducible (upon receipt of the ABCD data) via our code, which is avilable via OSF at https://osf.io/nr5h3/overview and via GitHub at https://github.com/Watts-MAP-Lab/psychopathology_nonlinearity.

References

1. First, M. B. Mutually exclusive versus co-occurring diagnostic categories: the challenge of diagnostic comorbidity. *Psychopathology* **38**, 206–210 (2005).
2. Widiger, T. A. & Frances, A. The DSM-III personality disorders: perspectives from psychology. *Arch. Gen. Psychiatry* **42**, 615 (1985).
3. Kessler, R. C., Chiu, W. T., Demler, O. & Walters, E. E. Prevalence, severity, and comorbidity of 12-month DSM-IV disorders in the national comorbidity survey replication. *Arch. Gen. Psychiatry* **62**, 617–627 (2005).
4. Deighton, J., Lereya, S. T. & Wolpert, M. Enduring mental health in childhood and adolescence: learning from the Millennium Cohort Study. *J. Am. Acad. Child Adolesc. Psychiatry* **60**, 1030–1039 (2021).
5. Schaefer, J. D. et al. Enduring mental health: prevalence and prediction. *J. Abnorm. Psychol.* **126**, 212–224 (2017).
6. Kotov, R. et al. The Hierarchical Taxonomy of Psychopathology (HiTOP): a dimensional alternative to traditional nosologies. *J. Abnorm. Psychol.* **126**, 454–477 (2017).
7. Hankin, B. L. et al. Understanding comorbidity among internalizing problems: integrating latent structural models of psychopathology and risk mechanisms. *Dev. Psychopathol.* **28**, 987–1012 (2016).
8. Clark, L. A., Nuzum, H. & Ro, E. Manifestations of personality impairment severity: comorbidity, course/prognosis, psychosocial dysfunction, and 'borderline' personality features. *Curr. Opin. Psychol.* **21**, 117–121 (2018).
9. Conway, C. C. et al. A hierarchical taxonomy of psychopathology can transform mental health research. *Perspect. Psychol. Sci.* **14**, 419–436 (2019).
10. Beauchaine, T. P. & Klein, D. N. Classifying psychopathology: the DSM, empirically based taxonomies, and the research domain criteria. In *Child and Adolescent Psychopathology* 3rd edn (eds Beauchaine, T. P. & Hinshaw S. P.) 33–67 https://doi.org/10.1002/9781394258932.ch2 (Wiley, 2017).
11. Wright, A. G. C. & Simms, L. J. A metastructural model of mental disorders and pathological personality traits. *Psychol. Med.* **45**, 2309–2319 (2015).
12. Carragher, N., Krueger, R. F., Eaton, N. R. & Slade, T. Disorders without borders: current and future directions in the meta-structure of mental disorders. *Soc. Psychiatry Psychiatr. Epidemiol.* **50**, 339–350 (2015).
13. Conway, C. C., Forbes, M. K. & South, S. C. A Hierarchical Taxonomy of Psychopathology (HiTOP) primer for mental health researchers. *Clin. Psychol. Sci.* **10**, 236–258 (2022).
14. Cicero, D. C. et al. State of the science: the Hierarchical Taxonomy of Psychopathology (HiTOP). *Behav. Ther.* **55**, 1114–1129 (2024).
15. Achenbach, T. M. The classification of children's psychiatric symptoms: s factor-analytic study. *Psychol. Monogr. Gen. Appl.* **80**, 1–37 (1966).
16. Krueger, R. F. & Markon, K. E. Reinterpreting comorbidity: a model-based approach to understanding and classifying psychopathology. *Annu. Rev. Clin. Psychol.* **2**, 111–133 (2006).
17. Krueger, R. F. The structure of common mental disorders. *Arch. Gen. Psychiatry* **56**, 921–926 (1999).
18. Krueger, R. F. et al. Validity and utility of Hierarchical Taxonomy of Psychopathology (HiTOP): II. Externalizing superspectrum. *World Psychiatry* **20**, 171–193 (2021).
19. Watson, D. et al. Validity and utility of Hierarchical Taxonomy of Psychopathology (HiTOP): III. Emotional dysfunction superspectrum. *World Psychiatry* **21**, 26–54 (2022).
20. Krueger, R. F., Markon, K. E., Patrick, C. J., Benning, S. D. & Kramer, M. D. Linking antisocial behavior, substance use, and personality: an integrative quantitative model of the adult externalizing spectrum. *J. Abnorm. Psychol.* **116**, 645–666 (2007).
21. Caspi, A. & Moffitt, T. E. All for one and one for all: mental disorders in one dimension. *Am. J. Psychiatry* **175**, 831–844 (2018).
22. Lahey, B. B., Krueger, R. F., Rathouz, P. J., Waldman, I. D. & Zald, D. H. Validity and utility of the general factor of psychopathology. *World Psychiatry* **16**, 142–144 (2017).
23. Smith, G. T., Atkinson, E. A., Davis, H. A., Riley, E. N. & Oltmanns, J. R. The general factor of psychopathology. *Annu. Rev. Clin. Psychol.* **16**, 75–98 (2020).
24. Romer, A. L. et al. Replicability of structural brain alterations associated with general psychopathology: evidence from a population-representative birth cohort. *Mol. Psychiatry* **26**, 3839–3846 (2021).
25. Gluschkoff, K., Jokela, M. & Rosenström, T. General psychopathology factor and borderline personality disorder: evidence for substantial overlap from two nationally representative surveys of U.S. adults. *Personal Disord.* **12**, 86–92 (2021).
26. Schweizer, T. H., Snyder, H. R., Young, J. F. & Hankin, B. L. The breadth and potency of transdiagnostic cognitive risks for psychopathology in youth. *J. Consult. Clin. Psychol.* **88**, 196–211 (2020).

27. Meier, M. A. & Meier, M. H. Clinical implications of a general psychopathology factor: a cognitive–behavioral transdiagnostic group treatment for community mental health. *J. Psychother. Integr.* **28**, 253–268 (2018).
28. Pettersson, E., Lahey, B. B., Larsson, H. & Lichtenstein, P. Criterion validity and utility of the general factor of psychopathology in childhood: predictive associations with independently measured severe adverse mental health outcomes in adolescence. *J. Am. Acad. Child Adolesc. Psychiatry* **57**, 372–383 (2018).
29. McElroy, E., Belsky, J., Carragher, N., Fearon, P. & Patalay, P. Developmental stability of general and specific factors of psychopathology from early childhood to adolescence: dynamic mutualism or *p* -differentiation?. *J. Child Psychol. Psychiatry* **59**, 667–675 (2018).
30. Carver, C. S., Johnson, S. L. & Timpano, K. R. Toward a functional view of the *p* factor in psychopathology. *Clin. Psychol. Sci.* **5**, 880–889 (2017).
31. Shields, A. N., Reardon, K. W., Brandes, C. M. & Tackett, J. L. The p factor in children: relationships with executive functions and effortful control. *J. Res. Pers.* **82**, 103853 (2019).
32. Brandes, C. M., Herzhoff, K., Smack, A. J. & Tackett, J. L. The *p* factor and the *n* factor: associations between the general factors of psychopathology and neuroticism in children. *Clin. Psychol. Sci.* **7**, 1266–1284 (2019).
33. Tackett, J. L. et al. Common genetic influences on negative emotionality and a general psychopathology factor in childhood and adolescence. *J. Abnorm. Psychol.* **122**, 1142–1153 (2013).
34. Christensen, W. F., Wall, M. M. & Moustaki, I. Assessing dimensionality in dichotomous items when many subjects have all-zero responses: an example from psychiatry and a solution using mixture models. *Appl. Psychol. Meas.* **46**, 167–184 (2022).
35. Hoffman, M. et al. The influence of sample selection on the structure of psychopathology symptom networks: an example with alcohol use disorder. *J. Abnorm. Psychol.* **128**, 473–486 (2019).
36. Steppan, M., Piontek, D. & Kraus, L. The effect of sample selection on the distinction between alcohol abuse and dependence. *Int. J. Alcohol Drug Res.* **3**, 2 (2014).
37. Brusco, M., Cradit, J. D. & Steinley, D. A comparison of 71 binary similarity coefficients: the effect of base rates. *PLoS ONE* **16**, e0247751 (2021).
38. Jernigan, T. L., Brown, S. A. & Dowling, G. J. The adolescent brain cognitive development study. *J. Res. Adolesc.* **28**, 154–156 (2018).
39. Lahey, B. B. et al. Is there a general factor of prevalent psychopathology during adulthood?. *J. Abnorm. Psychol.* **121**, 971–977 (2012).
40. Forbes, M. K. Implications of the symptom-level overlap among DSM diagnoses for dimensions of psychopathology. *J. Emot. Psychopathol.* **1**, 104–112 (2023).
41. Watts, A. L. et al. How robust is the *p* factor? Using multitrait-multimethod modeling to inform the meaning of general factors of youth psychopathology. *Clin. Psychol. Sci.* **10**, 640–661 (2022).
42. Watts, A. L., Greene, A. L., Bonifay, W. & Fried, E. I. A critical evaluation of the *p*-factor literature. *Nat. Rev. Psychol.* **3**, 108–122 https://doi.org/10.1038/s44159-023-00260-2 (2024).
43. Watts, A. L., Poore, H. E. & Waldman, I. D. Riskier tests of the validity of the bifactor model of psychopathology. *Clin. Psychol. Sci.* **7**, 1285–1303 (2019).
44. Watts, A. L., Lane, S. P., Bonifay, W., Steinley, D. & Meyer, F. A. C. Building theories on top of, and not independent of, statistical models: the case of the *p*-factor. *Psychol. Inq.* **31**, 310–320 (2020).
45. Warrens, M. J. On similarity coefficients for 2×2 tables and correction for chance. *Psychometrika* **73**, 487–502 (2008).
46. Janson, S. & Vegelius, J. Measures of ecological association. *Oecologia* **49**, 371–376 (1981).
47. Lilienfeld, S. O., Waldman, I. D. & Israel, A. C. A critical examination of the use of the term and concept of comorbidity in psychopathology research. *Clin. Psychol. Sci. Pract.* **1**, 71–83 (1994).
48. Eaton, N. R. et al. A review of approaches and models in psychopathology conceptualization research. *Nat. Rev. Psychol.* **2**, 622–636 (2023).
49. Spearman, C. General Intelligence,' objectively determined and measured. *Am. J. Psychol.* **15**, 201 (1904).
50. Caspi, A. et al. The *p* factor: one general psychopathology factor in the structure of psychiatric disorders?. *Clin. Psychol. Sci.* **2**, 119–137 (2014).
51. Spearman, C. *The Abilities Of Man* http://archive.org/details/abilitiesofman031969mbp (Macmillan, 1927).
52. Detterman, D. K. & Daniel, M. H. Correlations of mental tests with each other and with cognitive variables are highest for low IQ groups. *Intelligence* **13**, 349–359 (1989).
53. Maxwell, A. E. The WPPSI: a marked discrepancy in the correlations of the subtests for good and poor readers. *Br. J. Math. Stat. Psychol.* **25**, 283–291 (1972).
54. Jensen, A. R. Regularities in Spearman's law of diminishing returns. *Intelligence* **31**, 95–105 (2003).
55. Haeffel, G. J. et al. Folk classification and factor rotations: whales, sharks, and the problems with the Hierarchical Taxonomy of Psychopathology (HiTOP). *Clin. Psychol. Sci.* **10**, 259–278 (2022).
56. Reise, S. P., Du, H., Wong, E. F., Hubbard, A. S. & Haviland, M. G. Matching IRT models to patient-reported outcomes constructs: the graded response and log-logistic models for scaling depression. *Psychometrika* **86**, 800–824 (2021).
57. Flora, D. B. & Curran, P. J. An empirical evaluation of alternative methods of estimation for confirmatory factor analysis with ordinal data. *Psychol. Methods* **9**, 466–491 (2004).
58. Preacher, K. J. & MacCallum, R. C. Repairing Tom Swift's electric factor analysis machine. *Understand. Stat.* **2**, 13–43 (2003).
59. Emmons, O. *Linearity In Factor Analysis* https://doi.org/10.58809/VCNR9865 (Fort Hays State Univ., 1939).
60. Wall, M. M., Park, J. Y. & Moustaki, I. IRT modeling in the presence of zero-inflation with application to psychiatric disorder severity. *Appl. Psychol. Meas.* **39**, 583–597 (2015).
61. Kirisci, L., Hsu, T. chi & Yu, L. Robustness of item parameter estimation programs to assumptions of unidimensionality and normality. *Appl. Psychol. Meas.* **25**, 146–162 (2001).
62. Azevedo, C. L. N., Bolfarine, H. & Andrade, D. F. Bayesian inference for a skew-normal IRT model under the centred parameterization. *Comput. Stat. Data Anal.* **55**, 353–365 (2011).
63. Field, A. P. & Wilcox, R. R. Robust statistical methods: a primer for clinical psychology and experimental psychopathology researchers. *Behav. Res. Ther.* **98**, 19–38 (2017).
64. Reise, S. P., Rodriguez, A., Spritzer, K. L. & Hays, R. D. Alternative approaches to addressing non-normal distributions in the application of IRT models to personality measures. *J. Pers. Assess.* **100**, 363–374 (2018).
65. Haslam, N. The latent structure of mental disorders: a taxometric update on the categorical vs dimensional debate. *Curr. Psychiatry Rev.* **3**, 172–177 (2007).
66. Meehl, P. E. & Golden, R. Taxometric methods. In *Handbook of Research Methods in Clinical Psychology* Ch. 5 (eds Kendall, P. & Butcher, J.) 127–181 (Wiley, 1982).
67. Kam, C. C. S. Bifactor model is not the best-fitting model for self-esteem: investigation with a novel technique. *Assessment* **28**, 1799–1812 https://doi.org/10.1177/1073191120949916 (2021).
68. Molenaar, D., Dolan, C. V., Wicherts, J. M. & van der Maas, H. L. J. Modeling differentiation of cognitive abilities within the higher-order factor model using moderated factor analysis. *Intelligence* **38**, 611–624 (2010).

69. Molenaar, D., Kő, N., Rózsa, S. & Mészáros, A. Differentiation of cognitive abilities in the WAIS-IV at the item level. *Intelligence* **65**, 48–59 (2017).
70. Reynolds, M. R., Keith, T. Z. & Beretvas, S. N. Use of factor mixture modeling to capture Spearman's law of diminishing returns. *Intelligence* **38**, 231–241 (2010).
71. Hoffman, M., Steinley, D., Trull, T. J. & Sher, K. J. Estimating transdiagnostic symptom networks: the problem of 'skip outs' in diagnostic interviews. *Psychol. Assess.* **31**, 73–81 (2019).
72. Kazdin, A. E. & Kagan, J. Models of dysfunction in developmental psychopathology. *Clin. Psychol. Sci. Pract.* **1**, 35–52 (1994).
73. Muthén, B. O. Latent variable modeling in heterogeneous populations. *Psychometrika* **54**, 557–585 (1989).
74. Oquendo, M. A. et al. Initial strategy for the future of DSM. *Am. J. Psychiatry* **183**, 292–300 (2026).
75. Ruggero, C. J. et al. Integrating the Hierarchical Taxonomy of Psychopathology (HiTOP) into clinical practice. *J. Consult. Clin. Psychol.* **87**, 1069–1084 (2019).
76. de Ron, J., Fried, E. I. & Epskamp, S. Psychological networks in clinical populations: investigating the consequences of Berkson's bias. *Psychol. Med.* **51**, 168–176 https://doi.org/10.1017/S0033291719003209 (2021).
77. Hill, G. Neyman's bias re-visited. *J. Clin. Epidemiol.* **56**, 293–296 (2003).
78. Meredith, W. Notes on factorial invariance. *Psychometrika* **29**, 177–185 (1964).
79. McNeish, D. & Wolf, M. G. Thinking twice about sum scores. *Behav. Res. Methods* **52**, 2287–2305 (2020).
80. Ringwald, W. R., Forbes, M. K. & Wright, A. G. C. Meta-analytic tests of measurement invariance of internalizing and externalizing psychopathology across common methodological characteristics. *J. Psychopathol. Clin. Sci.* **131**, 847–856 (2022).
81. Conway, C. C. & Brown, T. A. Evaluating dimensional models of psychopathology in outpatients diagnosed with emotional disorders: a cautionary tale. *Depress. Anxiety* **35**, 898–902 (2018).
82. Kessler, R. C. Lifetime and 12-month prevalence of DSM-III-R psychiatric disorders in the United States: results from the National Comorbidity Survey. *Arch. Gen. Psychiatry* **51**, 8 (1994).
83. Kotov, R., Ruggero, C. J., Krueger, R. F., Watson, D. & Zimmerman, M. The perils of hierarchical exclusion rules: a further word of caution. *Depress. Anxiety* **35**, 903–904 (2018).
84. Mardia, K. V. Measures of multivariate skewness and kurtosis with applications. *Biometrika* **57**, 519–530 (1970).
85. Forbes, M. K. et al. Three recommendations based on a comparison of the reliability and validity of the predominant models used in research on the empirical structure of psychopathology. *J. Abnorm. Psychol.* **130**, 297–317 (2021).
86. Ringwald, W. R., Forbes, M. K. & Wright, A. G. C. Meta-analysis of structural evidence for the Hierarchical Taxonomy of Psychopathology (HiTOP) model. *Psychol. Med.* **14**, 1–14 (2021).
87. Littlefield, A. K., Lane, S. P., Gette, J. A., Watts, A. L. & Sher, K. J. The 'Big Everything': integrating and investigating dimensional models of psychopathology, personality, personality pathology, and cognitive functioning. *Personal. Disord.* **12**, 103–114 (2021).
88. Watts, A. L., Sher, K. J., Heath, A. C., Steinley, D. & Brusco, M. 'General addiction liability' revisited. *Clin. Psychol. Sci.* **13**, 242–260 (2025).
89. Garavan, H. et al. Recruiting the ABCD sample: design considerations and procedures. *Dev. Cogn. Neurosci.* **32**, 16–22 (2018).
90. Achenbach, T. M. & Rescorla, L. A. *Manual for the ASEBA School-Age Forms & Profiles* (Univ. of Vermont, Research Center for Children, Youth, & Families, 2001).
91. Achenbach, T. M. & Ruffle, T. M. The child behavior checklist and related forms for assessing behavioral/emotional problems and competencies. *Pediatr. Rev.* **21**, 265–271 (2000).
92. Kotov, R. et al. Validity and utility of Hierarchical Taxonomy of Psychopathology (HiTOP): I. Psychosis superspectrum. *World Psychiatry* **19**, 151–172 (2020).
93. Michelini, G. et al. Where do neurodevelopmental conditions fit in transdiagnostic psychiatric frameworks? Incorporating a new neurodevelopmental spectrum. *World Psychiatry* **23**, 333–357 (2024).
94. Van Buuren, S. Broken stick model for irregular longitudinal data. *J. Stat. Softw.* **106**, 1–51 https://doi.org/10.18637/jss.v106.i07 (2023).
95. McArdle, J. J. & Wang, L. Modeling age-based turning points in longitudinal life-span growth curves of cognition. In *Applied Data Analytic Techniques for Turning Points Research* (ed. P. Cohen) 105–127 (Taylor & Francis Group, 2008).
96. Muggeo, V. M. R. Modeling temperature effects on mortality: multiple segmented relationships with common break points. *Biostatistics* **9**, 613–620 (2008).
97. Gundermann, K. P. et al. Change-point models for identifying behavioral transitions in wild animals. *Mov. Ecol.* **11**, 65 (2023).
98. Qian, S. S., King, R. S. & Richardson, C. J. Two statistical methods for the detection of environmental thresholds. *Ecol. Model.* **166**, 87–97 (2003).
99. Aminikhanghahi, S. & Cook, D. J. A survey of methods for time series change point detection. *Knowl. Inf. Syst.* **51**, 339–367 (2017).
100. Imbens, G. W. & Lemieux, T. Regression discontinuity designs: a guide to practice. *J. Econom.* **142**, 615–635 (2008).
101. Hilbe, J. M. *Negative Binomial Regression* 2nd edn (Cambridge Univ. Press, 2011).
102. Gardner, W., Mulvey, E. P. & Shaw, E. C. Regression analyses of counts and rates: poisson, overdispersed poisson, and negative binomial models. *Psychol. Bull.* **118**, 392–404 (1995).
103. Stan Development Team. *The Stan Core Library* http://mc-stan.org/ (2023).
104. Hoffman, M. D. & Gelman, A. The no-U-turn sampler: adaptively setting path lengths in Hamiltonian Monte Carlo. *J. Mach. Learn. Res.* **15**, 1593–1623 (2014).
105. Gelman, A. & Rubin, D. B. Inference from iterative simulation using multiple sequences. *Stat. Sci.* **7**, 457–472 https://doi.org/10.1214/ss/1177011136 (1992).
106. Vehtari, A., Gelman, A. & Gabry, J. Practical Bayesian model evaluation using leave-one-out cross-validation and WAIC. *Stat. Comput.* **27**, 1413–1432 (2017).
107. Wood S. N. *Generalized Additive Models: An Introduction with R* 2nd edn https://doi.org/10.1201/9781315370279 (Chapman and Hall/CRC, 2017).
108. Hastie, T. & Tibshirani, R. *Generalized Additive Models* https://doi.org/10.1201/9780203753781 (Taylor & Francis, 2017).
109. Pines, A. et al. Psychiatric symptoms, cognition, and symptom severity in children. *JAMA Psychiatry* **81**, 1236 (2024).
110. Haist, F. & Jernigan, T. L. Adolescent Brain Cognitive Development Study (ABCD) - Annual Release 5.0. *NDA* https://doi.org/10.15154/8873-zj65 (2023).

Acknowledgements

This work was funded by the National Institutes of Health. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health.

Author contributions

We define author contribution using the Contributor Roles Taxonomy (CRediT). A.L.W. was responsible for conceptualization, data curation, formal analysis, investigation, methodology, project administration,

resources, software, supervision, validation, visualization and writing (original draft and editing). A.F.G.R. was responsible for formal analysis, methodology, visualization and writing (original draft and editing). A.L.G.K.M.K., F.A.C.M., T.J.T., K.J.S. and K.J.J. were responsible for writing (review and editing).

Funding

A.L.W. and A.F.G.R. are funded by the National Institute on Alcohol Abuse and Alcoholism (grant no. R00AA028306 to A.L.W.). A.L.G. was supported by the VISN 2 Mental Illness Research, Education, and Clinical Center and a Career Development Award–1 (award no. IK1RX004551-01A1) from the US Department of Veterans Affairs Rehabilitation Research and Development Service of the VA Office of Research and Development.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary information The online version contains supplementary material available at https://doi.org/10.1038/s44220-026-00727-0.

Correspondence and requests for materials should be addressed to Ashley L. Watts.

Peer review information *Nature Mental Health* thanks Thomas Olino and the other, anonymous, reviewer(s) for their contribution to the peer review of this work. Peer reviewer reports are available.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit http://creativecommons.org/licenses/by/4.0/.

© The Author(s) 2026

nature portfolio

Corresponding author(s): Ashley L. Watts

Last updated by author(s): Jul 7, 2026

Reporting Summary

Nature Portfolio wishes to improve the reproducibility of the work that we publish. This form provides structure for consistency and transparency in reporting. For further information on Nature Portfolio policies, see our Editorial Policies and the Editorial Policy Checklist.

Statistics

For all statistical analyses, confirm that the following items are present in the figure legend, table legend, main text, or Methods section.

n/a	Confirmed
☐	☒ The exact sample size (*n*) for each experimental group/condition, given as a discrete number and unit of measurement
☐	☒ A statement on whether measurements were taken from distinct samples or whether the same sample was measured repeatedly
☐	☒ The statistical test(s) used AND whether they are one- or two-sided *Only common tests should be described solely by name; describe more complex techniques in the Methods section.*
☐	☒ A description of all covariates tested
☐	☒ A description of any assumptions or corrections, such as tests of normality and adjustment for multiple comparisons
☐	☒ A full description of the statistical parameters including central tendency (e.g. means) or other basic estimates (e.g. regression coefficient) AND variation (e.g. standard deviation) or associated estimates of uncertainty (e.g. confidence intervals)
☐	☒ For null hypothesis testing, the test statistic (e.g. *F*, *t*, *r*) with confidence intervals, effect sizes, degrees of freedom and *P* value noted *Give P values as exact values whenever suitable.*
☐	☒ For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings
☐	☒ For hierarchical and complex designs, identification of the appropriate level for tests and full reporting of outcomes
☐	☒ Estimates of effect sizes (e.g. Cohen's *d*, Pearson's *r*), indicating how they were calculated

Our web collection on statistics for biologists contains articles on many of the points above.

Software and code

Policy information about availability of computer code

Data collection	*Provide a description of all commercial, open source and custom code used to collect the data in this study, specifying the version used OR state that no software was used.*
Data analysis	All of our analysis code is available on Github (https://github.com/Watts-MAP-Lab/psychopathology_nonlinearity). The Adolescent Brain and Cognitive Development Study data are not publicly available but are attainable upon receipt of a data use agreement. With access to the data, our results are fully reproducible.

For manuscripts utilizing custom algorithms or software that are central to the research but not yet described in published literature, software must be made available to editors and reviewers. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Portfolio guidelines for submitting code & software for further information.

Data

Policy information about availability of data

All manuscripts must include a data availability statement. This statement should provide the following information, where applicable:

- Accession codes, unique identifiers, or web links for publicly available datasets
- A description of any restrictions on data availability
- For clinical datasets or third party data, please ensure that the statement adheres to our policy

The Adolescent Brain and Cognitive Development Study data are not publicly available but are attainable upon receipt of a data use agreement. With access to the

data, our results are fully reproducible.

Data used in the preparation of this article were obtained from the Adolescent Brain Cognitive Study (ABCD; https://abcdstudy.org), held in the NIMH Data Archive (NDA). This is a multisite, longitudinal study designed to recruit more than 10,000 children aged 9-10 and follow them over 10 years into early adulthood. The ABCD Study® is supported by the National Institutes of Health and additional federal partners under award numbers U01DA041048, U01DA050989, U01DA051016, U01DA041022, U01DA051018, U01DA051037, U01DA050987, U01DA041174, U01DA041106, U01DA041117, U01DA041028, U01DA041134, U01DA050988, U01DA051039, U01DA041156, U01DA041025, U01DA041120, U01DA051038, U01DA041148, U01DA041093, U01DA041089, U24DA041123, U24DA041147. A full list of supporters is available at https://abcdstudy.org/federal-partners.html. A listing of participating sites and a complete listing of the study investigators can be found at https://abcdstudy.org/consortium_members/. ABCD consortium investigators designed and implemented the study and/or provided data but did not necessarily participate in the analysis or writing of this report. This manuscript reflects the views of the authors and may not reflect the opinions or views of the NIH or ABCD consortium investigators. The ABCD data repository grows and changes over time. The ABCD data used in this report came from DOI 10.15154/8873-zj65.

Research involving human participants, their data, or biological material

Policy information about studies with human participants or human data. See also policy information about sex, gender (identity/presentation), and sexual orientation and race, ethnicity and racism.

Reporting on sex and gender	Biological sex was used as a covariate.
Reporting on race, ethnicity, or other socially relevant groupings	We did not control for race, ethnicity, or other socially relevant groupings.
Population characteristics	*Describe the covariate-relevant population characteristics of the human research participants (e.g. age, genotypic information, past and current diagnosis and treatment categories). If you filled out the behavioural & social sciences study design questions and have nothing to add here, write "See above."*
Recruitment	Study participants were recruited across 21 study sites in the US for in-person study participation.
Ethics oversight	The IRB for the ABCD Study is managed by the University of California, San Diego.

Note that full information on the approval of the study protocol must also be provided in the manuscript.

Field-specific reporting

Please select the one below that is the best fit for your research. If you are not sure, read the appropriate sections before making your selection.

☐ Life sciences ☒ Behavioural & social sciences ☐ Ecological, evolutionary & environmental sciences

For a reference copy of the document with all sections, see nature.com/documents/nr-reporting-summary-flat.pdf

Behavioural & social sciences study design

All studies must disclose on these points even when the disclosure is negative.

Study description	The data are quantitative.
Research sample	Data were drawn from the first four waves of ABCD (Data Release 5.0), a collaboration between 21 sites across the US38,39 that tracks a sample of 11,868 9- and 10-year-olds from the US annually across development (wave 1: Mage = 9.94 years [SD = .62 years]; 48% female). Seventy-four percent of youth participants identified as White (14% Hispanic), 21% as Black (2% Hispanic), 7% as Asian, and 11% as Other (see Table B1 in Appendix B for demographics of all other waves). Twenty percent identified as Hispanic.
Sampling strategy	The sample size of ABCD was determined by the study investigators. We conducted secondary data analysis of the existing data.
Data collection	*Provide details about the data collection procedure, including the instruments or devices used to record the data (e.g. pen and paper, computer, eye tracker, video or audio equipment) whether anyone was present besides the participant(s) and the researcher, and whether the researcher was blind to experimental condition and/or the study hypothesis during data collection.*
Timing	*Indicate the start and stop dates of data collection. If there is a gap between collection periods, state the dates for each sample cohort.*
Data exclusions	*If no data were excluded from the analyses, state so OR if data were excluded, provide the exact number of exclusions and the rationale behind them, indicating whether exclusion criteria were pre-established.*
Non-participation	*State how many participants dropped out/declined participation and the reason(s) given OR provide response rate OR state that no participants dropped out/declined participation.*

Randomization

If participants were not allocated into experimental groups, state so OR describe how participants were allocated to groups, and if allocation was not random, describe how covariates were controlled.

Reporting for specific materials, systems and methods

We require information from authors about some types of materials, experimental systems and methods used in many studies. Here, indicate whether each material, system or method listed is relevant to your study. If you are not sure if a list item applies to your research, read the appropriate section before selecting a response.

Materials & experimental systems

n/a	Involved in the study
☐	☐ Antibodies
☐	☐ Eukaryotic cell lines
☐	☐ Palaeontology and archaeology
☐	☐ Animals and other organisms
☐	☐ Clinical data
☐	☐ Dual use research of concern
☐	☐ Plants

Methods

n/a	Involved in the study
☐	☐ ChIP-seq
☐	☐ Flow cytometry
☐	☐ MRI-based neuroimaging

Antibodies

Antibodies used

Describe all antibodies used in the study; as applicable, provide supplier name, catalog number, clone name, and lot number.

Validation

Describe the validation of each primary antibody for the species and application, noting any validation statements on the manufacturer's website, relevant citations, antibody profiles in online databases, or data provided in the manuscript.

Eukaryotic cell lines

Policy information about cell lines and Sex and Gender in Research

Cell line source(s)

State the source of each cell line used and the sex of all primary cell lines and cells derived from human participants or vertebrate models.

Authentication

Describe the authentication procedures for each cell line used OR declare that none of the cell lines used were authenticated.

Mycoplasma contamination

Confirm that all cell lines tested negative for mycoplasma contamination OR describe the results of the testing for mycoplasma contamination OR declare that the cell lines were not tested for mycoplasma contamination.

Commonly misidentified lines
(See ICLAC register)

Name any commonly misidentified cell lines used in the study and provide a rationale for their use.

Palaeontology and Archaeology

Specimen provenance

Provide provenance information for specimens and describe permits that were obtained for the work (including the name of the issuing authority, the date of issue, and any identifying information). Permits should encompass collection and, where applicable, export.

Specimen deposition

Indicate where the specimens have been deposited to permit free access by other researchers.

Dating methods

If new dates are provided, describe how they were obtained (e.g. collection, storage, sample pretreatment and measurement), where they were obtained (i.e. lab name), the calibration program and the protocol for quality assurance OR state that no new dates are provided.

☐ Tick this box to confirm that the raw and calibrated dates are available in the paper or in Supplementary Information.

Ethics oversight

Identify the organization(s) that approved or provided guidance on the study protocol, OR state that no ethical approval or guidance was required and explain why not.

Note that full information on the approval of the study protocol must also be provided in the manuscript.

Animals and other research organisms

Policy information about studies involving animals; ARRIVE guidelines recommended for reporting animal research, and Sex and Gender in Research

Laboratory animals	*For laboratory animals, report species, strain and age OR state that the study did not involve laboratory animals.*
Wild animals	*Provide details on animals observed in or captured in the field; report species and age where possible. Describe how animals were caught and transported and what happened to captive animals after the study (if killed, explain why and describe method; if released, say where and when) OR state that the study did not involve wild animals.*
Reporting on sex	*Indicate if findings apply to only one sex; describe whether sex was considered in study design, methods used for assigning sex. Provide data disaggregated for sex where this information has been collected in the source data as appropriate; provide overall numbers in this Reporting Summary. Please state if this information has not been collected. Report sex-based analyses where performed, justify reasons for lack of sex-based analysis.*
Field-collected samples	*For laboratory work with field-collected samples, describe all relevant parameters such as housing, maintenance, temperature, photoperiod and end-of-experiment protocol OR state that the study did not involve samples collected from the field.*
Ethics oversight	*Identify the organization(s) that approved or provided guidance on the study protocol, OR state that no ethical approval or guidance was required and explain why not.*

Note that full information on the approval of the study protocol must also be provided in the manuscript.

Clinical data

Policy information about clinical studies

All manuscripts should comply with the ICMJE guidelines for publication of clinical research and a completed CONSORT checklist must be included with all submissions.

Clinical trial registration	*Provide the trial registration number from ClinicalTrials.gov or an equivalent agency.*
Study protocol	*Note where the full trial protocol can be accessed OR if not available, explain why.*
Data collection	*Describe the settings and locales of data collection, noting the time periods of recruitment and data collection.*
Outcomes	*Describe how you pre-defined primary and secondary outcome measures and how you assessed these measures.*

Dual use research of concern

Policy information about dual use research of concern

Hazards

Could the accidental, deliberate or reckless misuse of agents or technologies generated in the work, or the application of information presented in the manuscript, pose a threat to:

No	Yes	
☒	☐	Public health
☒	☐	National security
☒	☐	Crops and/or livestock
☒	☐	Ecosystems
☒	☐	Any other significant area

Experiments of concern

Does the work involve any of these experiments of concern:

No	Yes	
☒	☐	Demonstrate how to render a vaccine ineffective
☒	☐	Confer resistance to therapeutically useful antibiotics or antiviral agents
☒	☐	Enhance the virulence of a pathogen or render a nonpathogen virulent
☒	☐	Increase transmissibility of a pathogen
☒	☐	Alter the host range of a pathogen
☒	☐	Enable evasion of diagnostic/detection modalities
☒	☐	Enable the weaponization of a biological agent or toxin
☒	☐	Any other potentially harmful combination of experiments and agents

Plants

Seed stocks

Report on the source of all seed stocks or other plant material used. If applicable, state the seed stock centre and catalogue number. If plant specimens were collected from the field, describe the collection location, date and sampling procedures.

Novel plant genotypes

Describe the methods by which all novel plant genotypes were produced. This includes those generated by transgenic approaches, gene editing, chemical/radiation-based mutagenesis and hybridization. For transgenic lines, describe the transformation method, the number of independent lines analyzed and the generation upon which experiments were performed. For gene-edited lines, describe the editor used, the endogenous sequence targeted for editing, the targeting guide RNA sequence (if applicable) and how the editor was applied.

Authentication

Describe any authentication procedures for each seed stock used or novel genotype generated. Describe any experiments used to assess the effect of a mutation and, where applicable, how potential secondary effects (e.g. second site T-DNA insertions, mosiacism, off-target gene editing) were examined.

ChIP-seq

Data deposition

☐ Confirm that both raw and final processed data have been deposited in a public database such as GEO.

☐ Confirm that you have deposited or provided access to graph files (e.g. BED files) for the called peaks.

Data access links
May remain private before publication.

For "Initial submission" or "Revised version" documents, provide reviewer access links. For your "Final submission" document, provide a link to the deposited data.

Files in database submission

Provide a list of all files available in the database submission.

Genome browser session
(e.g. UCSC)

Provide a link to an anonymized genome browser session for "Initial submission" and "Revised version" documents only, to enable peer review. Write "no longer applicable" for "Final submission" documents.

Methodology

Replicates

Describe the experimental replicates, specifying number, type and replicate agreement.

Sequencing depth

Describe the sequencing depth for each experiment, providing the total number of reads, uniquely mapped reads, length of reads and whether they were paired- or single-end.

Antibodies

Describe the antibodies used for the ChIP-seq experiments; as applicable, provide supplier name, catalog number, clone name, and lot number.

Peak calling parameters

Specify the command line program and parameters used for read mapping and peak calling, including the ChIP, control and index files used.

Data quality

Describe the methods used to ensure data quality in full detail, including how many peaks are at FDR 5% and above 5-fold enrichment.

Software

Describe the software used to collect and analyze the ChIP-seq data. For custom code that has been deposited into a community repository, provide accession details.

Flow Cytometry

Plots

Confirm that:

☐ The axis labels state the marker and fluorochrome used (e.g. CD4-FITC).

☐ The axis scales are clearly visible. Include numbers along axes only for bottom left plot of group (a 'group' is an analysis of identical markers).

☐ All plots are contour plots with outliers or pseudocolor plots.

☐ A numerical value for number of cells or percentage (with statistics) is provided.

Methodology

Sample preparation	*Describe the sample preparation, detailing the biological source of the cells and any tissue processing steps used.*
Instrument	*Identify the instrument used for data collection, specifying make and model number.*
Software	*Describe the software used to collect and analyze the flow cytometry data. For custom code that has been deposited into a community repository, provide accession details.*
Cell population abundance	*Describe the abundance of the relevant cell populations within post-sort fractions, providing details on the purity of the samples and how it was determined.*
Gating strategy	*Describe the gating strategy used for all relevant experiments, specifying the preliminary FSC/SSC gates of the starting cell population, indicating where boundaries between "positive" and "negative" staining cell populations are defined.*

☐ Tick this box to confirm that a figure exemplifying the gating strategy is provided in the Supplementary Information.

Magnetic resonance imaging

Experimental design

Design type	*Indicate task or resting state; event-related or block design.*
Design specifications	*Specify the number of blocks, trials or experimental units per session and/or subject, and specify the length of each trial or block (if trials are blocked) and interval between trials.*
Behavioral performance measures	*State number and/or type of variables recorded (e.g. correct button press, response time) and what statistics were used to establish that the subjects were performing the task as expected (e.g. mean, range, and/or standard deviation across subjects).*

Acquisition

Imaging type(s)	*Specify: functional, structural, diffusion, perfusion.*
Field strength	*Specify in Tesla*
Sequence & imaging parameters	*Specify the pulse sequence type (gradient echo, spin echo, etc.), imaging type (EPI, spiral, etc.), field of view, matrix size, slice thickness, orientation and TE/TR/flip angle.*
Area of acquisition	*State whether a whole brain scan was used OR define the area of acquisition, describing how the region was determined.*

Diffusion MRI ☐ Used ☐ Not used

Preprocessing

Preprocessing software	*Provide detail on software version and revision number and on specific parameters (model/functions, brain extraction, segmentation, smoothing kernel size, etc.).*
Normalization	*If data were normalized/standardized, describe the approach(es): specify linear or non-linear and define image types used for transformation OR indicate that data were not normalized and explain rationale for lack of normalization.*
Normalization template	*Describe the template used for normalization/transformation, specifying subject space or group standardized space (e.g. original Talairach, MNI305, ICBM152) OR indicate that the data were not normalized.*
Noise and artifact removal	*Describe your procedure(s) for artifact and structured noise removal, specifying motion parameters, tissue signals and physiological signals (heart rate, respiration).*

Volume censoring

Define your software and/or method and criteria for volume censoring, and state the extent of such censoring.

Statistical modeling & inference

Model type and settings

Specify type (mass univariate, multivariate, RSA, predictive, etc.) and describe essential details of the model at the first and second levels (e.g. fixed, random or mixed effects; drift or auto-correlation).

Effect(s) tested

Define precise effect in terms of the task or stimulus conditions instead of psychological concepts and indicate whether ANOVA or factorial designs were used.

Specify type of analysis: ☐ Whole brain ☐ ROI-based ☐ Both

Statistic type for inference

Specify voxel-wise or cluster-wise and report all relevant parameters for cluster-wise methods.

(See Eklund et al. 2016)

Correction

Describe the type of correction and how it is obtained for multiple comparisons (e.g. FWE, FDR, permutation or Monte Carlo).

Models & analysis

n/a	Involved in the study
☐	☐ Functional and/or effective connectivity
☐	☐ Graph analysis
☐	☐ Multivariate modeling or predictive analysis

Functional and/or effective connectivity

Report the measures of dependence used and the model details (e.g. Pearson correlation, partial correlation, mutual information).

Graph analysis

Report the dependent variable and connectivity measure, specifying weighted graph or binarized graph, subject- or group-level, and the global and/or node summaries used (e.g. clustering coefficient, efficiency, etc.).

Multivariate modeling and predictive analysis

Specify independent variables, features extraction and dimension reduction, model, training and evaluation metrics.