

Network temperature as a metric of stability in depression symptoms across adolescence

Received: 9 August 2024

Accepted: 19 March 2025

Published online: 29 April 2025



Poppy Z. Grimes¹✉, Aja L. Murray², Keith Smith³, Andrea G. Allegrini^{4,5}, Giulia G. Piazza⁴, Henrik Larsson^{6,7}, Sacha Epskamp^{8,11}, Heather C. Whalley^{1,9,11} & Alex S. F. Kwong^{1,10,11}

Depression is characterized by diverse symptom combinations that can be represented as dynamic networks. While previous research has focused on central symptoms for targeted interventions, less attention has been given to whole-network properties. Here we show that ‘network temperature’, a novel measure of psychological network stability, captures symptom alignment across adolescence—a critical period for depression onset. Network temperature reflects system stability, with higher values indicating less symptom alignment and greater variability. In three large longitudinal adolescent cohorts (total $N = 35,901$), we found that network temperature decreases across adolescence, with the steepest decline during early adolescence, particularly in males. This suggests that depression symptom networks stabilize throughout development via increased symptom alignment, potentially explaining why adolescence is a crucial period for depression onset. These findings highlight early adolescence as a key intervention window and underscore the importance of sex-specific and personalized interventions.

The network theory of psychopathology conceptualizes depression as emerging from the interactions among its composing symptoms¹. Network science has provided an alternative approach to psychiatric conditions that allows us to tease apart symptom-level heterogeneity for more precise etiological understanding and intervention insights². Though most studies have examined cross-sectional adult symptom networks, interpretation of these is limited, as depression symptoms evolve and fluctuate in severity and interconnectedness over time, forming a dynamic temporal network^{1,3}. Intraindividual networks are particularly important to examine across adolescence, where depression onset is most severe and symptoms follow diverse trajectories, but intervention may be most fruitful^{4–7}.

In both cross-sectional and longitudinal networks, network evaluation is focused on centrality estimates—measures of node importance in the network⁸. Commonly reported estimates are node strength, closeness and betweenness. However, centrality estimates were originally developed for social network analysis and have more limited interpretability within psychological networks due to issues such as node indistinctiveness (symptom overlap) and the lack of consideration for symptom severity⁹.

Network studies also allude to individual symptoms as treatment targets^{5,10–13}. However, targeting a specific node (symptom) may not have the desired effect, potentially failing to address system complexity. For example, a symptom might appear central and be targeted for

¹Division of Psychiatry, Centre for Clinical Brain Sciences, University of Edinburgh, Edinburgh, UK. ²Department of Psychology, University of Edinburgh, Edinburgh, UK. ³Computer and Information Sciences, University of Strathclyde, Glasgow, UK. ⁴Department of Clinical, Educational and Health Psychology, Division of Psychology and Language Sciences, University College London, London, UK. ⁵Social, Genetic and Developmental Psychiatry Centre, Institute of Psychiatry, Psychology and Neuroscience, King's College London, London, UK. ⁶Department of Medical Epidemiology and Biostatistics, Karolinska Institute, Stockholm, Sweden. ⁷School of Medical Sciences, Örebro University, Örebro, Sweden. ⁸Department of Psychology, National University of Singapore, Singapore, Singapore. ⁹Generation Scotland, Centre for Genomic and Experimental Medicine, Institute of Genetics and Molecular Medicine, University of Edinburgh, Edinburgh, UK. ¹⁰MRC Integrative Epidemiology Unit, University of Bristol, Bristol, UK. ¹¹These authors jointly supervised this work: Sacha Epskamp, Heather C. Whalley, Alex S. F. Kwong. ✉e-mail: p.grimes@ed.ac.uk

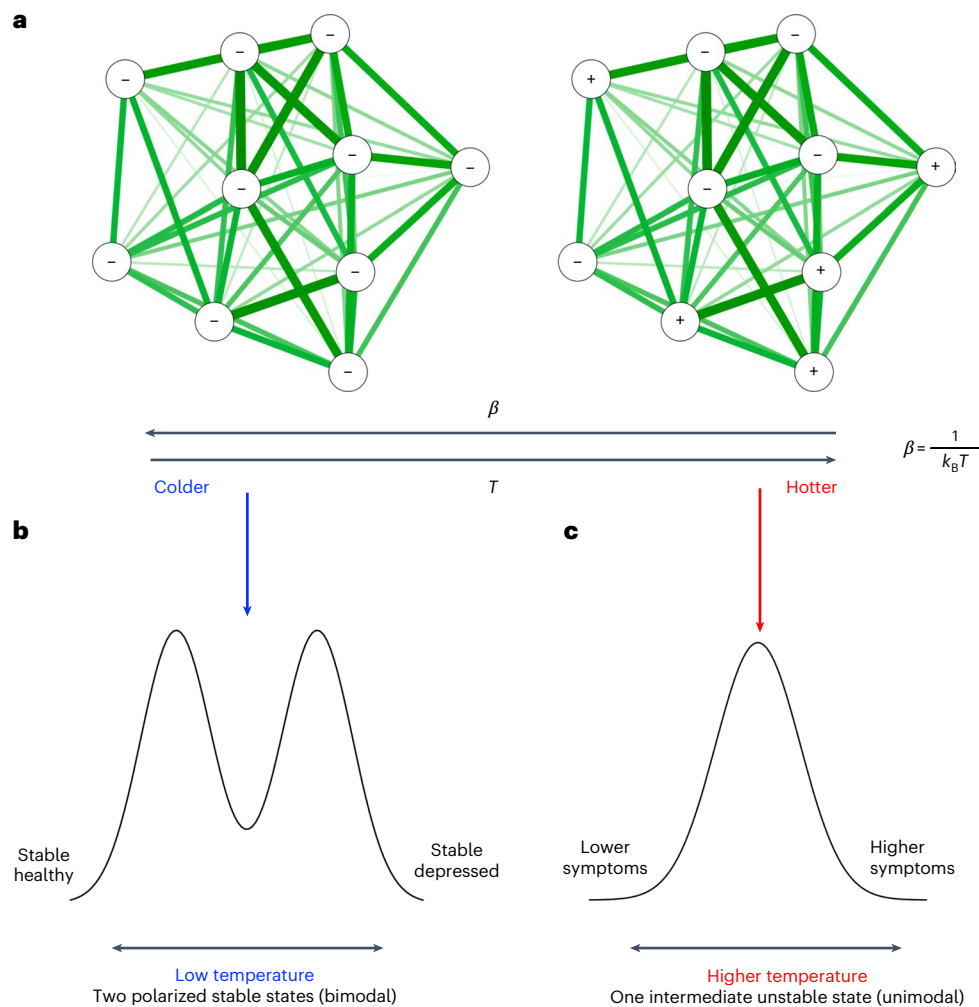


Fig. 1 | The temperature parameter controlling network behavior in depression. **a**, A lower temperature (higher β) characterizes more stable networks with a greater alignment of node states, leading to less variability in symptom configurations. **b**, Low-temperature networks occupy two stable

states: one corresponding to a healthy (nondepressed) state and the other to a high-depression state, resulting in a bimodal distribution. **c**, High-temperature networks exhibit greater within-person variability, existing in a single, intermediate unstable state with more fluctuation in symptom configurations.

intervention, but if it is actually a common effect of other symptoms, intervening on it may not disrupt the overall network, as it does not influence other nodes¹⁴. Effective intervention must also consider timing, hence the importance of longitudinal investigation. Consequently, estimating node centrality becomes less relevant longitudinally as symptoms fluctuate. Instead, it is suggested that we disregard centrality estimates and refocus on global measures that observe the network as a complex system, such as network structure comparison and global connectivity^{2,9,15,16}. As such, while centrality measures may be useful for questions based on symptom-specific insights, such as identifying the most strongly connected symptoms, further examination of network-wide characteristics to better understand depression as a system is needed.

According to symptom network theory, more densely connected networks tend to be associated with poorer outcomes^{1,17}. Persistent depression is thought to arise from a feedback loop of reinforcing connected symptoms. Supporting this, more densely connected networks were observed in individuals with persistent major depression versus those who remitted¹². Contrastingly, however, major depression treatment responders showed a larger increase in connectivity after treatment than nonresponders for both pharmacological¹⁸ and therapeutic interventions¹⁹. Increased network connectivity may therefore not necessarily be 'bad' or pathological but tell us about overall

system dynamics such as stability and variability²⁰. Just as adverse symptoms can propagate and negatively reinforce a tightly connected network, 'good' resilience could also propagate in a positively reinforcing network^{19,21}.

A potentially more valuable metric is network temperature²². Network temperature is analogous to physical temperature in the solid \rightarrow liquid \rightarrow gas energetic state transition described by statistical mechanics, which can be modeled using the Ising model^{23,24}. In the Ising model, temperature refers to the level of randomness in the possible states (for example, on or off, 1 or -1 and present or absent) that nodes can take²⁵. In depression networks, this randomness can be interpreted as symptom fluctuation. Entropy measures the disorder or uncertainty in the system: high entropy in depression networks indicates a wider and more unpredictable variation in symptoms, whereas low entropy suggests more consistent and predictable symptoms. Therefore, temperature influences global entropy and can provide insights into network or depression stability²². When considering a network as an interdependent energetic system, temperature tells us about the energetic organization.

At low network temperatures, the node states are aligned with a low number of possible configurations (within-person variability), resulting in low entropy and stable, predictable symptom patterns in depression networks. Conversely, high network temperatures lead

Table 1 | Preimputed and complete case sample characteristics

		Wave							
Demographics		1	2	3	4	5	6	7	8
ABCD	N	11,232	10,568	10,977	10,510	9,952	9,985	8,218	4,648
	Mean age (years)	10.4	10.9	11.4	12	12.4	12.9	13.4	14.1
	Female (%)	47.68	47.75	47.66	47.63	47.84	47.37	47.59	47.56
	BPM, mean	1.82	1.74	1.54	1.81	1.83	2.02	2.2	2.27
	BPM, s.d.	2.09	2.11	1.94	2.23	2.29	2.38	2.6	2.56
	Complete cases, mean	1.58	1.56	1.32	1.64	1.69	1.9	2.01	2.2
	Complete cases, s.d.	1.89	1.97	1.77	2.08	2.19	2.27	2.47	2.54
ALSPAC	N	7,364	6,716	6,019	4,997	4,497	3,335		
	Mean age (years)	10.7	12.8	13.8	16.7	17.8	18.7		
	Female (%)	50.56	51.2	51.52	58.82	56.94	63.83		
	SMFQ, mean	4.04	3.97	4.92	5.91	6.59	6.83		
	SMFQ, s.d.	3.51	3.86	4.49	5.64	5.25	5.93		
	Complete cases, mean	3.77	3.96	4.85	5.53	6.11	6.33		
	Complete cases, s.d.	3.24	3.74	4.4	5.16	4.9	5.58		
MCS	N (complete)	10,680	9,378	7,139					
	Mean age (years)	11.2	13.8	16.7					
	Female (%)	51.39	50.85	51.97					
	SDQ Emotional, mean	1.88	2.04	2.03					
	SDQ, s.d.	2	2.13	2.23					
	Complete cases, mean	1.82	1.99	2.03					
	Complete cases, s.d.	1.98	2.11	2.24					

N, number of individuals at each wave; s.d., standard deviation.

to more random activation patterns, higher entropy and less stable symptom profiles. In this case, symptoms might vary unpredictably over time, reflecting a system susceptible to change (Fig. 1a).

Network temperature has emerged in causal attitude networks (for example, political attitudes where nodes represent thoughts and feelings)^{26–28} and cognitive belief networks²⁹ that apply the Ising model representation of nodes in binary states²⁴. These systems tend toward an overall low-energy state of node alignment, where the probability of the overall network state depends on the amount of energy within the network. In attitude networks, directed attention (lower temperature) results in lower attitudinal entropy and increased stability²⁷. For example, when a political voter directs their cognitive attention on a specific issue, such as healthcare reform, the variability (entropy) of their attitude toward this issue decreases, leading to a more stable and consistent stance. In belief networks, individuals with high dissonance (belief inconsistency) are associated with low-temperature networks, where tightly connected beliefs amplify the drive to reduce dissonance by changing specific beliefs. Conversely, high temperature networks are characterized by a more malleable belief system but less pressure to resolve inconsistencies²⁹. Consequently, temperature represents a global parameter of the network state, scaling both the connectedness of the network, which indicates a general predisposition to an attitude or belief and the thresholds of the network, reflecting external influences on thoughts.

We can bring this idea to psychological symptom networks. The concepts of network temperature and entropy may be valuable in better describing stability in depression and enhancing understanding of risk and resilience. The model of psychiatric disorders as dynamical systems parallels this theory by describing two alternative ‘basins of

attraction’ characterized by a resilient healthy and resilient disorder state in early and recent work^{15,21,30}. The theory would suggest that both someone who is persistently depressed and someone who never experiences depressive symptoms have low temperature, stable networks in more fixed states (Fig. 1b). Conversely, someone with consistently fluctuating depression state or symptoms would have a high temperature network which exhibits increased variability (Fig. 1c).

We explore the hypothesis that symptom networks stabilize across adolescence. Research shows that emotional regulation improves with age³¹, while emotional fluctuations are pronounced during puberty and adolescence³². Depression trajectories also reveal increased symptom heterogeneity during adolescence, but as individuals approach adulthood, they tend to diverge toward two groups: nondepressed and persistently depressed^{4,7,33}, suggesting that individual symptoms stabilize over time.

The purpose of this work is to gain insight into the underlying dynamics of adolescent depression by integrating network analysis with longitudinal data. For the first time, we formalize the concept of network temperature in psychological symptom networks and assess its utility for network comparison using a novel approach of multigroup Ising models. We examine how network temperature changes across the development of adolescent depression in three large, independent cohorts, providing a robust understanding of temporal changes. This approach may identify critical developmental periods for intervention. In addition, as a further illustration of this method, we investigate how temperature changes under sex stratification. Sex is a key predictor of depression, with females suffering higher rates and severity across development³⁴. This demonstrates the potential application of network temperature to different subgroups or risk factors.

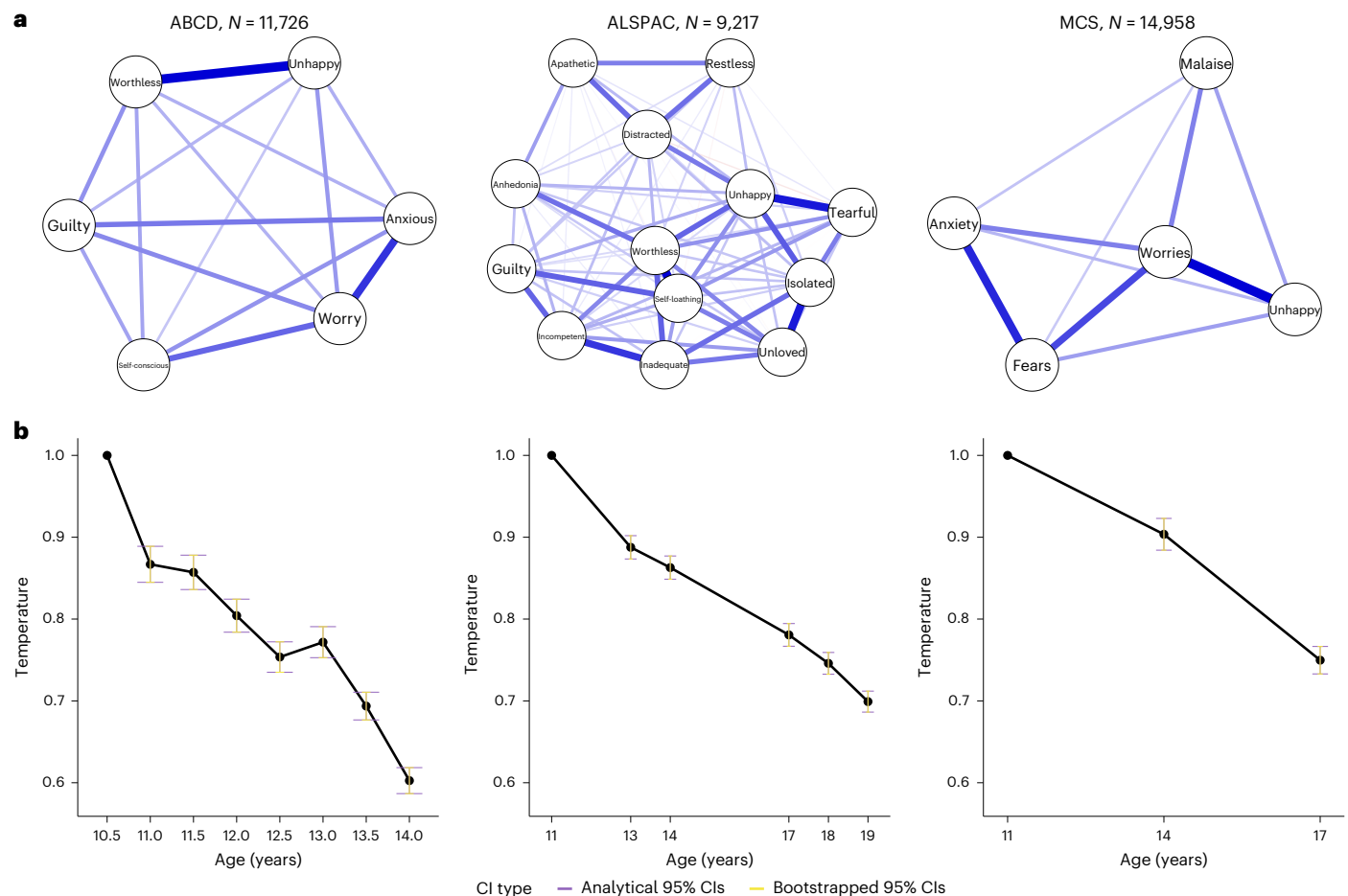


Fig. 2 | Depression symptom network temperature change across adolescence in three cohorts. **a**, The estimated Ising model networks with equal edges constrained over time. ABCD: $N = 11,726$; ALSPAC: $N = 9,217$; MCS: $N = 14,958$. **b**, The network temperature values were derived from the Ising model at each age,

with the first occasion in each cohort fixed to 1 for identification. The error bars represent the 95% CIs around the network temperature parameter estimate, with purple indicating the analytical CIs and yellow indicating the bootstrapped CIs.

Results

Sample characteristics

The imputed sample sizes for each cohort were $N = 11,726$ in the Adolescent Brain and Cognitive Development Study (ABCD), $N = 9,217$ in the Avon Longitudinal Study of Parents and Children Study (ALSPAC) and $N = 14,958$ in the Millennium Cohort Study (MCS) for individuals with at least one depression measure. ABCD is an ethnically diverse cohort and observed self-reported ethnicity of 250 Asian, 1,743 Black, 2,364 Hispanic, 6,134 White and 1,234 other individuals. ALSPAC and MCS are White European cohorts.

Preimputation sample demographics, attrition information for each cohort at each wave and complete case statistics are reported in Table 1. Both the depression summary score means and standard deviations increase with age in the original missing data and complete case samples, indicating an increase in between-person depression score variability. Preimputation symptom frequencies are provided in Supplementary Tables 8–10. Missing symptom data rates were 28.8% in ABCD, 40.1% in ALSPAC and 32.6% in MCS. For ABCD and ALSPAC, the missingness patterns indicated unit nonresponse, whereas in MCS, there was a general pattern of missingness. Full information on missing data patterns is shown in Supplementary Figs. 1–3.

Network structure

Networks were estimated at multiple waves across adolescent development and constructed from Ising models of binary symptoms. For all cohorts, a dense network structure with fixed edge weights that

allowed external fields and temperature to vary over time provided the best model fit. The network structures are depicted in Fig. 2a. The full model fit results for models increasing in constraint are reported in Supplementary Tables 11–13. The network model covariance matrices are provided in Supplementary Tables 14–16.

Network temperature

The results are reported as T for temperature with the 95% confidence intervals (CI). Across all three cohorts, the network temperature decreased across adolescence, albeit at different rates (Fig. 2b). Temperature decrease indicates an increase in node state alignment in the network as the system becomes more ordered (less random), and the symptoms tend to ‘all on’ or ‘all off’. In ABCD, the sharpest drop in temperature was between 10.5 ($T = 1$) and 11 years ($T = 0.87$ (0.85–0.88)), with continued cooling to age 14 years ($T = 0.60$ (0.59–0.61)). In ALSPAC, we observed the sharpest drop in temperature between 11 ($T = 1$) and 13 years ($T = 0.89$ (0.88–0.90)) and a consistent cooling between 13 and 19 years ($T = 0.70$ (0.69–0.71)). In MCS, we saw a consistent cooling rate between age 11 ($T = 1$) and 17 years ($T = 0.75$ (0.74–0.76)).

Network entropy decreased consistently with temperature cooling across cohorts, reflecting reduced instability through increased node state alignment (Supplementary Fig. 4). The mean symptom scores and symptom variance showed distinct patterns (increasing in ALSPAC but decreasing in ABCD and MCS), indicating that network-level changes (temperature and entropy) provide insights beyond static symptom measures (Supplementary Tables 17–19).

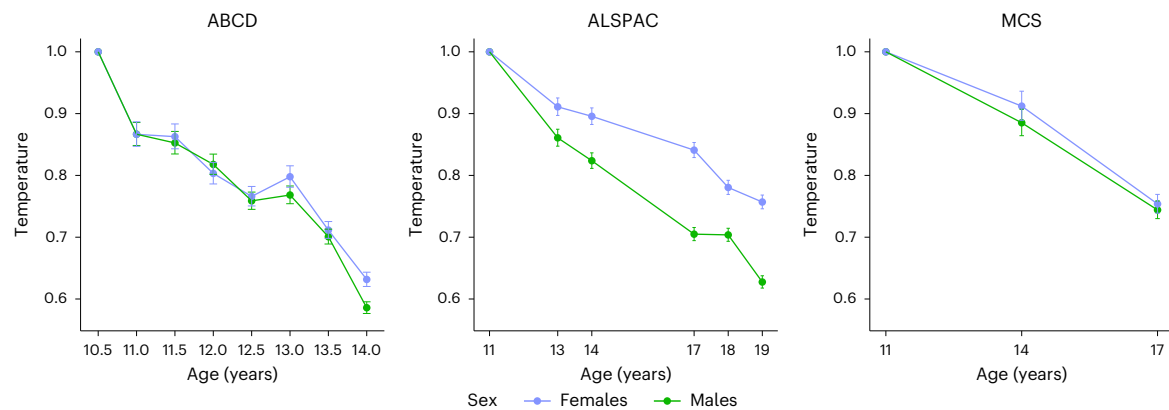


Fig. 3 | Sex-stratified depression network temperature change across three cohorts. The error bar plots show the relative rate of temperature decrease with age, with the first occasion in each cohort fixed to 1 for identification. ABCD: $N = 5,603$ (females), $N = 6,120$ (males); ALSPAC: $N = 4,784$ (females), $N = 4,417$

(males); MCS: $N = 6,382$ (females), $N = 6,442$ (males). The error bars represent the analytical 95% CIs around the network temperature parameter estimate. The lines illustrate relative rates of cooling across age, not absolute temperature values.

Sex differences in depression temperature across adolescence

The sex stratification of network temperature revealed an overall trend of faster cooling and lower temperature in males than females across adolescence in all three cohorts (Fig. 3). Female temperature was higher across adolescence, indicating increased symptom variability and, therefore, reduced network stability. This means females observed a longer period of high temperature where the network is more susceptible to change, for example, potential impact by environmental stressors acting as the thresholds.

In the ABCD cohort, temperature decreased at a similar rate for both sexes until mid-adolescence. By age 12.5 years, males exhibited a slightly faster cooling trend compared with females. At this age, the temperatures showed a near identical decrease (males, $T = 0.76$ (0.75–0.77); females, $T = 0.77$ (0.75–0.78)). By age 13 years, the females showed a minor increase in temperature to $T = 0.80$ (0.78–0.82), while the temperature of the males remained stable. By age 14 years, both sexes demonstrated rapid cooling. The males exhibited the greatest cooling rate, with a decrease of $T = 0.59$ (0.58–0.60) compared with the females ($T = 0.63$ (0.62–0.64)). The effect size of the interaction between age and sex was -0.0108 , indicating a trend toward faster cooling in males. The evidence for this effect was modest ($P = 0.050$).

In the ALSPAC cohort, the males displayed a rapid cooling rate from age 11 ($T = 1$) to age 14 years ($T = 0.82$ (0.81–0.84)), continuing into late adolescence ($T = 0.63$ (0.62–0.64) at age 19 years). In contrast, the females showed a slower rate of cooling, with temperatures decreasing to $T = 0.90$ (0.88–0.91) by age 14 years. By age 19 years, the relative temperature for the females remained higher compared with the males, at $T = 0.79$ (0.75–0.77), exhibiting less stability in their networks. The effect size of the interaction between age and sex was -0.0140 , indicating a faster cooling trend in males, with the evidence supporting this observation being robust ($P = 0.025$).

In the MCS cohort, both the male and female networks exhibited cooling, though males at a slightly faster rate. The difference was most apparent between ages 11 and 14 years, with males cooling to $T = 0.88$ (0.86–0.91) and females to $T = 0.92$ (0.89–0.94). By age 17 years, both sexes reached similar temperatures (males, $T = 0.75$ (0.83–0.76); females, $T = 0.75$ (0.74–0.77)). The effect size for the interaction between age and sex was -0.00162 , suggesting a trend toward faster cooling in males. The evidence for this effect was weak ($P = 0.769$).

Sensitivity analyses

Estimated network structures remained consistent when analyzing nonimputed data (complete cases) only, with the maximum-likelihood estimator (Supplementary Tables 20–22 and Supplementary Figs. 5–7). The network temperature also decreased across all cohorts, albeit to a

lesser extent than with the imputed samples. As expected, the network structure remained equivalent with [0,1] encoding. Surprisingly, we also saw consistent results for the temperature parameter suggesting either encoding approach is suitable for analysis of network temperature (Supplementary Tables 23–25 and Supplementary Figs. 8–10). The global network connectivity was approximately proportional to the product of connectivity at the first time point and beta (Supplementary Tables 26–28). The minor discrepancies of calculated values probably reflect numerical variations arising from parameter estimation, aggregation or rounding effects. The overall patterns demonstrate that connectivity is scaled by beta, reinforcing that temperature captures distinct information on system dynamics beyond what connectivity alone can reveal.

Discussion

For the first time, we apply the concept of network temperature to psychological symptom networks. Specifically, we observed a decrease in the temperature of depressive symptom networks across adolescent development, indicating increased stability and decreased variability of symptoms. While earlier studies have begun to interpret mental health networks as complex systems, moving away from node-focused evaluations, these studies typically identify cross-sectional differences such as network structure or density. We demonstrate network temperature as a novel metric that captures network dynamics over time, a measure currently missing from symptom network studies and which can be applied to group comparisons of network stability.

In three adolescent cohorts with three different depression scales, we show a consistent pattern of network temperature decrease across adolescence. This ‘cooling’ indicates increased within-person depression stability as symptoms become more aligned and less variable with age. This finding is supported by previous work that found global connectivity of depression–anxiety networks increased from childhood through adolescence³⁵.

Our results confirm that network temperature and entropy decrease together, supporting the hypothesis of reduced instability as networks settle into fewer configurations. Importantly, changes in symptom variance were cohort specific, highlighting that network temperature provides unique insights into system dynamics not captured by symptom variability alone.

To illustrate these cohort-specific patterns, we observed that symptom variation decreased in ABCD and MCS, suggesting increasing synchronization of symptoms across individuals and more homogeneous samples. In contrast, the increase in symptom variance in ALSPAC indicates greater heterogeneity in symptom dynamics, potentially reflecting a polarization into two subgroups: individuals

with persistently high symptoms and those stabilizing at low levels. This aligns with the theory of low-temperature network states, where symptoms converge into ‘all-on’ or ‘all-off’ configurations, consistent with prior trajectory findings in this cohort^{7,33,36}.

Temperature may explain why we see strongly connected networks in both depression patients and healthy individuals—offering a dynamical explanation for the conflicting evidence as to whether connectivity is ‘good’ or ‘bad’. Rather than being pathological in symptom networks, as previously thought¹⁵, increased connectivity may instead represent lower temperature and increased homogeneity of symptom states within a group.

In the study suggesting connectivity as pathological, the individuals with persistent depression had a higher network connectivity compared with those whose depression had remitted¹². This difference may be due to persisters having stable low-temperature networks, while remitters exhibit higher temperatures and more symptom heterogeneity.

In the study that saw increased connectivity as a marker of improvement, major depression treatment responders had a weaker baseline network connectivity than nonresponders but then saw the greatest increase in connectivity with treatment¹⁹. This could be understood as treatment responders initially having less stable, ‘hotter’ networks, which were more susceptible to intervention. This increased system randomness may have enabled treatment to tip them into a positive stable state, heightening stability or connectivity, whereas the nonresponders began with lower temperature networks and remained ‘stuck’ in these negative low-temperature states.

Connectivity and temperature are conceptually distinct measures in network dynamics. Connectivity quantifies the aggregate strength of pairwise interactions in a static network, while temperature governs the system’s stability and susceptibility to fluctuations, reflecting its dynamism and likelihood of transitioning between configurations. By influencing edge weights and thresholds (external fields), temperature shapes the probability distribution of node states, making it especially relevant in psychological networks where symptom instability is critical.

While connectivity captures the overall strength of relationships, temperature highlights fluctuations and transitions in symptom patterns, offering a nuanced perspective on system dynamics. This distinction is particularly valuable when sudden changes or instability are key phenomena. By identifying temperature and enabling group comparisons, our study complements traditional connectivity measures and provides new insights into psychological network behavior.

Our findings are also consistent with research on longitudinal depression trajectories. Across adolescence, studies observe two stable groups: one displaying persistent symptoms (stable high) and a second with low or no symptoms (stable low)^{7,33,37}. Individuals in both of these stable groups would be characterized by low-temperature networks. The individuals belonging to more intermediate trajectories (increasing and decreasing) who experience fluctuating symptom patterns eventually converge into one of these stable high- or stable low-trajectory groups by early adulthood. This aligns with the theory of statistical thermodynamics for complex systems tending to lower-energy states. As adolescents age, within-individual networks homogenize, and the symptom states align.

Conversely, between-individual networks exhibited increased variation across development, as individuals tend toward either pole of the depression spectrum. We observed this polarization as the increased variance in depression summary scores. This supports the theory of depression as a bistable system with two attractor states^{15,21}. The novelty that we propose is that these attractor states are both characterized by low-energy states but at opposite poles of the depression spectrum (symptoms ‘all on’ or ‘all off’), as depicted in Fig. 1.

Importantly, both the pattern of network temperature cooling and increased polarization drive home the cruciality of early intervention

for mitigating depression onset. The sharp changes in temperature between early to mid-adolescence indicate that when the system is at higher temperatures and displays the most randomness, we are presented with a critical period for intervention before network states become ‘stuck’. Etiologically, this period corresponds to the onset of puberty and hormone changes, synaptic pruning of neural connections and social changes, such as the transition to secondary education³⁸.

While our study emphasizes network dynamics over structure, we observed consistent network structure across all cohorts and time points. However, since our model tests are frequentist, this consistency suggests a lack of evidence for differences rather than definitive equality. Nonetheless, the large sample sizes bolster the evidence for truly equal network structures. We found the observed data fit best models with constrained edges but allowing thresholds and temperature to vary with time. This consistent equality was also found in a longitudinal network of depressed mothers³⁹. The equal network structure means centrality estimates are also equal over time. The consistency of centrality across development, as indicated by stable network structures over time, supports the transition from a node-focused approach to overall properties of the network, which offers further insight for psychological network analysis.

We also observed sex differences in the temperature of depression networks across adolescence. The males generally exhibited faster cooling and lower temperatures across adolescence compared with the females, though this sex difference was not significant in all cohorts. This pattern suggests that the symptom networks of the males stabilize earlier, while the symptom networks of the females remain more variable and unstable, potentially contributing to higher rates of depression observed in adolescent females³⁴. Our findings are supported by longitudinal work in depressive symptoms that show females follow more heterogeneous trajectory patterns with higher peaks across adolescence⁴⁰. Clinically, this may highlight an even stronger need to intervene early in males and continuously in females.

Sex is one predictor of mental health trait differences, but there are many other risk and protective factors that could make use of this analytical approach in either a cross-sectional two-group approach or longitudinally. For example, how does temperature differ in someone with poor sleep versus good sleep? How does temperature vary with substance use? How does temperature respond to depression interventions? These are opportunities for future work to explore both for depression and other psychological conditions.

Our study emphasizes the unique contributions of network models to understanding psychological systems. Unlike traditional approaches such as structural equation modeling, which infer latent constructs from aggregated questionnaire responses, network models focus on dynamic interrelations between individual symptoms, making them useful for exploring the structure and behavior of psychological systems.

While structural equation modeling relies on acyclic (no loops), theory-driven path models⁴¹, network approaches such as the Ising model can capture potentially cyclic and data-driven dynamics, offering new insights such as network temperature to enrich our understanding of system behavior.

This study had a few limitations. As expected with longitudinal research, there were many missing symptom observations (Table 1). Though our approach to handling missing data using multiple imputation and pooling mode symptom combinations is appropriate for the nature of the data, we cannot guarantee data are missing at random. For example, it is known that those with severe depression are less likely to complete longitudinal follow-up, thus leading to selective attrition bias in the imputed data⁴². Imputation for network analysis data, particularly longitudinal, is still a gap in the methodology literature^{43,44}. However, our sensitivity analyses with complete cases only found consistent results for both the model selection and pattern of temperature decrease, albeit less steep, across cohorts (Supplementary

Tables 20–22 and Supplementary Figs. 5–7). Though these approaches all have their own biases, the consistent general pattern of results provides robustness.

We leveraged three adolescent cohorts with different symptom scales (Short Mood and Feelings Questionnaire (SMFQ), Strength and Difficulties Questionnaire (SDQ) and Brief Problem Monitoring Youth Scale (BPM)), which had a mix of self-report (ABCD and ALSPAC) and parent report (MCS) and included individuals born in different generations and from different countries. Despite this heterogeneity, we show replicated results of decreasing temperature and sex differences in depression networks strengthening the generalizability of our findings. The diversity in the ABCD cohort allows for broader generalizability across different ethnic groups.

Conclusions

Overall, our work provides more evidence for the emerging dynamical systems perspective of psychological disorders. We demonstrate network temperature as a novel parameter for assessing the stability of symptom networks over time and reveal early adolescence as a critical window for intervention. Sex differences exemplify how temperature may provide more specificity of subgroup-based dynamics and prevention approaches.

Methods

Sample and depression phenotype

The participants were from the ABCD^{45,46} study from the USA, the ALSPAC^{47,48} study from the UK and the MCS^{49–51} study from the UK.

We included all individuals with at least one completed validated depression questionnaire. The longitudinal symptom data available for each cohort differed in measure, frequency and age range. In ABCD, the self-report depression symptoms were acquired from the internalizing subscale of the self-report BPM (six items) collected at eight waves between the ages of 10 and 14 years. In ALSPAC, we used the SMFQ (13 items) collected at six waves between the ages of 11 and 19 years. In MCS, we used the emotional problems subscale from the parent-report SDQ (five items) collected at three waves between the ages of 11 and 17 years. Full information on cohorts, questionnaires and validation of methods can be found in Supplementary Information and Supplementary Tables 1–6.

Symptom cleaning and multiple imputation

The symptom data for all three cohorts was scored on a three-point Likert scale (not true, sometimes and true). However, due to the Ising model specification, we binarize symptoms to $[-1, 1]$ encoding with -1 meaning ‘not true’ and 1 meaning ‘true’, where ‘sometimes’ is subsumed into ‘true’. Using the MICE R package, we used multiple imputations by chained equations with predictive mean matching to impute missing symptom data with m imputed datasets, where m is the percentage of missing data and 20 iterations. The imputed datasets were aggregated and pooled point estimates of the mode symptom combination were selected for the final imputed dataset. To determine the impact of both attrition and imputation, we estimated the networks and extracted temperature change for all cohorts with nonimputed (complete cases) datasets using the maximum-likelihood estimator.

The Ising model

The Ising model provides a framework to understand the probability distribution of different configurations in a network based on the interactions between nodes (symptoms) and external (environmental) influences. It captures how aligned node states (lower energy) are more probable and how temperature affects the randomness in the system. The Ising model equation is given by equation (1) (ref. 25)

$$P(X = x) = \frac{e^{-\beta H(x)}}{Z}, \quad (1)$$

where $P(X = x)$ is the probability of the network nodes being in a specific configuration. A configuration x represents a specific state of all nodes in the network, where each node x_i can be in one of two states (for example, -1 or 1). $H(x)$ is the potential function, which measures the energy of the configuration x with lower-energy configurations more likely to occur. The potential function captures the effect of external fields on nodes and the interactions between connected nodes. β is the inverse temperature parameter that scales the influence of $H(x)$ on the probability. A high β (low temperature) means that the system favors lower-energy states and less randomness. Z is the partition function, which sums over all possible network configurations to ensure probabilities sum to 1, acting as a normalization constant.

Network temperature

Network formulation requires all parameters—edge weights, thresholds and temperature—to be estimated in a series of cross-sectional networks across time. In the first group (time point), β (inverse temperature) is unidentified, so we fix it to 1—analogue to the latent variable identification in structural equation modeling⁵². Adding equality constraints to the groups (here, the model at different time points) means that temperature becomes identified from the second group onwards (t_2 – t_n). Consequently, temperature can increase or decrease relative to this starting value. To determine network temperature change, we extracted the thermodynamic β value from the best-fitting Ising model output at each time point. Thermodynamic β is the inverse of the temperature of the network represented in the Ising model given by equation (2) (ref. 53)

$$\beta = \frac{1}{k_B T}, \quad (2)$$

where T is the temperature and k_B is the Boltzmann constant^{53,54}. β is also termed the ‘dependence parameter’, as it scales the configuration and, thus, entropy of the Ising model^{27,54}. As β increases, the probability of the network configuration depends more on its energy, thus resulting in a higher configuration stability and lower within-system variability. The system tends toward lower Gibbs entropy due to the reduced likelihood of higher-energy configurations. Conversely, when β is at its lowest, all network configurations are equally probable, indicating a maximum Gibbs entropy.

We computed temperature as $1/\beta$ and plot the change in temperature across adolescence for each cohort. Importantly, temperature and entropy, while related, are distinct: temperature reflects the stability of the system and can fluctuate independently of overall entropy. For example, even if temperature increases (indicating higher instability), it does not necessarily imply an increase in the system’s Gibbs entropy if the overall distribution of configurations remains relatively stable.

We performed 1,000 bootstraps of network temperature estimates to derive bootstrapped 95% CIs in addition to analytical CIs.

Network estimation and model selection

Depression networks were estimated using multigroup Ising models of binary symptoms. We used the Ising model function in the psychonetrics R package to estimate network models independently at different time points. We estimated four network models in each cohort in an iterative approach with increasing constraints: (1) fully saturated with all parameters free; (2) set omega equal between groups for equal edges (structure) over time; (3) set omega and tau equal between groups for equal edges and external information (thresholds) over time, respectively; and (4) set omega, tau and beta equal between groups for equal edges, thresholds and temperature over time, respectively (Supplementary Table 7). For each of these four models, we also tested both a dense network (all nodes connected) and a sparse network (pruning of edges) structure, resulting in eight models estimated for each cohort. This approach has been detailed in previous Ising model

methodology²². The estimated network structures were plotted with the qgraph R package.

The best model was selected on the basis of fit parameters including low Akaike information criterion, low Bayesian information criterion, root mean square error of approximation <0.05 and P value ($\alpha = 0.05$) of the chi-squared test. The root mean square error of approximation compares subsequent models of increasing constraints with the fully connected (saturated) Ising model of pairwise interactions, which has 0 degrees of freedom. This approach is analogous to structural equation modeling. We performed a pairwise nested comparison between dense and sparse networks, where model 1 is the reference model. The model descriptions and nesting structure are detailed in the Supplementary Table 7.

We derived network entropy using the IsingSampler::IsingEntropy R package and function and calculated the mean symptom scores and variance across time to compare observable symptom changes with temperature and entropy.

Stratification application and sex differences

We computed the temperature change independently in males and females for the three cohorts to provide an example of how network temperature can be investigated in stratified groups. The networks were estimated, and the temperature was extracted as described for the main analysis.

We used a linear mixed-effects regression model with temperature as the outcome, age, sex and their interaction (age \times sex) as fixed effects, and a random intercept for age to examine the interaction between age and sex on network temperature. This allowed us to examine how the rate of change in network temperature differs between males and females over time. The P values are two-sided with $\alpha = 0.05$.

Sensitivity analyses

The classical Ising model denotes binary nodes as $[-1, 1]$; however, network psychometricians often prefer $[0, 1]$ encoding where symptoms are conceptualized as present or absent¹⁵. Though differently encoded networks can be transformed to produce statistically equivalent models, network parameters generate different dynamical interpretations⁵⁵. Under $[-1, 1]$ encoding, the dependence parameter (β) controls the probability of node state alignment. As β increases, the probability of nodes aligning as $[1, 1]$ or $[-1, -1]$ increases if these nodes are connected positively in the network. However, under $[0, 1]$ encoding, only the probability of $[1, 1]$ increases with β , not $[0, 0]$. We tested the impact of $[0, 1]$ encoding on temperature to determine whether this parameter requires specific encoding. We followed advised reporting methods for psychological network analysis where applicable⁵⁶.

We used the NetworkComparisonTest R package to derive global network connectivity and compare it with temperature across time points to further assess the validity and novelty of temperature. Since the networks in our analysis are structurally identical with equal edge weights, we expect connectivity to provide no additional information, with any observed differences driven by beta (inverse temperature). If network structures differed, connectivity could vary independently of temperature, with temperature potentially amplifying or suppressing connectivity effects.

Reporting summary

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

Data availability

The ALSPAC study website contains details of all the data that are available through a fully searchable data dictionary at <http://www.bristol.ac.uk/alspac/researchers/access/>. Permission to use the ALSPAC data is obtained through a proposal system managed by the ALSPAC executive. The ABCD data used in the preparation of this article were obtained

from the ABCD study at <https://abcdstudy.org>, held in the NIMH Data Archive. The ABCD data repository grows and changes over time. The ABCD data used in this report came from the NIMH Data Archive at <https://doi.org/10.15154/8873-zj65>. The DOIs can be found at <https://nda.nih.gov/abcd/>. The MCS data are freely available and can be downloaded on the UK Data Service website at <https://ukdataservice.ac.uk/>.

Code availability

All code and analysis scripts are available via GitHub at https://github.com/poppyzenzi/network_temperature.

References

- Borsboom, D. A network theory of mental disorders. *World Psychiatry* **16**, 5–13 (2017).
- van Borkulo, C. D. et al. Comparing network structures on three aspects: a permutation test. *Psychol. Methods* **28**, 1273–1285 (2023).
- Borsboom, D. & Cramer, A. O. J. Network analysis: an integrative approach to the structure of psychopathology. *Annu. Rev. Clin. Psychol.* **9**, 91–121 (2013).
- Costello, D. M., Swendsen, J., Rose, J. S. & Dierker, L. C. Risk and protective factors associated with trajectories of depressed mood from adolescence to early adulthood. *J. Consult. Clin. Psychol.* **76**, 173–183 (2008).
- Rubin, M. et al. Distinct trajectories of depression symptoms in early and middle adolescence: preliminary evidence from longitudinal network analysis. *J. Psychiatric Res.* **142**, 198–203 (2021).
- Adkins, D. E., Wang, V., Dupre, M. E., van den Oord, E. J. C. G. & Elder, G. H. Jr. Structure and stress: trajectories of depressive symptoms across adolescence and young adulthood. *Soc. Forces* **88**, 31–60 (2009).
- Kwong, A. S. F. et al. Genetic and environmental risk factors associated with trajectories of depression symptoms from adolescence to young adulthood. *JAMA Netw. Open* **2**, e196587 (2019).
- Epskamp, S., Borsboom, D. & Fried, E. I. Estimating psychological networks and their accuracy: a tutorial paper. *Behav. Res.* **50**, 195–212 (2018).
- Bringmann, L. F. et al. What do centrality measures measure in psychological networks? *J. Abnorm. Psychol.* **128**, 892–903 (2019).
- Bringmann, L. F., Lemmens, L. H. J. M., Huibers, M. J. H., Borsboom, D. & Tuerlinckx, F. Revealing the dynamic network structure of the Beck Depression Inventory-II. *Psychol. Med.* **45**, 747–757 (2015).
- Mullarkey, M. C., Marchetti, I. & Beevers, C. G. Using network analysis to identify central symptoms of adolescent depression. *J. Clin. Child Adolesc. Psychol.* **48**, 656–668 (2019).
- van Borkulo, C. et al. Association of symptom network structure with the course of depression. *JAMA Psychiatry* **72**, 1219–1226 (2015).
- Wang, K. et al. Network analysis links adolescent depression with childhood, peer, and family risk environment factors. *J. Affect. Disord.* **330**, 165–172 (2023).
- Epskamp, S. *Network Psychometrics* (Univ. Amsterdam, 2017).
- Cramer, A. O. J. et al. Major depression as a complex dynamic system. *PLoS ONE* **11**, e0167490 (2016).
- Öngür, D. & Paulus, M. P. Embracing complexity in psychiatry—from reductionistic to systems approaches. *Lancet Psychiatry* **12**, 220–227 (2025).
- Cramer, A. O. J. & Borsboom, D. Problems attract problems: a network perspective on mental disorders. in *Emerging Trends in the Social and Behavioral Sciences: An Interdisciplinary, Searchable, and Linkable Resource* (eds Scott, R. A. et al.) 1–15 (Wiley, 2015).

18. Bos, F. M. et al. Cross-sectional networks of depressive symptoms before and after antidepressant medication treatment. *Soc. Psychiatry Psychiatr. Epidemiol.* **53**, 617–627 (2018).
19. McElroy, E., Napoleone, E., Wolpert, M. & Patalay, P. Structure and connectivity of depressive symptom networks corresponding to early treatment response. *eClin. Med.* **8**, 29–36 (2019).
20. Kelley, S. W. et al. Elevated emotion network connectivity is associated with fluctuations in depression. *Proc. Natl Acad. Sci. USA* **120**, e2216499120 (2023).
21. Scheffer, M. et al. A dynamical systems view of psychiatric disorders—theory: a review. *JAMA Psychiatry* **81**, 618–623 (2024).
22. Borsboom, D. et al. Network analysis of multivariate data in psychological science. *Nat. Rev. Methods Primers* **1**, 58 (2021).
23. BRUSH, S. G. History of the Lenz–Ising Model. *Rev. Mod. Phys.* **39**, 883–893 (1967).
24. van Borkulo, C. D. et al. A new method for constructing networks from binary data. *Sci. Rep.* **4**, 5918 (2014).
25. Finnemann, A., Borsboom, D., Epskamp, S. & van der Maas, H. L. J. The theoretical and statistical Ising model: a practical guide in R. *Psych* **3**, 593–617 (2021).
26. Dalege, J., Borsboom, D., van Harreveld, F., Waldorp, L. J. & van der Maas, H. L. J. Network structure explains the impact of attitudes on voting decisions. *Sci. Rep.* **7**, 4909 (2017).
27. Dalege, J., Borsboom, D., van Harreveld, F. & van der Maas, H. L. J. The attitudinal entropy (AE) framework as a general theory of individual attitudes. *Psychol. Inquiry* **29**, 175–193 (2018).
28. Dalege, J., Borsboom, D., van Harreveld, F. & van der Maas, H. L. J. A network perspective on attitude strength: testing the connectivity hypothesis. *Soc. Psychol. Personal. Sci.* **10**, 746–756 (2019).
29. Dalege, J. & van der Does, T. Using a cognitive network model of moral and social beliefs to explain belief change. *Sci. Adv.* **8**, eabm0137 (2022).
30. Schmittmann, V. D. et al. Deconstructing the construct: a network perspective on psychological phenomena. *New Ideas Psychol.* **31**, 43–53 (2013).
31. Silvers, J. A. Adolescence as a pivotal period for emotion regulation development. *Curr. Opin. Psychol.* **44**, 258–263 (2022).
32. Bailen, N. H., Green, L. M. & Thompson, R. J. Understanding emotion in adolescents: a review of emotional frequency, intensity, instability, and clarity. *Emotion Rev.* **11**, 63–73 (2019).
33. Grimes, P. Z. et al. Genetic architectures of adolescent depression trajectories in 2 longitudinal population cohorts. *JAMA Psychiatry* **81**, 807–816 (2024).
34. Thapar, A., Eyre, O., Patel, V. & Brent, D. Depression in young people. *Lancet* **400**, 617–631 (2022).
35. McElroy, E., Fearon, P., Belsky, J., Fonagy, P. & Patalay, P. Networks of depression and anxiety symptoms across development. *J. Am. Acad. Child Adolesc. Psychiatry* **57**, 964–973 (2018).
36. López-López, J. A. et al. Trajectories of depressive symptoms and adult educational and employment outcomes. *BJPsych Open* **6**, e6 (2020).
37. Rice, F. et al. Characterizing developmental trajectories and the role of neuropsychiatric genetic risk variants in early-onset depression. *JAMA Psychiatry* **76**, 306–313 (2019).
38. Blakemore, S.-J. Adolescence and mental health. *Lancet* **393**, 2030–2031 (2019).
39. Santos, H. P. Jr, Kossakowski, J. J., Schwartz, T. A., Beebe, L. & Fried, E. I. Longitudinal network structure of depression symptoms and self-efficacy in low-income mothers. *PLoS ONE* **13**, e0191675 (2018).
40. Kwong, A. S. F. et al. Identifying critical points of trajectories of depressive symptoms from childhood to young adulthood. *J. Youth Adolesc.* **48**, 815–827 (2019).
41. Epskamp, S., Rhemtulla, M. & Borsboom, D. Generalized network psychometrics: combining network and latent variable models. *Psychometrika* **82**, 904–927 (2017).
42. Taylor, A. E. et al. Exploring the association of genetic factors with participation in the Avon Longitudinal Study of Parents and Children. *Int. J. Epidemiol.* **47**, 1207–1216 (2018).
43. Huisman, M. Imputation of missing network data: some simple procedures. *J. Soc. Struct.* **10**, 1–29 (2009).
44. Huisman, M. & Steglich, C. Treatment of non-response in longitudinal network studies. *Soc. Netw.* **30**, 297–308 (2008).
45. Volkow, N. D. et al. The conception of the ABCD study: from substance use to a broad NIH collaboration. *Dev. Cogn. Neurosci.* **32**, 4–7 (2018).
46. Jernigan, T. L., Brown, S. A. & Dowling, G. J. The Adolescent Brain Cognitive Development Study. *J. Res. Adolesc.* **28**, 154–156 (2018).
47. Fraser, A. et al. Cohort profile: the Avon Longitudinal Study of Parents and Children: ALSPAC mothers cohort. *Int. J. Epidemiol.* **42**, 97–110 (2013).
48. Boyd, A. et al. Cohort profile: the ‘children of the 90s’—the index offspring of the Avon Longitudinal Study of Parents and Children. *Int. J. Epidemiol.* **42**, 111–127 (2013).
49. Connelly, R. & Platt, L. Cohort profile: UK Millennium Cohort Study (MCS). *Int. J. Epidemiol.* **43**, 1719–1725 (2014).
50. Joshi, H. E. & Fitzsimons, E. The UK Millennium Cohort: the making of a multipurpose resource for social science and policy. *Longit. Life Course Stud.* **7**, 409–430 (2016).
51. University College London, UCL Social Research Institute, Centre for Longitudinal Studies. Millennium Cohort Study. *UK Data Service* <https://beta.ukdataservice.ac.uk/datacatalogue/series/series?id=2000031> (2024).
52. Kline, R. B. *Principles and Practice of Structural Equation Modeling* (Guilford Publications, 2023).
53. Meixner, J. Coldness and temperature. *Arch. Rational Mech. Anal.* **57**, 281–290 (1975).
54. Kindermann, R. & Snell, J. L. *Markov Random Fields and Their Applications* (American Mathematical Society, 1980).
55. Haslbeck, J. M. B., Epskamp, S., Marsman, M. & Waldorp, L. J. Interpreting the Ising model: the input matters. *Multivar. Behav. Res.* **56**, 303–313 (2021).
56. Burger, J. et al. Reporting standards for psychological network analyses in cross-sectional data. *Psychol. Methods* <https://doi.org/10.1037/met0000471> (2022).

Acknowledgements

This work has made use of the resources provided by the Edinburgh Compute and Data Facility, 2023 (<http://www.ecdf.ed.ac.uk/>). We are extremely grateful to all the families who took part in the ALSPAC study, the midwives for their help in recruiting them and the whole ALSPAC team, which includes interviewers, computer and laboratory technicians, clerical workers, research scientists, volunteers, managers, receptionists and nurses. The ABCD is a multisite, longitudinal study designed to recruit more than 10,000 children aged 9–10 years and follow them over 10 years into early adulthood. The ABCD is supported by the National Institutes of Health and additional federal partners (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 is available 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 paper reflects the views of the authors and may not reflect the opinions or views of the NIH or ABCD consortium investigators. We are grateful for the cooperation of the MCS families who voluntarily participated in the study. We are grateful to the Centre for Longitudinal Studies, UCL Social Research Institute, for the use of the MCS data and to the UK Data Service for making them available. However, neither the Centre for Longitudinal Studies nor the UK Data Service bears any responsibility for the analysis or interpretation of these data. This work was funded by a UKRI Medical Research Council grant for the Precision Medicine MRC DTP awarded to P.Z.G. (grant no. MR/W006804/1). The UK Medical Research Council and Wellcome (grant no. 217065/Z/19/Z) and the University of Bristol provide core support for ALSPAC. A.S.F.K. is supported by a Wellcome Early Career Award (grant no. 227063/Z/23/Z). This publication is the work of the authors, and P.Z.G. and A.S.F.K. serve as guarantors for the contents of this study.

Author contributions

P.Z.G. proposed and designed the study, analyzed the data and wrote the paper. S.E. and A.L.M. helped to conceive the experimental design. S.E. designed the implementation model and contributed to the code. S.E., A.S.F.K., H.C.W., A.L.M. and K.S. contributed to the interpretation of results and gave conceptual advice. S.E., A.S.F.K. and H.C.W. supervised the study. All authors (including H.L., A.G.A. and G.G.P.) contributed to the editing and revising of the paper and approved the final version.

Competing interests

H.L. reports receiving grants from Shire Pharmaceuticals; personal fees from and serving as a speaker for Medice, Shire/Takeda Pharmaceuticals and Evolan Pharma AB, all outside the submitted work. H.L. is Editor-in-Chief of JCPP Advances. The other authors declare no competing interests.

Additional information

Supplementary information The online version contains supplementary material available at <https://doi.org/10.1038/s44220-025-00415-5>.

Correspondence and requests for materials should be addressed to Poppy Z. Grimes.

Peer review information *Nature Mental Health* thanks Jonas Dalege, Haibo Yang and the other, anonymous reviewer(s) for their contribution to the peer review of this work.

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) 2025

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
<input type="checkbox"/>	<input checked="" type="checkbox"/> The exact sample size (<i>n</i>) for each experimental group/condition, given as a discrete number and unit of measurement
<input type="checkbox"/>	<input checked="" type="checkbox"/> A statement on whether measurements were taken from distinct samples or whether the same sample was measured repeatedly
<input type="checkbox"/>	<input checked="" type="checkbox"/> The statistical test(s) used AND whether they are one- or two-sided <i>Only common tests should be described solely by name; describe more complex techniques in the Methods section.</i>
<input type="checkbox"/>	<input checked="" type="checkbox"/> A description of all covariates tested
<input type="checkbox"/>	<input checked="" type="checkbox"/> A description of any assumptions or corrections, such as tests of normality and adjustment for multiple comparisons
<input type="checkbox"/>	<input checked="" type="checkbox"/> 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)
<input type="checkbox"/>	<input checked="" type="checkbox"/> For null hypothesis testing, the test statistic (e.g. <i>F</i> , <i>t</i> , <i>r</i>) with confidence intervals, effect sizes, degrees of freedom and <i>P</i> value noted <i>Give P values as exact values whenever suitable.</i>
<input checked="" type="checkbox"/>	<input type="checkbox"/> For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings
<input checked="" type="checkbox"/>	<input type="checkbox"/> For hierarchical and complex designs, identification of the appropriate level for tests and full reporting of outcomes
<input checked="" type="checkbox"/>	<input type="checkbox"/> Estimates of effect sizes (e.g. Cohen's <i>d</i> , Pearson's <i>r</i>), 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	No software was used
Data analysis	Analyses were performed in R 4.4.1. We used the following R packages: MICE (3.16.0), psychometrics (0.13), NetworkComparisonTest (v2.2.2), qqgraph (1.9.8), IsingSampler(0.2.3). All code and analysis scripts are available on GitHub: https://github.com/poppyzenzi/network_temperature .

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](#)

ALSPAC: The ALSPAC study website contains details of all the data that are available through a fully searchable data dictionary at <http://www.bristol.ac.uk/alspac/researchers/access/>. Permission to use the ALSPAC data is obtained through a proposal system managed by the ALSPAC executive.

ABCD: Data used in the preparation of this article were obtained from the Adolescent Brain Cognitive Development (ABCD) Study (<https://abcdstudy.org>), held in the NIMH Data Archive (NDA). The ABCD data repository grows and changes over time. The ABCD data used in this report came from <https://dx.doi.org/10.15154/8873-zj65>. DOIs can be found at <https://nda.nih.gov/abcd/>.
MCS: Millennium Cohort Study data are freely available and can be downloaded on the UK Data Service website (<https://ukdataservice.ac.uk/>).

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	We use the term sex, as provided by the studies themselves. Overall numbers are presented in the manuscript for percentage female at each wave of data for each study.
Reporting on race, ethnicity, or other socially relevant groupings	We use the terms as provided by the studies themselves. ALSPAC and MCS are White European cohorts. ABCD is an ethnically diverse cohort and has the following categories of self-reported ethnicity: Asian, Black, Hispanic, White and Other.
Population characteristics	<p>ALSPAC: Pregnant women resident in Avon, UK with expected dates of delivery between 1st April 1991 and 31st December 1992 were invited to take part in the study. The initial number of pregnancies enrolled was 14,541 with 13,988 children who were alive at 1 year of age. The total sample size for analyses using any data collected after the age of seven is therefore 15,447 pregnancies, resulting in 15,658 fetuses. Of these 14,901 children were alive at 1 year of age. In ALSPAC, we used the Short Mood and Feelings Questionnaire (SMFQ; 13 items) collected at 6 waves between ages 11 and 19 years. 55% of participants were female on average across the waves. ALSPAC is a White European cohort.</p> <p>ABCD: The Adolescent Brain and Cognitive Development (ABCD) study cohort has recruited a total number of 11,876 children starting in 2015 across 21 different North American study sites. In ABCD, the self-report depression symptoms were acquired from the internalising subscale of the self-report Brief Problem Monitoring scale (BPM; 6 items) collected at 8 waves between 10 and 14 years. 47% of participants were female on average across the waves. ABCD is an ethnically diverse cohort and observed self-reported ethnicity of 250 Asian, 1743 Black, 2364 Hispanic, 6134 White and 1234 Other individuals.</p> <p>MCS: The Millennium Cohort Study (MCS) which began in 2000 (birth year of participants between 2000-02), is conducted by the Centre for Longitudinal Studies (CLS). It aims to chart the conditions of social, economic and health advantages and disadvantages facing children born at the start of the 21st century. The study has been tracking the 'Millennium children' through their early childhood years and plans to follow them into adulthood. In MCS, we used the emotional problems subscale from the parent-report Strength and Difficulties Questionnaire (SDQ; 5 items) collected at 3 waves between ages 11 and 17 years. 51% of participants were female on average across the waves. MCS is a White European cohort.</p>
Recruitment	Recruitment from each study can be found in the corresponding ABCD, ALSPAC and MCS cohort study websites.
Ethics oversight	<p>ABCD: The study was approved by the National Institute of Mental Health Data Archive, United States (NIMH). Written consent was obtained from all participants.</p> <p>ALSPAC: Ethical approval for the study was obtained from the ALSPAC Ethics and Law Committee and the Local Research Ethics Committees. Consent for biological samples has been collected in accordance with the Human Tissue Act (2004). Informed consent for the use of data collected via questionnaires and clinics was obtained from participants following the recommendations of the ALSPAC Ethics and Law Committee at the time.</p> <p>MCS: The MCS was approved by the South West and London Multi-Centre Research Ethics Committees. The Millennium Cohort Study obtained informed written consent from parent/ guardians of the cohort children in order to participate in the study, children themselves as they grow-up and other participants as necessary.</p>

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	This is a quantitative study that uses network analysis with longitudinal cohort data
Research sample	<p>Participants were from the ABCD, ALSPAC and MCS longitudinal cohort studies (existing datasets).</p> <p>ALSPAC: Pregnant women resident in Avon, UK with expected dates of delivery between 1st April 1991 and 31st December 1992 were invited to take part in the study. The initial number of pregnancies enrolled was 14,541 with 13,988 children who were alive at 1 year of age. The total sample size for analyses using any data collected after the age of seven is therefore 15,447 pregnancies, resulting in 15,658 fetuses. Of these 14,901 children were alive at 1 year of age. In ALSPAC, we used the Short Mood and Feelings Questionnaire (SMFQ; 13 items) collected at 6 waves between ages 11 and 19 years. 55% of participants were female on average across the waves. ALSPAC is a White European cohort. The ALSPAC study website contains details of all the data that are available through a fully searchable data dictionary at http://www.bristol.ac.uk/alspac/researchers/access/. Permission to use the ALSPAC data</p>

is obtained through a proposal system managed by the ALSPAC executive.

ABCD: The Adolescent Brain and Cognitive Development (ABCD) study cohort has recruited a total number of 11,876 children starting in 2015 across 21 different North American study sites. In ABCD, the self-report depression symptoms were acquired from the internalising subscale of the self-report Brief Problem Monitoring scale (BPM; 6 items) collected at 8 waves between 10 and 14 years. 47% of participants were female on average across the waves. ALSPAC is a White European cohort. ABCD is an ethnically diverse cohort and observed self-reported ethnicity of 250 Asian, 1743 Black, 2364 Hispanic, 6134 White and 1234 Other individuals. Data used in the preparation of this article were obtained from the Adolescent Brain Cognitive Development (ABCD) Study (<https://abcdstudy.org>), held in the NIMH Data Archive (NDA). The ABCD data repository grows and changes over time. The ABCD data used in this report came from <https://dx.doi.org/10.15154/8873-zj65>. DOIs can be found at <https://nda.nih.gov/abcd/>.

MCS: The Millennium Cohort Study (MCS) which began in 2000 (birth year of participants between 2000-02), is conducted by the Centre for Longitudinal Studies (CLS). It aims to chart the conditions of social, economic and health advantages and disadvantages facing children born at the start of the 21st century. The study has been tracking the 'Millennium children' through their early childhood years and plans to follow them into adulthood. In MCS, we used the emotional problems subscale from the parent-report Strength and Difficulties Questionnaire (SDQ; 5 items) collected at 3 waves between ages 11 and 17 years. 51% of participants were female on average across the waves. MCS is a White European cohort. Millennium Cohort Study data are freely available and can be downloaded on the UK Data Service website (<https://ukdataservice.ac.uk/>).

Sampling strategy

ALSPAC: Participants were recruited from a geographically defined region in the UK, where pregnant women were invited to join the study. The cohort was primarily composed of mothers living in the Avon area who were due to give birth between 1991 and 1992. ABCD: Participants were selected through stratified random sampling from across the United States, ensuring representation across diverse demographic characteristics, such as race, ethnicity, and socio-economic status. MCS: Participants were selected using stratified random sampling from across the UK, ensuring a nationally representative sample that accounted for geographical regions, social class, and ethnicity.

Data collection

Data collection varied between cohorts. See the individual cohorts themselves. Briefly, ALSPAC used the Short Mood and Feelings questionnaire (SMFQ), ABCD used the Brief Problem Monitoring (BPM) scale and MCS used the Strength and Difficulties Questionnaire (SDQ). The SMFQ is a 13-item questionnaire that measures the presence of depressive symptoms in the last two weeks, and has a clinical cut-off as a total score of ≥ 11 . The SMFQ has been clinically validated against ICD-10 depression diagnosis. The Brief Problem Monitoring (internalising) is a 6-symptom subscale with items derived using factor analysis and item response theory from CBCL and YSR. BPM has a clinical T-score cut off of >65 (from the ASEBA manual). Internal consistency and correspondence with CBCL has been demonstrated. The SDQ is a brief behavioural screening questionnaire about 2- to 17-year-olds. The emotional symptoms subscale is used to measure information about mood and feelings related to depression. This subscale has a clinical cut-off of ≥ 5 for depression status. The researcher was not blinded to the study hypothesis or experimental condition, as the study aimed to analyze existing data from cohort studies.

Timing

ABCD: 09-2015 - ongoing, ALSPAC: 04-1991 - ongoing, MCS: 09-2000 - ongoing.

Data exclusions

No data were excluded from these analyses

Non-participation

All participants are included in these analyses

Randomization

Participants were not allocated into experimental groups, as this study used existing data from observational cohort studies (ALSPAC, ABCD, MCS). There was no randomisation. This was a longitudinal investigation so age was not included as a covariate. Sex was stratified for one analysis as part of the investigation.

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
<input checked="" type="checkbox"/>	<input type="checkbox"/> Antibodies
<input checked="" type="checkbox"/>	<input type="checkbox"/> Eukaryotic cell lines
<input checked="" type="checkbox"/>	<input type="checkbox"/> Palaeontology and archaeology
<input checked="" type="checkbox"/>	<input type="checkbox"/> Animals and other organisms
<input checked="" type="checkbox"/>	<input type="checkbox"/> Clinical data
<input checked="" type="checkbox"/>	<input type="checkbox"/> Dual use research of concern
<input checked="" type="checkbox"/>	<input type="checkbox"/> Plants

Methods

n/a	Involved in the study
<input checked="" type="checkbox"/>	<input type="checkbox"/> ChIP-seq
<input checked="" type="checkbox"/>	<input type="checkbox"/> Flow cytometry
<input checked="" type="checkbox"/>	<input type="checkbox"/> MRI-based neuroimaging

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, mosaicism, off-target gene editing) were examined.