

<https://doi.org/10.1038/s44184-024-00106-6>

Development of the psychopathological vulnerability index for screening at-risk youths: a Rasch model approach

Check for updates

Yujing Liao^{1,6}, Haitao Shen^{2,6}, Wenjie Duan^{1,6}✉, Shanshan Cui³, Chunxiu Zheng³, Rong Liu⁴ & Yawen Jia⁵

Accumulating research on mental health emphasizes the general factor of psychopathology (p -factor) that unites various mental health issues. This study develops a psychopathological vulnerability assessment for youths, evaluating its psychometric properties and clinical utility. An umbrella review conceptualized multifactor psychopathological vulnerability, leading to a 57-item pool. A total of 11,224 individuals participated in this study. The resulting 22-item psychopathological vulnerability index (PVI) fitted the unidimensional Rasch model, demonstrating a person separation reliability of 0.78 and a Cronbach's alpha of 0.84. Cut-off points of 11 and 5, derived from latent class analysis, were used to distinguish vulnerable and high-protection populations. The PVI's concurrent and predictive hit rates ranged from 36.00% to 53.57% in clinical samples. The PVI concretized the vulnerability–stress model for identifying at-risk youths and may facilitate universal interventions by integrating the theoretical foundations of bifactor S-1 models with key symptoms from network models for theoretically grounded approaches.

Youths undergo a developmental transition from adolescence to adulthood; this transition poses a significant challenge to youths and makes them vulnerable to mental health issues¹. The WHO World Mental Health International College Student Project suggested that over 30% of youths met the diagnostic criteria for at least one common mental health issue. The pooled prevalence of depression, anxiety symptoms, and burnout among youths was 33.6%, 39.0%, and 12.1%, respectively^{2,3}. A nationwide survey in the United States revealed clinical levels of insomnia at 17.6% and 19.0% before and during the pandemic. Additionally, the alcohol use disorder rates were 35.8% and 41.7%⁴. Experiencing mental health issues is associated with poor academic achievements⁵, increased academic drop-out rates⁶, and a high prevalence of suicide-related outcomes⁷ and non-suicidal self-injury⁸. Improving early detection and pinpointing vulnerable populations in schools, the community, and clinical settings has been highlighted⁹.

Numerous widely used measurements for detecting specific mental health issues exist. Previous literature reviews^{7,10,11} have revealed that commonly employed specific symptom instruments include the Depression Anxiety and Stress Scale-21 (DASS-21)¹², Patient Health Questionnaire-9 (PHQ-9)¹³, Generalized Anxiety Disorder 7-item Scale (GAD-7)¹⁴, and Center for Epidemiological Studies-Depression Scale¹⁵. The PHQ-9 mainly

assesses depression symptoms, whereas the GAD-7 centers on anxiety. Other instruments, such as DASS-21, can measure specific components, such as anxiety, depression, and stress. Several composite assessments have been used as alternatives. They include the Symptom Checklist-90 (SCL-90)¹⁶, General Health Questionnaire (GHQ)¹⁷, Psychological Screening Inventory¹⁸, and the Counseling Center Assessment of Psychological Symptoms (CCAPS)¹⁹.

However, none of the above specific symptom instruments can fully capture the main aspects of psychopathological status within the context of high rates of comorbidity²⁰. The aforementioned composite assessments exhibited limitations related to the measurement framework and psychometric performance. The test construction steps were somewhat unclear, with insufficient attention paid to international standards for item development^{21,22}. Additionally, the reliance on classical test theory (CTT) has limited discussions surrounding score interpretation, item bias, and the assumption of equal intervals between Likert scale response categories^{23,24}.

Developing a practical composite assessment should directly respond to a classical controversy in psychopathology: the debate over whether mental disorders are discrete clinical conditions or distinctions along continuous dimensions of functioning²⁵. Categorical diagnostic systems, such as

¹Social and Public Administration School, East China University of Science and Technology, Shanghai, China. ²Institute of Marxism, East China University of Science and Technology, Shanghai, China. ³The Office of Student Services, East China University of Science and Technology, Shanghai, China. ⁴Student Counseling and Mental Health Center, East China University of Science and Technology, Shanghai, China. ⁵School of Journalism and Communication, Nanchang University, Nanchang, China. ⁶These authors contributed equally: Yujing Liao, Haitao Shen, Wenjie Duan. ✉e-mail: duan.w@outlook.com; duan.w@ecust.edu.cn

the various editions of the American Psychiatric Association's Diagnostic and Statistical Manual (DSM) and the Mental, Behavioral or Neurodevelopmental Disorders (MBND) chapter of the International Statistical Classification of Diseases and Related Health Problems (ICD), are currently used in research and practice^{20,26}. The DSM-5, published in 2013, has expanded diagnostic criteria and classifications, addressing the complexities of mental health disorders more comprehensively than its predecessors²⁷. For decades, the ICD has been an essential tool for ensuring international comparability in public health. Initially intended primarily for the classification of causes of death, its scope has progressively expanded to include the coding and tabulation of morbidity causes, as well as medical record indexing and retrieval. The ICD-10, released by the World Health Organization in 1992, has also played a pivotal role in the global classification of diseases, including mental health disorders. Currently, the ICD-11, accepted by the World Health Assembly in May 2019, updates the classification system with changes to diagnostic criteria (e.g., a higher diagnostic threshold for schizoaffective disorders) and introduces new diagnostic entities (e.g., prolonged grief disorder and complex post-traumatic stress disorder)²⁸. Such categorical models are preferred because they offer pragmatic advantages. A diagnostic label can effectively convey important clinical information and guide treatment²⁹. However, these categorical diagnostic systems are being questioned. The greatest challenge of categorical diagnostic systems is comorbidity²⁵, wherein two or more conditions or disorders, each presumably with its own etiology, pathology, and treatment implications, coexist³⁰. Although new diagnoses have been introduced to represent the discovery of a previously unrecognized disease, they may also create additional boundary confusion³¹.

Against this background, there is increasing interest in transdiagnostic dimensional models that include the National Institute of Mental Health Research Domain Criteria (RDoC) initiative³², as well as models of the general factor of psychopathology (p -factor)²⁰. Within the RDoC framework, a "bottom-up" approach grounded in neuroscience is employed to identify biologically-based constructs and units of analysis that likely span across current diagnostic categories³³. However, the evidence incorporating this framework is relatively limited and requires further advancement to refine the model³⁴.

Models of the p -factor typically are constructed to characterize the comorbidity of common mental health issues, and three main approaches for capturing the structure of the p -factor exist: the hierarchical, bifactor, and unidimensional one-factor models^{20,35}. Many studies have used higher-order models to capture correlations among first-order factors (internalizing, externalizing, and thought disorders) and indicated that higher-order factor models tend to show superior goodness of fit over other models^{36,37}. Similarly, evidence showing that the bifactor models of psychopathology specifying the p -factor and specific factors fit better than competing structural models exists³⁸. However, hierarchical and bifactor models have limitations in the measurement and interpretability of the p -factor. The first-order factors act as mediators between the p -factor and observed symptoms in hierarchical models, prompting inquiries into how the influence of second-order factors on observed indicators remains obscured³⁹. Concerns have been raised about the bifactor model, particularly regarding overfitting, reliability, and interpretational challenges when applied to psychopathological structures^{40,41}. Furthermore, current research exploring the p -factor structure of psychopathology was predominantly based on CTT, which is possibly less effective than the Rasch measurement in reducing complex data matrices into unidimensional variables⁴², exploring item bias, and explaining the meaning of scores⁴³.

The unidimensional one-factor model can be employed to conceptualize and measure the p -factor. In such a model, each item is allowed to be loaded on the single general p -factor⁴⁴, which shows less ambiguity in terms of interpretability than higher-order and bifactor models⁴⁵. Additionally, the psychometric properties for determining the scoring and interpretation of the p -factor can be examined through unidimensional item response theory (IRT) analysis⁴⁶.

A number of unidimensional one-factor measures have been developed. They include the 12-item revised Neurological Disorders Depression Inventory-Epilepsy for Youth (NDDI-E-Y)⁴⁷, Functional Posttraumatic Cognition Questionnaire in Children and Adolescents⁴⁸, and Climate Change Anxiety Scale-Short Form (CCAS-S)⁴⁹. A unidimensional one-factor instrument specifically designed to assess the continuous general factor of psychopathology (p -factor) in youths is unavailable. In addition, most of the above instruments have not undergone item analysis, limiting the ability to estimate item biases^{47,50}. Given that the p -factor emphasizes the exploration of common etiological factors^{51,52}, existing research identifies vulnerability as a predictor of mental illness and subjective well-being^{53,54}, particularly within a developed and refined risk categorization scheme^{7,53}. In this context, vulnerability was defined as the risk of developing psychopathology and susceptibility to undesirable outcomes^{54,55}. Moreover, the vulnerability-stress model of mental disorders may explain the onset of mental health issues in youths by stating that persons have an intrinsic vulnerability, leading in combination with an extrinsic stressor to the development of health problems⁵⁶. Therefore, we argue that psychopathological vulnerability more accurately reflects the intrinsic attributes of the p -factor instrument we aim to develop.

This study aims to develop a general p -factor (i.e., the PVI) to capture a wide range of mental health issues among youths. The secondary goals include identifying the unidimensional items for the PVI, establishing cut-off points, and assessing the concurrent and predictive hit rates. The hypotheses for the present study are as follows: (1) Rasch analysis will reveal a unidimensional one-factor structure (p -factor, PVI) that represents various mental issues among youths; (2) PVI scores are expected to effectively classify youths into distinct levels of vulnerability, with corresponding cut-off points established for these levels; (3) PVI scores are positively correlated with the SCL-90 scores; and (4) the clinical hit rates of PVI are expected to be higher than that of the SCL-90 at the sample level.

Results

Item fit statistics

The infit and outfit statistics of all items were assessed. The infit ranged from 0.81 (c1) to 1.29 (i2), and the outfit ranged from 0.55 (b5) to 2.08 (f7). Following the criteria outlined in the Methods section, 35 items beyond the acceptable range of the infit and outfit were deleted⁵⁷. Finally, a 22-item PVI was developed (see Table 1). As shown in Table 2, the remaining 22 items had acceptable infit and outfit mean square error (MNSQ) statistics that ranged from 0.80 to 1.20 and were close to 1.0, indicating that the observed response pattern fits the model.

Dimensionality and local dependence

Rasch principal component analyses for residuals (PCAR) revealed that the Rasch model estimates on the logit scale accounted for 34.46% of the variance in item responses. Meanwhile, the first contrast exhibited an eigenvalue of 1.92. These results indicated the underlying one-dimensionality construct.

Differential item functioning

The differential item functioning (DIF) results are shown in Table 2. No evidence of substantive DIF for any item by academic year or age group was found because the absolute values of freshman, senior, and age DIF contrasts were less than 0.5 logits between the comparison groups. This difference in t -values was insignificant. However, four items (j1, i3, j2, and f1) showed moderate sex differences because the absolute values of sex DIF contrast were all greater than 0.5 between the comparison groups. This difference in t -values was significant, indicating systematic biases favoring either males or females in responses. Specifically, item i3 was found to be significantly more challenging for males to endorse than for females, whereas items f1, j1, and j2 were significantly more difficult for females to endorse than for males.

Table 1 | Items of PVI

Dimension		Item content
		Considering your experiences in the past month, respond to the following items by selecting either “Yes” or “No”.
Anxiety	a1	I found myself in situations that made me so anxious and I was most relieved when they ended.
	a2	I had a feeling of faintness.
Depression	b2	I felt that I had nothing to look forward to.
Learning burnout	d1	I feel nervous whenever I have an exam.
	d4	I do not have common topics and interests with my classmates.
	d5	I feel that my attitude towards studying is not as good as it used to be.
	d6	I often have trouble sleeping due to things related to my schoolwork.
Internet addiction	e2	I am short of sleep because of the Internet.
	e3	I often neglect my daily obligations (work, school, or family life) because I prefer to spend time on the Internet.
	e4	When I am feeling down, I often go on the Internet.
	e5	I find it difficult to stop using the Internet when I am online.
	f1	I have found that I was not able to stop drinking once I had started.
Sleep disorders	g3	I often have difficulty getting back to sleep after a nocturnal awakening.
	g4	I often wish for more sleep after getting up.
	g5	I am satisfied with my sleep.
Aggressive behavior	i3	Sometimes I fly off the handle for no good reason.
	i5	Other people always seem to get breaks.
Social inhibition	j1	I find it hard to start a conversation.
	j2	I do not find the right things to talk about.
	j3	I have difficulty making contact.
	j4	I often worry that others may disapprove of me.
	j5	I feel very insecure when I do not know another’s thoughts.

Item g5 was subject to reverse scoring.

Sources of validity and reliability evidence

The mean person location logit value (logit = 0.0036) indicated that the PVI was well targeted, with participants on average at a marginally higher level of psychopathological vulnerability than those tapped by the PVI items (0 logits). The person separation reliability was 0.78. The Cronbach’s alpha was 0.84. The PVI was positively correlated with the global severity index (GSI) T-score ($r = 0.60, p < 0.001$).

Latent class analysis results

Table 3 presents the fit statistics for the two- to six-class solutions identified using latent class analysis (LCA). The Akaike information criteria (AIC), Bayesian information criteria (BIC), and sample size-adjusted BIC (saBIC) continued to decrease across five classes, and the Lo-Mendell-Rubin (LMR) and bootstrap likelihood ratio test (BLRT) were significant for all comparators. The classification quality evaluated by the entropy and average latent class probabilities for the most likely class showed that among all solutions, the two-class solution was the highest, and the three- and four-class solutions were also adequate (entropy > 0.70; average latent profile probabilities > 0.80).

The three-class solution exhibited clearer boundaries than the two- and four-class solutions within the range of the PVI total scores. This situation allowed for meaningful and quantitative distinctions. Therefore, the three-class solution was adopted. For the three-class solution, the first class included 20.30% youths ($n = 2278$), the second class encompassed 46.68% youths ($n = 5239$), and the final class contained 33.03% youths ($n = 3707$). The three-class solution is presented in Fig. 1.

Receiver operator characteristic (ROC) results

ROC analysis indicated 4.50 and 10.50 as two cut-off points for distinguishing three classes. A cut-off point was determined to be 10.50, which corresponded to the maximum Youden index (96.15%) and an AUC value

of 99.75% (95% CI: 99.71–99.79%; $p < 0.001$), indicating excellent discriminatory ability. Similarly, the cut-off point of 4.50 corresponded to the maximum Youden index (97.50%) and an AUC value of 99.76% (95% CI: 99.72–99.80%; $p < 0.001$), indicating excellent discriminative performance. Table 4 presents the diagnostic indices for the two cut-off points. Given that PVI is dichotomous (i.e., “yes” or “no”), its two cut-off points were determined to be 5 and 11.

The Kruskal-Wallis (KW) and post hoc tests, as presented in Table 5, further found significant differences in PVI scores among the three classes differentiated by the two cut-off points ($p < 0.001$). Consequently, the three classes were defined as the vulnerable class (PVI scores ≥ 11 , 22.88%), medium-protection class ($5 \leq$ PVI scores < 11, 42.55%), and high-protection class (PVI scores < 5, 34.57%). Significant differences in SCL-90 scores were also observed among the three classes.

Further exploration revealed that the PVI outperformed the SCL-90 in terms of concurrent and predictive hit rates in two clinical samples, as reported in Table 6. In the concurrent sample, which consisted of data collected simultaneously with the PVI questionnaire, the PVI, and SCL-90 had concurrent hit rates of 36.00% and 28.00%, respectively, for identifying vulnerable youths with definitive diagnosis ($n = 50$) and 46.94% and 44.90%, respectively, for youths with provisional diagnosis ($n = 49$). Regarding predictive hit rates, in another sample diagnosed 1 year after the PVI questionnaire was administered, the PVI and SCL-90 had predictive hit rates of 37.50% and 28.13% for youths with definitive diagnosis ($n = 32$), respectively, and 53.57% and 46.43% for youths with provisional diagnosis ($n = 28$). In the overall diagnosis sample, the PVI and SCL-90 had hit rates of 36.59% and 28.05% for youths with definitive diagnosis ($n = 82$), respectively, and 49.35% and 45.45% for youths with provisional diagnosis ($n = 77$), respectively.

Table 2 | Item Rasch analysis results of the PVI (sorted by item difficulty)

Item	Difficulty	SE	Infit	Outfit	Sex DIF contrast	Freshman DIF contrast	Senior DIF contrast	Age DIF contrast
f1	4.94	0.08	1.02	1.20	-0.51**	0.16	-0.15	-0.20
g3	3.31	0.04	1.00	0.99	-0.15	0.02	0.05	-0.20 [†]
d4	2.60	0.03	1.04	1.16	-0.42***	0.01	-0.04	0.14 [†]
d6	1.96	0.03	1.00	0.93	0.40***	-0.01	0.05	-0.34***
a2	1.91	0.03	1.04	1.05	0.15**	-0.07	0.10	0.08
i3	1.86	0.03	1.04	1.03	0.61***	0.03	-0.04	0.15**
e3	1.78	0.03	0.96	0.87	-0.39***	-0.09	0.02	-0.06
e2	1.65	0.03	0.98	0.88	0.27***	-0.01	0.05	0.06
b2	1.38	0.02	1.01	0.95	0.02	-0.05	-0.08	-0.23***
j3	1.27	0.02	0.99	0.95	-0.14**	0.01	0.02	-0.05
i5	1.26	0.02	1.00	0.98	-0.15**	-0.04	0.01	0.04
j1	1.10	0.02	0.97	0.93	-0.66***	-0.04	0.00	0.07
j2	1.01	0.02	1.00	0.99	-0.53***	0.03	-0.02	0.00
a1	0.85	0.02	0.97	0.94	0.23***	-0.02	-0.03	-0.07
g5	0.77	0.02	1.08	1.11	0.13**	0.00	0.02	-0.31***
e5	0.56	0.02	0.97	0.93	0.20***	-0.02	-0.01	0.15**
j5	0.56	0.02	0.99	0.97	0.04	0.03	0.06	0.27***
j4	0.35	0.02	0.92	0.87	0.10 [†]	0.07	0.04	0.35***
d1	-0.32	0.02	1.08	1.12	0.39***	-0.03	-0.06	0.25***
d5	-0.42	0.02	0.96	0.95	-0.20***	-0.04	-0.02	-0.21***
e4	-0.44	0.02	1.02	1.03	0.37***	0.06	0.02	-0.06
g4	-0.99	0.02	1.05	1.13	0.22***	0.02	0.02	0.17***

Items are ordered from the least-endorsable item f1 (top) to the most-endorsable item g4 (bottom). DIF contrast above 0.50 logits are in bold.
[†] $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

Table 3 | Fit indices from latent class analysis

Model	AIC	BIC	saBIC	LMR	BLRT	Entropy	AvePP _K
2 Classes	245,642.65	245,972.31	245,829.30	<0.001	<0.001	0.83	>0.94
3 Classes^a	241,157.99	241,656.14	241,440.05	<0.001	<0.001	0.78	>0.89
4 Classes	238,337.22	239,003.87	238,714.68	<0.001	<0.001	0.78	>0.83
5 Classes	237,086.30	237,921.44	237,559.16	<0.001	<0.001	0.75	>0.77
6 Classes	236,095.44	237,099.08	236,663.71	<0.001	<0.001	0.74	>0.74

^aThe final selected solution and its fit indices from the LCA are highlighted in bold.

Discussion

This study developed a unidimensional 22-item PVI using the Rasch model. The cut-off points of 11 and 5 were identified to distinguish vulnerable and high-protection youths. At the sample level, the concurrent hit rates of the PVI were 36.00% and 46.94% for youths with definitive and provisional diagnoses, respectively. The predictive hit rates of the PVI were 37.50% and 53.7% for youths with definitive and provisional diagnoses, respectively. Psychometric evidence shows that the 22-item PVI has the potential to detect vulnerable youths.

Previously used composite instruments for screening symptoms, such as the Counseling Center Assessment of Psychological Symptoms-62 (CCAPS-62) and SCL-90, have been questioned. The CCAPS-62 was initially developed as a correlated eight-factor model¹⁹. A later study then identified two correlated second-order factors ($r = 0.11$, $p < 0.01$) in a college sample²⁴. These studies highlighted the importance of thoroughly examining the higher-order structure of the CCAPS-62. Existing research attempted to evaluate psychological vulnerability on the basis of the total score of the individual items in the SCL-90^{58,59} or SCL-90-R⁶⁰. However, such evaluation entails challenges. The SCL-90 is based on

CTT, which is well suited for identifying structures but it is not a strict test of the summed scale score being a sufficient statistic⁶¹. A study based on Rasch analysis testing the dimensionality of the SCL-90-R suggested that the total score of the SCL-90-R is multidimensional and cannot be used as an overall indicator of psychological distress⁶². The PVI, developed based on the p -factor, provides significant advantages in elucidating psychopathological phenomena, primarily due to its ability to capture the shared variance among symptoms across various disorders and diagnostic domains²⁰. A crucial interpretation that may enhance the understanding of the results of this study is that the p -factor reflects the influence of common etiological factors, including pleiotropic genetic effects, personality traits, and extensive environmental exposures^{51,52}. For instance, a genetic p -factor has been identified that correlates with the risk of developing multiple mental disorders, such as bipolar disorder, post-traumatic stress disorder, and anorexia nervosa, with a single genetic factor potentially explaining 20–60% of the variance⁶³. Building on this, ongoing research aimed at exploring how a general factor may drive the emergence of various symptoms and conditions can further enhance our comprehension of the findings presented in this paper. The

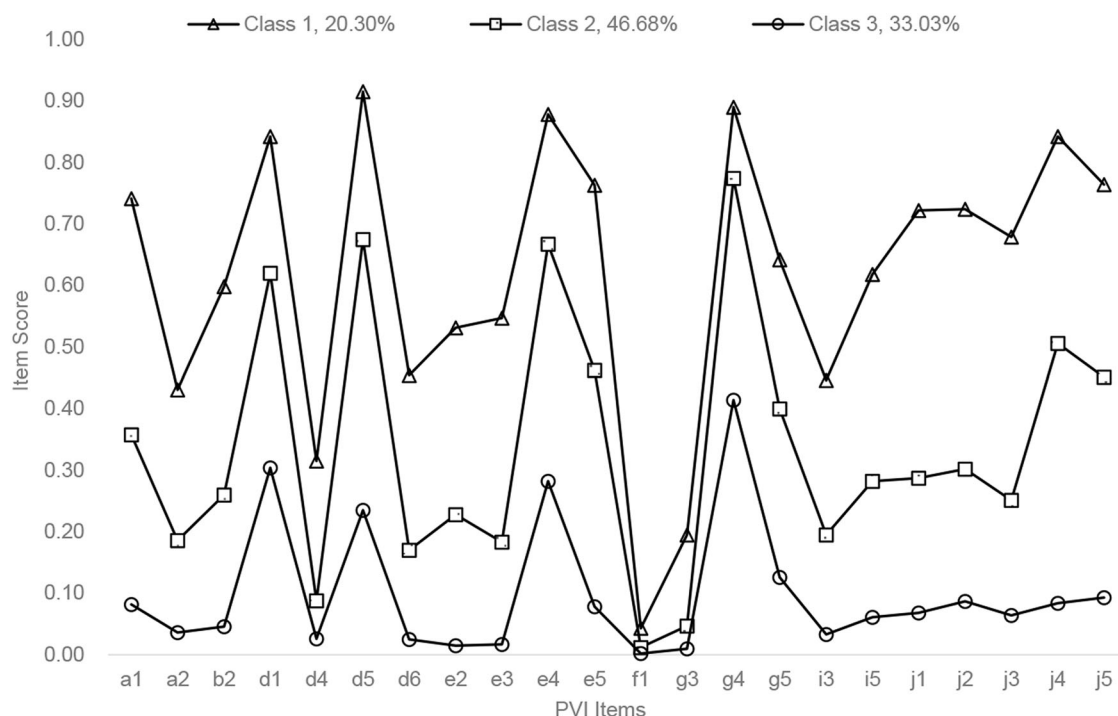


Fig. 1 | Three classes were obtained from the latent class analysis. In the figure, the youth sample of 11,224 participants is divided into three distinct classes following a latent class analysis based on their PVI scores.

Table 4 | Cut-off points and coordinates of the ROC curve for PVI

Cut-off point	Specificity, (%)	Sensitivity, (%)	Accuracy, (%)	PPV, (%)	NPV, (%)	Youden index, (%)
10.50	96.64	99.52	97.22	88.28	99.87	96.15
4.50	99.87	97.63	98.37	99.93	95.41	97.50

Table 5 | Characteristics of the latent classes

	Vulnerable class, (n = 2568)		Medium protection class, (n = 4776)		High protection class, (n = 3880)		KW test		
	Mean (SD)	Frequency	Mean, (SD)	Frequency	Mean, (SD)	Frequency	χ^2	df	p
Sex (Male)		1456 (56.70%)		2642 (55.32%)		2202 (56.75%)			
Age	19.63 (1.31)		19.57 (1.31)		19.61 (1.40)				
Grade									
Freshman		752 (29.28%)		1528 (31.99%)		1384 (35.67%)			
Sophomore		690 (26.87%)		1278 (26.76%)		892 (22.99%)			
Junior		639 (24.88%)		1088 (22.78%)		794 (20.46%)			
Senior		487 (18.96%)		882 (18.47%)		810 (20.88%)			
PVI	13.51 (2.28)		7.35 (1.69)		2.18 (1.38)		9760.28	2	<0.001
SCL-90	59.14 (13.76)		49.37 (7.19)		44.72 (4.06)		3449.86	2	<0.001

mechanism, referred to as pleiotropy, posits that DNA variants may causally relate to multiple traits associated with psychopathology. Additionally, “mediated pleiotropy” describes how genetic factors can heighten the risk for one disorder, which subsequently leads to others. This framework also considers external influences, such as adverse life events, that initiate interaction processes within these networks⁵², as emphasized by the vulnerability–stress model. Furthermore, it is suggested that DNA variants may contribute to general cognitive impairments, subsequently increasing susceptibility to various conditions⁶⁴.

Although the PVI was embodied as a unidimensional construct through the approach of IRT, the 22 retained items covered eight aspects of mental health issues, including anxiety (two items), depression (one item), learning burnout (four items), internet addiction (four items), alcohol use (one item), sleep quality (three items), aggressive behavior (two items), and social inhibition (five items). These aspects reflect the prevalent mental health issues among contemporary youths in China^{65–69}. The original reviews revealed ten aspects of mental health. However, the items related to stress and antisocial behavior were removed because they were beyond the

Table 6 | Hit rates from diagnostic samples

	Concurrent sample, (n = 99)		Predictive sample, (n = 60)		Overall diagnosis sample, (n = 159)	
	Definitive diagnosis, (n = 50) Frequency (hit rate)	Provisional diagnosis, (n = 49) Frequency (hit rate)	Definitive diagnosis, (n = 32) Frequency (hit rate)	Provisional diagnosis (n = 28) Frequency (hit rate)	Definitive diagnosis, (n = 82) Frequency (hit rate)	Provisional diagnosis, (n = 77) Frequency (hit rate)
PVI						
Vulnerable	18 (36.00%)	23 (46.94%)	12 (37.50%)	15 (53.57%)	30 (36.59%)	28 (49.35%)
Medium Protection	19 (38.00%)	20 (40.82%)	14 (43.75%)	9 (32.14%)	33 (40.24%)	29 (37.66%)
High Protection	13 (26.00%)	6 (12.24%)	6 (18.75%)	4 (14.29%)	19 (23.17%)	10 (12.99%)
SCL-90						
High-risk	14 (28.00%)	22 (44.90%)	9 (28.13%)	13 (46.43%)	23 (28.05%)	35 (45.45%)
Low-risk	36 (72.00%)	27 (55.10%)	23 (71.88%)	15 (53.57%)	59 (71.95%)	42 (54.55%)

acceptable outfit MNSQ values. Specifically, the stress-related items were removed because their outfit MNSQ values were less than 0.80, indicating a nontrivial interdependence with other items. Conversely, the antisocial behavior-related items exceeded 1.20, suggesting a poor fit within the unidimensional model⁷⁰. The results were not unexpected because stress-related items are widely recognized to overlap with anxiety and depression, and some items measure anxiety and depression, as well as stress⁷¹. The removal of antisocial behavior-related items may be due to the broader definition of antisocial behavior, which includes both violent (e.g., aggression and verbal abuse) and non-violent actions (e.g., absenteeism, truancy, and cheating)⁷². This definition differs from more commonly accepted views and likely introduced elements that extend beyond the scope of the *p*-factor construct in this study, reducing the overlap between these items and other indicators.

Among the 22 items, few showed sex-based DIF that should be discussed. Items f1 (“I have found that I was not able to stop drinking once I had started”), j1 (“I find it hard to start a conversation”), and j2 (“I do not find the right things to talk about”) showed negative sex DIF contrast values, indicating that at equal levels of psychopathological vulnerability, males were slightly more likely than females to endorse these items. Conversely, item i3 (“Sometimes I fly off the handle for no good reason”) exhibited a positive sex DIF contrast value, indicating that at equivalent levels of psychopathological vulnerability, females were slightly more likely than males to endorse this particular item. These differences may partly stem from the societal expectations and norms related to emotional expression⁷³. Taking item i3 as an example, social role theory and research related to normative expectations consistently identified women as the emotional gender^{74,75}. This phenomenon has been extensively documented in psychological and youth research, indicating that females were significantly more likely to report emotional outbursts than males⁷⁶. Research suggests that women may internalize societal expectations regarding emotional expressiveness⁷⁷, leading them to be more attuned to their emotional states and more inclined to report experiences of anger or frustration than men. This heightened emotional responsiveness among women may contribute to the endorsement of item i3. Items with DIF are retained mainly because DIF contrasts of less than 1.00 logits are considered statistically nonsignificant, and the content measured by the items represents important symptoms of the measured concept⁷⁸. Retaining items with slight DIF has been widely considered reasonable. Previous studies have retained one item each from the PHQ-9 and GAD-7 with DIF contrasts greater than 0.50 logits but less than 1.00 logits⁷⁹. Hence, items (i.e., f1, i3, j1, and j2) with slight sex DIF were retained in the PVI.

Practically, although the PVI involves eight aspects and differences across classes in various aspects exist, we recommend that future intervention strategies employ universal transdiagnostic treatment protocols. We made this recommendation because universal therapy makes theoretical and pragmatic sense^{20,80}. From a theoretical standpoint, the general *p*-

factor implies that numerous different disorders share common risk factors⁸¹. Interventions targeting different disorders comprise the same evidence-based modules (e.g., psychoeducation, setting treatment goals, and cognitive restructuring), suggesting that they treat the constituents of the *p*-factor and possibly the *p*-factor itself⁸⁰. From a practical perspective, transdiagnostic treatment protocols are being tested with several notable successes^{80,82}. Certain therapies originally developed for specific disorders have been applied in transdiagnostic interventions due to their demonstrated effectiveness across various conditions⁸³. Other protocols have been intentionally designed to address multiple related disorders, such as the unified protocol for the transdiagnostic treatment of emotional disorders⁸⁴. A meta-analysis of transdiagnostic treatments for anxiety and depressive disorders has found that these treatments resulted in significant improvement in mental health⁸².

The hit rates of the 22-item PVI and SCL-90 for identifying vulnerable youths were compared by using two clinical samples. In the overall diagnosis sample, the PVI and SCL-90 had hit rates of 36.59% and 28.05%, respectively, for youths with definitive diagnosis (*n* = 82), and 49.35% and 45.45%, respectively, for youths with provisional diagnosis (*n* = 77). PVI was found to have a better performance than the SCL-90 at the sample level. Only a few early studies have recorded the efficiency of psychological measurements among youth. The hit rate of the Children’s Depression Inventory (CDI) was reported to be 24.00%⁸⁵. A succeeding study reported hit rates of 24.00% to 28.00% for the CDI and Beck youth inventory—depression, respectively, in screening⁸⁶. Importantly, PVI identifies the majority of the intermediate olive-shaped population. Such a result may yield new insights into interventions for mental health issues, indicating a need to prioritize a preventive perspective in interventions. This preventive perspective also aligns with the *p*-factor structure given that primary prevention cannot be solely indicated by diagnosis, emphasizing the necessity of transdiagnostic prevention strategies rather than disorder-specific ones⁸⁷. Scholars have noted that a uniform etiology does not guarantee a uniform treatment response²⁰. The PVI based on the *p*-factor structure facilitated the classification of individuals in accordance with their vulnerability levels; customized interventions for each group⁸⁸, and provided the first-line intervention in a staged, stepped-care system, wherein individuals who do not improve proceed to specialized intervention²⁰. One strategy for achieving a universal psychotherapy protocol aiming to reduce high *p*-factor levels involves identifying and distilling the therapeutic elements common across disorder-specific protocols. Subsequently, these elements, which have proven efficacy, can be cobbled together into a single universal *p*-therapeutic protocol^{89,90}. While the PVI has shown evidence in effectively classifying individuals according to their vulnerability levels, incorporating alternative approaches such as network models and bifactor S-1 models could further refine the psychotherapy protocol. Network models conceptualize psychopathology as the result of interactions between specific symptoms, providing a more dynamic understanding of how symptoms influence one another²⁰. These

models emphasize the identification of “key symptoms” that may trigger other symptoms, thus allowing for targeted interventions aimed within a staged, stepped-care system. Additionally, the bifactor S-1 model can complement the PVI by defining the general factor through indicators of one facet, chosen as the reference based on strong theoretical reasoning⁹¹. This approach offers clearer interpretations of both the general and specific factors, leading to more theoretically grounded interventions.

Some limitations of this work warrant mention. First, there are concerns regarding the representativeness of the sample. Although it was drawn from a nationwide cohort, selection bias may have occurred due to the Chinese college enrollment system, leading to non-representative numbers from various provinces⁹². Furthermore, the lack of information about non-responses limited the generalizability of the study’s conclusions. Future research should aim to recruit a randomly stratified sample of youth to mitigate the effects of sample selection bias. Second, this study did not include more detailed sociodemographic characteristics of the sample, which limits the understanding of how these factors may influence the results. Nevertheless, the participants were drawn from 34 provincial-level administrative units in China, providing a diverse representation of the youth population and reflecting the broader societal context in which the PVI will be applied, future research should consider collecting additional sociodemographic data to better profile distinct vulnerability groups. Third, the concurrent and predictive validity evidence for the PVI was analyzed by using only samples collected from the student counseling and mental health center. However, it is important to note that without conducting statistical tests, no conclusions can be drawn about whether these differences are statistically significant. Therefore, future studies should incorporate inferential statistical methods to rigorously test the significance of the hit rate differences. Moreover, predictive validity evidence was examined using a 1-year follow-up sample, highlighting the importance of long-term follow-up studies in the future to further assess hit rates. The fourth limitation of this study is that among the three classes delineated by the PVI, the changing trajectory of the intermediate class in the olive-shaped distribution was not evident. In fact, this class is more deserving of attention than other classes because it has the potential to transition to either being vulnerable or highly protected. Finally, the existing PVI lacks objective indicators, such as frequencies of psychiatric admissions and emergency department visits⁹³, attention level⁹⁴, brainwaves, remote measurement technologies⁹⁵, and electroencephalography^{96,97}. The inclusion of these indicators could further enhance the predictive effectiveness of PVI.

Methods

Participants

A priori power analyses were conducted to calculate the minimum sample size needed to address the research questions in this study. Previous research suggested that for stable item calibration within ± 0.5 logits, a sample size of 64 participants was adequate if the sample was well-targeted. However, this requirement increased to 144 participants for poorly targeted samples⁹⁸. Additionally, sample size recommendations for LCA indicated a ratio of 1:3 to 1:4 between items and participants was needed to achieve a minimum power of 0.8. With an expected 30 items, our target minimum sample size was 1100 participants⁹⁹. Furthermore, G*Power 3.1 was utilized to determine the minimum necessary sample size for correlational analysis, considering an effect size of 0.6, an alpha level of 0.05, and a desired power of 0.95, the required sample size was 26 participants¹⁰⁰. Therefore, the final estimated minimum sample size for this study was 1100 participants.

Participants were voluntarily recruited from universities that admitted students through the National College Entrance Exam, which covers 34 provincial-level administrative units in China. The ratios of male students to female students across different grades have ranged from 1.26:1 to 1.63:1. This type of sample has been employed in previous studies to develop psychological instruments¹⁰¹. Students were invited to participate in an online test, which included questions on demographic information (i.e., age, sex, and grade), the self-developed initial items of the PVI, and the SCL-90.

This study was approved by the Ethics Committee of the East China University of Science and Technology (20240329010). Respondents were informed about the study’s aim and confidentiality. As recommended, all participants provided informed consent by clicking on an “I agree” button in a separate consent screen that appeared before they could access the survey¹⁰². Since participation was voluntary, some potential respondents chose not to participate for various reasons, and information about those who declined is unavailable. Furthermore, although we utilized a reliable online survey platform, the online format limited our control over test administration errors, such as test environment and potential technical issues¹⁰³. We addressed these challenges by enhancing the standardization of test administration, thereby reducing inconsistencies in administrator procedures and questionnaires.

To ensure standardized administration of the test, we collaborated with academic advisors and counselors from the student counseling and mental health center, who served as test administrators. Each administrator participated in a comprehensive training session, during which they were provided with a standardized poster containing a QR code, a link to the questionnaire, and information on the expected time commitment. Alongside the poster, a scripted invitation was provided, outlining the study’s purpose and giving instructions for inviting students to participate. This uniform communication strategy ensured that all students received consistent information about the study^{103,104}. Following the initial invitation, an operational guide was distributed to potential respondents, and administrators were instructed to frequently remind students to complete the questionnaire within the designated response period from 17 October 2022 to 25 October 2022. Additionally, to enhance data quality, we embedded attention checks throughout the questionnaire to identify careless respondents, allowing researchers to exclude them prior to analysis^{105,106}. Due to the online nature of the questionnaire, the response times varied significantly, with a median response time of 10.78 min. As indicated by a meta-analysis, the threshold of response times exhibited non-negligible differences, and there was substantial inconsistency among studies in determining the boundaries between guessing behavior and normal responses¹⁰⁷. Therefore, we primarily relied on quality control methods (i.e., attention checks) to maintain data quality^{105,106}.

The initial sample included 13,582 responses. To select qualified samples, the following exclusion criteria were adopted: (1) repetitive responses; (2) lack of identification information; (3) age outside the range of 15–24 years; (4) failure to complete the SCL-90 questionnaire; (5) failed attention checks; and (6) participants with missing values on items in the PVI item pool. Further details on the sample selection can be found in Fig. 2.

The final sample of 11,224 youth (overall sample) from all academic units had an average age of 19.60 (SD = 1.34, range = 15–24) and a male-to-female ratio of 1.28:1. Additionally, the student counseling and mental health center of the East China University of Science and Technology provided an anonymous list containing 159 participants with a definitive or provisional diagnosis of psychological disorders. Among these participants, 99 individuals (concurrent sample) underwent diagnoses concurrently with questionnaire collection. They comprised 50 individuals with a definite clinical diagnosis and 49 individuals categorized with a provisional diagnosis of psychological issues by counseling center psychologists. Another 60 individuals (predictive sample) were recruited as the diagnostic sample within 1 year after questionnaire collection. They included 32 and 28 individuals with definite clinical and provisional diagnoses, respectively. The demographic characteristics of the participants are presented in Table 7.

Development steps for PVI

The steps for the construction of the test were developed according to the Standards for Educational and Psychological Testing, which delineates four phases of the test development process: test specifications, item development and review, assembling and evaluating test forms, and developing procedures and materials for administration and scoring²¹. Additionally, the

Fig. 2 | Flowchart of the study sample. This figure provides a visual representation of the participant selection process for the study. It outlines the initial number of participants, the application of exclusion criteria, and the final number of individuals included in the analysis. PVI: psychopathological vulnerability index.

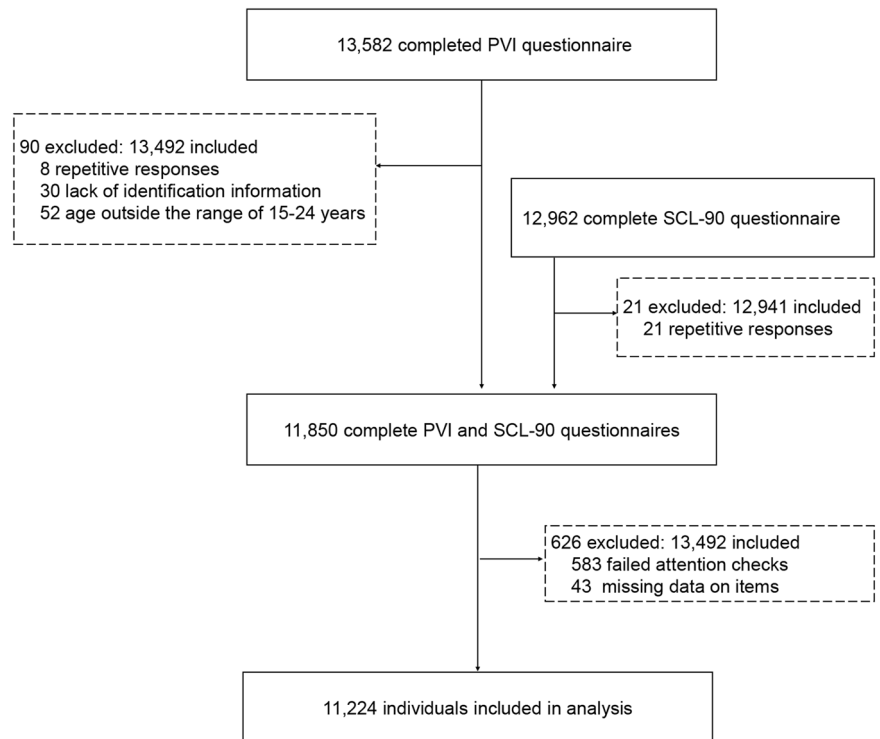


Table 7 | Demographic characteristics of the participants

Variable	Overall sample, (N = 11,224)		Concurrent sample, (n = 99)		Predictive sample, (n = 60)	
	Mean, (SD)	Frequency	Mean, (SD)	Frequency	Mean, (SD)	Frequency
Sex						
Male		6300 (56.13%)		52 (52.53%)		36 (60.00%)
Female		4924 (43.87%)		47 (47.47%)		24 (40.00%)
Age	19.60 (1.34)		20.51 (1.62)		19.08 (1.23)	
Grade						
Freshman		3664 (32.64%)		17 (17.17%)		30 (50.00%)
Sophomore		2860 (25.48%)		17 (17.17%)		16 (26.67%)
Junior		2521 (22.46%)		32 (32.32%)		10 (16.67%)
Senior		2179 (19.41%)		33 (33.33%)		4 (6.67%)

steps inspired by the test development guidelines proposed by Downing (2006), which encompass an overall plan, content definition, test specifications, item development, test design and assembly, test production, test administration, scoring responses, determining passing scores, reporting results, item banking, and producing a technical report. The test construction process also reflected the author’s experience in instrument development. Integrating these frameworks was essential because the standards for educational and psychological testing provide international benchmarks for test and item construction, while Downing’s guidelines offer best practices grounded in years of experience and research literature from various areas of psychometrics and educational measurement. By combining these sources with the author’s practical experience, this approach has been successfully applied in the development of other scales^{19,108}.

The first step involved clearly defining the purpose and rationale for the development of the PVI, which was designed to capture a broad range of mental health issues among youth. Synthesized findings from the literature helped to establish the empirical framework for PVI. The “umbrella review” has frequently been employed to identify and classify the dimensions of

measurements¹⁰⁹. This study adopted an umbrella review to identify mental health issues that youths widely encounter, as well as the most appropriate instruments and representative items. A targeted search was conducted on Web of Science by using three sets of keywords: a set of words related to mental health (e.g., psychological health, psychology, psychopathology, and psychiatric); a set of words related to systematic review (e.g., systematic review, meta-analysis, and review); and a set of words related to youth (e.g., youth, student, adolescence, and teenage). The obtained literature was screened based on titles and abstracts. Twenty-four review articles were included (as shown in Table 8). These review articles mainly summarized ten types of mental health issues (see Table 9), including anxiety, depression, stress, learning burnout, sleep disorders, alcohol abuse, Internet addiction, antisocial behavior, aggressive behavior, and social inhibition.

The second step focused on developing the item pool. The initial item pool was constructed by selecting instruments that had accumulated validated evidence in Chinese youth populations, with a focus on mental health issues identified in prior umbrella reviews. Previous research has shown that psychological measurements with multiple published translations have accumulated more evidence of cross-cultural reliability and validity than

Table 8 | Overview of the selected studies (n = 24)

Review	Location	Number of studies reviewed	Population	Sample size	Age range	Mental health issues
Pacheco, et al. ¹⁵⁹	Brazil	59	Medical students	8015	Not reported	Stress, anxiety, depression, burnout, alcohol use, low sleep quality, and excessive daytime sleepiness
Savaglio, et al. ¹⁶⁰	Australia	37	Youth	Not reported	10-25	Depression, anxiety, alcohol, and other drug use
Raccanello, et al. ¹⁶¹	Worldwide	26	Children and adolescents	Not reported	6-20	Depression, anxiety, posttraumatic stress disorder (PTSD), sleep disturbances, irritability, and social anxiety
Chan, et al. ¹⁶²	Bahrain, Kuwait, Oman, Qatar, Saudi Arabia, and the United Arab Emirates	33	Children and adolescents	Not reported	Not reported	Depression, anxiety, stress, attention deficit hyperactivity disorder (ADHD), eating disorders (ED), and tobacco use disorder
Schwarz, et al. ¹⁶³	Worldwide	28	Children and adolescents	Not reported	7-20	Depression, social anxiety, PTSD, ED, ADHD, autism, phobia, bipolar disorder, psychotic disorder, borderline personality disorder, obsessive-compulsive disorder, agoraphobia, and body dysmorphic disorder
Xie, et al. ¹⁶⁴	Worldwide	35	Children and adolescents with atopic dermatitis	521,976	0-18	Anxiety, depression, sleep disorders, conduct disorder (CD), ADHD, autism
Ma, et al. ¹⁶⁵	China and Turkey	23	Children and adolescents	57,927	Not reported	Depression, anxiety, sleep disorders, and PTSD
Li, et al. ¹⁶⁶	Worldwide	27	University student	706,415	Not reported	Depression, anxiety, stress, PTSD, tired, sleeplessness, suicidal ideation, fear, panic, somatic symptom
Jia, et al. ¹⁶⁷	Worldwide	41	Medical students	36,608	Not reported	Depression and anxiety
Ebrahim, et al. ¹⁶⁸	Worldwide	90	University students	46,284	Not reported	Depression and anxiety
Dunning, et al. ¹⁶⁸	Worldwide	33	Children and adolescents	1762	0-18	Depression and anxiety/stress
Cielo, et al. ¹⁶⁹	Worldwide	15	Young adults	18,495	18-30	Anxiety, stress, depression, and sleep disturbances
Bronsard, et al. ¹⁷⁰	Worldwide	8	Children and adolescents	Not reported	4-19	Anxiety, depression, CD, oppositional defiant disorder (ODD), ADHD, and PTSD
Becker, et al. ¹⁷¹	Worldwide	21	Children and adolescents with ADHD	Not reported	0-18	Anxiety, depression, CD, and ODD
Blackmore, et al. ¹⁷²	Worldwide	8	Refugee children and adolescents	779	0-18	Depression, anxiety, PTSD, ADHD, and ODD
Marino, et al. ¹⁷³	Worldwide	23	Adolescents and young adults	13,929	16.50-32.40	Depression and anxiety
Rodriguez-Ayllon, et al. ¹⁷⁴	Worldwide	12	Preschoolers, children, and adolescents	Not reported	2-18	Depression, anxiety, and stress
Polanczyk, et al. ¹⁷⁵	Worldwide	41	Children and adolescents	Not reported	4-18	Anxiety, depression, ADHD, and disruptive disorder (DD)
Cuijpers, et al. ¹⁷⁶	Worldwide	31	College students	Not reported	Not reported	Depression, anxiety, stress, alcohol use, smoking cessation, test anxiety, internet addiction, and procrastination
Martinez, et al. ¹⁷⁷	Worldwide	21	Adolescents with type 1 diabetes	Not reported	Not reported	Depression, social anxiety, stress, and fear
Buizza, et al. ¹⁷⁸	Worldwide	17	College students	20,108	Not reported	Anxiety, depression, anger, mania, sleep problems, alcohol use, and internet addiction
Kuder, et al. ¹⁷⁹	Worldwide	23	University students on the autism spectrum	695	17-35	Anxiety, depression, loneliness, social anxiety, and learning disorders
Zhang, et al. ¹⁸⁰	Worldwide	29	Children and adolescents	148,427	Not reported	Depression, anxiety, stress, school burnout, social anxiety
Chawla, et al. ¹⁸¹	Worldwide	102	Children and adolescents	Not reported	Not reported	Anxiety, depression, internet addiction, and sleep quality

Table 9 | Frequency of mental health issues

Mental health issues	Frequency
Anxiety	22
Depression	24
Stress	8
Sleep disorders, sleep quality, sleep problems	8
Attention deficit hyperactivity disorder (ADHD)	6
Posttraumatic stress disorder (PTSD)	6
Social anxiety	5
Disruptive disorder (DD), conduct disorder (CD), oppositional defiant disorder (ODD), aggressive and antisocial behavior	4
Alcohol use	4
Burnout, school burnout, test anxiety, tired	4
Internet addiction	3

We only included mental health issues mentioned in more than 2 reviews, except for ADHD and PTSD. The prevalences of ADHD and PTSD are primarily based on the children and adolescents population under 18-years-old^{162,163}.

tools available in only a single language^{110,111}. Ultimately, the selected instruments included: (1) the depression anxiety stress scale for anxiety, depression, and stress¹². The Cronbach's alpha values for the depression, anxiety, and stress subscales were 0.97, 0.92, and 0.95, respectively. Evidence for concurrent validity, indicated by correlations with other measures, ranged from 0.42 to 0.85, which was considered acceptable to excellent¹¹². This scale has shown evidence of good to excellent internal consistency reliability (i.e., Cronbach's alpha and omega) and evidence of convergent validity across youth samples¹¹³; (2) learning burnout scale for learning burnout¹¹⁴, showed a Cronbach's alpha of 0.77 in youth populations, with a test-retest reliability coefficient of 0.79 after two months. Confirmatory factor analysis revealed a good overall model fit, with inter-dimension correlations ranging from 0.10 to 0.48, and correlations with a total score between 0.68 and 0.82; (3) sleep quality scale for sleep disorders has shown concurrent validity evidence, evidenced by a significant correlation with the Pittsburgh sleep quality index ($r = 0.72$)¹⁰⁸. Recent findings indicated good internal consistency and test-retest reliability across diverse populations¹¹⁵; (4) alcohol use disorders identification test for alcohol use disorders¹¹⁶, has accumulated validated evidence in numerous studies¹¹⁷; (5) compulsive internet use scale for compulsive Internet use¹¹⁸, exhibited a high internal reliability (Cronbach's alpha = 0.88) and the evidence of concurrent validity was established among adolescents; (6) antisocial behavior scale for antisocial behaviors⁷², developed for youth populations, reported reliability alphas above 0.70, with sufficient predictive validity evidence established; (7) aggression questionnaire for aggressive behaviors¹¹⁹, provided evidence of construct validity and showed reliability coefficients ranging from 0.72 to 0.85 in adolescent groups¹²⁰; and (8) social inhibition questionnaire for social inhibition¹²¹, generated evidence of internal, construct, convergent, and predictive validity, with a Cronbach's alpha of 0.91 in youth populations¹²².

Previous recommendations have suggested that incorporating perspectives from researchers, clinicians, and the target population¹²³ is critical for forming an item pool. In the first stage of forming an item pool, researchers and counseling psychologists established the criteria for item selection on the basis of clinical experience in instrument development¹²⁴⁻¹²⁶ and intervention¹²⁷⁻¹²⁹. For ensuring sufficient identification and testing for homogeneity within each construct, 5-8 items with the highest factor loadings were retained for each mental health issue¹³⁰. In the second stage, in-person interviews were conducted among youths, counseling psychologists, and officers from the Office of Student Services to balance the phrasing, pertinence, lucidity, and difficulty of items. Previous research in developing the brief strengths scale for individuals with mental health issues¹³¹ and

three-dimensional inventory of character strengths for clinical and non-clinical samples¹³² have laid the foundation for the above processes. Finally, an item pool containing 57 items was developed.

The third step was to select the most appropriate items and evaluate their psychometric properties. Rasch model was used to select and evaluate items. To further provide the validity evidence, LCA was conducted to differentiate statistically distinct subgroups among the participants. Additionally, ROC analysis was used to determine cut-off points, establishing thresholds that categorize youth into different groups. A correlation analysis was performed between PVI scores and the composite scores of SCL-90 to examine the validity evidence of the PVI in relation to other variables. The SCL-90 utilized in this study was derived from the introductory paper by Derogatis, et al.¹³³, and numerous studies have shown its evidence of high reliability and validity among youth populations in China¹³⁴. All items on the SCL-90 adopted a five-point Likert scale. In accordance with previous research, the composite score used in this study was the GSI T-scores. The GSI T-scores were derived from the GSI, calculated as the mean of all 90 items of the SCL-90, and transformed into GSI T-scores using a standard conversion method with a mean of 50 and a standard deviation of 10¹³⁵. Participants who obtained GSI T-scores of 63 or above were identified as the high-risk group for psychological symptoms and those with GSI T-scores below 63 were considered as the low-risk group for psychological symptoms⁵⁹. As a widely recognized instrument, the SCL-90 could serve as a reliable benchmark for establishing the validity evidence of the PVI due to its comprehensive assessment of mental health issues^{59,72,133}.

The fourth step determined the format of the measurement. This study employed a 1-month time frame for measurement to enable a highly accurate and robust capture of vulnerable youths. A monthly interval aligns with the theoretical perspective that regards psychopathological vulnerability as a general p -factor, ensuring comprehensive assessment by encompassing most of the nonspecific mental health consequences associated with the evolution of the p -factor^{20,81}. Empirically, responses between weekly and monthly time frames at the item and scale levels were found to be no significant difference¹³⁶. By contrast, adopting long periods tends to increase recall bias¹³⁷.

In this study, we adopted dichotomous measures (i.e., "yes" or "no") instead of the original Likert scale not only for ease of use^{138,139} but also because these selected items are the manifestation of a general p -factor that exhibits an interplay pattern⁸¹. The differentiation between adjacent rating points on the Likert scale is complicated depicting distinct boundaries in the presentation of symptoms¹³⁸. Statistically, the assessment of psychopathology using the two-point scale showed high accuracy without the need to account for chance guessing and disagreements¹³⁸ and yielded a robust outcome in cases of quantity violation conditions¹⁴⁰.

Statistical analysis strategy

The study ultimately analyzed 61 variables, including sex, age, grade, 57 items from the PVI item pool, and the composite score of the SCL-90. No item in the PVI item pool exhibited extreme scores outside the range of 0 to 1, nor did any item display an extreme score of all 0s or all 1s. Only the composite score of the SCL-90 had missing data (0.14%), with no outliers present. Rasch analysis was used to derive objective, basic, and additive measures from categorical data and estimate latent traits on the basis of the unidimensional model. The dichotomous Rasch model was designed for use with ordinal data with two categories. All items were initially submitted to the Rasch model to evaluate whether different measures could be integrated into a single composite measure. First, the fit of the items to the model was examined by infit and outfit MNSQ statistics, with values between 0.80 and 1.20 indicating acceptable fit¹⁴¹. Items outside this range were considered for deletion^{57,142}. Second, Rasch PCAR was conducted to examine the assumptions of unidimensionality and local independence. The following criteria were employed to determine whether additional factors were likely to be present in residuals: (a) a cut-off of 20% of the variance explained by the trait that the index intends to measure¹⁴³ and (b) an eigenvalue of the first residual factor smaller than 2¹⁴⁴. Third, DIF was evaluated using the item

location parameter, commonly referred to as the difficulty or endorsability of an item, to examine whether measurement invariance was present¹⁴⁵. Typically, measurement invariance can be examined through IRT or CTT¹⁴⁶. In the context of IRT, a lack of measurement invariance is indicated by DIF^{145–147}. DIF analysis was applied to assess whether items functioned equivalently across groups based on sex, academic year, and age. Consistent with prior studies^{148,149}, gender comparisons were made between males and females, while age was categorized into two groups: low (\leq mean age of 19.60) and high ($>$ mean age of 19.60). Academic years were divided into two comparison groups: freshman vs non-freshman, and senior vs non-senior. DIF contrasts $<$ 0.50 logits are considered negligible, those 0.50 to 1.00 are considered moderate, and those $>$ 1.00 are considered substantial¹⁵⁰. Fourth, the mean person score should be close to 0 to indicate the good targeting of the items¹⁵¹. Finally, the person separation index ($>$ 0.70) was calculated to estimate how well one can differentiate between different individuals' performances on the measured variables¹⁴¹. The above analyses were performed in R version 4.1.2 by using the *eRm* and *TAM* packages.

The source of validity evidence in relation to other variables was then investigated. The Spearman correlation coefficient was calculated between the PVI and GSI T-score when the normal distribution assumption was violated, with values above 0.60 being preferred. Given the presence of missing values in the GSI T-score and the small proportion of missing data (less than 5%), the listwise deletion approach was employed at this stage¹⁵².

LCA was introduced to distinguish classes statistically. The goodness of fit of the model was assessed by using AIC, BIC, and saBIC. The lowest values of these criteria indicate good model fit. Furthermore, the LMR and BLRT were used to determine the appropriate number of classes. A significant LMR or BLRT suggests that a model with k classes fits the data better than a model with one less class. Finally, model classification quality was evaluated by using entropy and the average posterior class probability with values ranging from 0 to 1 and high values indicating good individual classification¹⁵³. An entropy value of 0.70 or higher represents the adequate quality of classification¹⁵⁴. LCA was performed by using Mplus 8.3. Cut-off points were determined through ROC analysis¹⁵⁵. The statistical performance of the cut-off points, such as the sensitivity, specificity, accuracy, positive predictive value (PPV), and negative predictive value (NPV), can be calculated and compared to identify the cut-off points with the best indicators¹⁵⁶. ROC analysis was conducted by using the *pROC* package in R version 4.1.2.

The KW non-parametric test was utilized to compare differences between classes determined on the basis of the PVI or SCL-90 scores. The concurrent and predictive hit rates of the PVI and SCL-90 were calculated to assess the screening performance of the two instruments in identifying psychopathological vulnerability among youths. Hit rates were defined as the percentage of correctly detected vulnerable youths relative to the overall number of diagnosed individuals^{157,158}.

Data availability

All data and code used in this analysis are publicly available at <https://osf.io/pebkh/>.

Code availability

The underlying code for this study is available in the Open Science Framework at <https://osf.io/pebkh/>.

Received: 4 May 2024; Accepted: 24 November 2024;

Published online: 02 December 2024

References

- Patel, V., Flisher, A. J., Hetrick, S. & McGorry, P. Mental health of young people: a global public-health challenge. *Lancet* **369**, 1302–1313 (2007).
- Li, W., Zhao, Z., Chen, D., Peng, Y. & Lu, Z. Prevalence and associated factors of depression and anxiety symptoms among college students: a systematic review and meta-analysis. *J. Child Psychol. Psychiatry*. **63**, 1222–1230 (2022).
- Kaggwa, M. M. et al. Prevalence of burnout among university students in low- and middle-income countries: a systematic review and meta-analysis. *PLoS One* **16**, e0256402 (2021).
- Kim, H. et al. College mental health before and during the COVID-19 pandemic: results from a nationwide survey. *Cognit. Ther. Res.* **46**, 1–10 (2022).
- Renshaw, T. L. & Cohen, A. S. Life satisfaction as a distinguishing indicator of college student functioning: further validation of the two-continua model of mental health. *Soc. Indic. Res.* **117**, 319–334 (2014).
- Mohr, C. et al. Insufficient coping behavior under chronic stress and vulnerability to psychiatric disorders. *Psychopathology* **47**, 235–243 (2014).
- Sheldon, E. et al. Prevalence and risk factors for mental health problems in university undergraduate students: a systematic review with meta-analysis. *J. Affect. Disord.* **287**, 282–292 (2021).
- Kiekens, G. et al. Non-suicidal self-injury among first-year college students and its association with mental disorders: results from the World Mental Health International College Student (WMH-ICS) initiative. *Psychol. Med.* **53**, 875–886 (2023).
- Chen, S.-P., Chang, W.-P. & Stuart, H. Self-reflection and screening mental health on Canadian campuses: validation of the mental health continuum model. *BMC Psychol.* **8**, 76 (2020).
- Ebrahim, A. H., Dhahi, A., Husain, M. A. & Jahrami, H. The psychological well-being of university students amidst COVID-19 pandemic: scoping review, systematic review and meta-analysis. *Sultan Qaboos Univ. Med. J.* **22**, 179–197 (2022).
- Irie, T., Yokomitsu, K. & Sakano, Y. Relationship between cognitive behavioral variables and mental health status among university students: a meta-analysis. *PLoS One* **14**, e0223310 (2019).
- Lovibond, P. F. & Lovibond, S. H. The structure of negative emotional states: comparison of the depression anxiety stress scales (DASS) with the Beck depression and anxiety inventories. *Behav. Res. Ther.* **33**, 335–343 (1995).
- Kroenke, K., Spitzer, R. L. & Williams, J. B. W. The PHQ-9: validity of a brief depression severity measure. *J. Gen. Intern. Med.* **16**, 606–613 (2001).
- Spitzer, R. L., Kroenke, K., Williams, J. B. W. & Löwe, B. A brief measure for assessing generalized anxiety disorder: the GAD-7. *Arch. Intern. Med.* **166**, 1092–1097 (2006).
- Radloff, L. S. The CES-D scale: a self-report depression scale for research in the general population. *Appl. Psychol. Meas.* **1**, 385–401 (1977).
- Derogatis, L. R. & Cleary, P. A. Factorial invariance across gender for the primary symptom dimensions of the SCL-90. *Br. J. Soc. Clin. Psychol.* **16**, 347–356 (1977).
- Sriram, T. G., Chandrashekar, C. R., Isaac, M. K. & Shanmugham, V. The general health questionnaire (GHQ). *Soc. Psychiatry Psychiatr. Epidemiol.* **24**, 317–320 (1989).
- Lanyon, R. I. Development and validation of a psychological screening inventory. *J. Consult. Clin. Psychol.* **35**, 1–24 (1970).
- Locke, B. D. et al. Development of the counseling center assessment of psychological symptoms-62 (CCAPS-62). *J. Couns. Psychol.* **58**, 97–109 (2011).
- Caspi, A. & Moffitt, T. E. All for one and one for all: mental disorders in one dimension. *Am. J. Psychiatry* **175**, 831–844 (2018).
- American Educational Research Association, American Psychological Association & National Council on Measurement in Education. The standards for educational and psychological testing. <https://www.testingstandards.net/open-access-files.html> (American Educational Research Association, 2014).
- Downing, S. M. Twelve steps for effective test development. In *Handbook of Test Development* (eds. Downing, S. M. & Haladyna, T. M.) 3–25 (Lawrence Erlbaum Associates, 2006).
- Locke, B. D. et al. Development and initial validation of the counseling center assessment of psychological symptoms-34. *Meas. Eval. Couns. Dev.* **45**, 151–169 (2012).

24. McAleavey, A. A. et al. Clinical validity of the counseling center assessment of psychological symptoms-62 (CCAPS-62): further evaluation and clinical applications. *J. Couns. Psychol.* **59**, 575–590 (2012).
25. Widiger, T. A. & Samuel, D. B. Diagnostic categories or dimensions? A question for the diagnostic and statistical manual of mental disorders-fifth edition. *J. Abnorm. Psychol.* **114**, 494–504 (2005).
26. Gaebel, W., Stricker, J. & Kerst, A. Changes from ICD-10 to ICD-11 and future directions in psychiatric classification. *Dialogues Clin. Neurosci.* **22**, 7–15 (2020).
27. Francés, L. et al. Current state of knowledge on the prevalence of neurodevelopmental disorders in childhood according to the DSM-5: a systematic review in accordance with the PRISMA criteria. *Child Adolesc. Psychiatry Ment. Health.* **16**, 27 (2022).
28. Nafarieh, K. et al. Prevalence of mental, behavioural or neurodevelopmental disorders according to the International Classification of diseases 11: a scoping review protocol. *BMJ Open.* **14**, e081082 (2024).
29. Trull, T. J. & Durrett, C. A. Categorical and dimensional models of personality disorder. *Annu. Rev. Clin. Psychol.* **1**, 355–380 (2004).
30. Mineka, S., Watson, D. & Clark, L. A. COMORBIDITY OF ANXIETY AND UNIPOLAR MOOD DISORDERS. *Annu. Rev. Psychol.* **49**, 377–412 (1998).
31. Wakefield, J. C. Diagnostic issues and controversies in DSM-5: return of the false positives problem. *Annu. Rev. Clin. Psychol.* **12**, 105–132 (2016).
32. Cuthbert, B. N. & Insel, T. R. Toward the future of psychiatric diagnosis: the seven pillars of RDoC. *BMC Med.* **11**, 126 (2013).
33. Becker, S. P. & Willcutt, E. G. Advancing the study of sluggish cognitive tempo via DSM, RDoC, and hierarchical models of psychopathology. *Eur. Child Adolesc. Psychiatry* **28**, 603–613 (2019).
34. Carcone, D. & Ruocco, A. C. Six years of research on the National Institute of Mental Health's Research Domain Criteria (RDoC) initiative: a systematic review. *Front. Cell. Neurosci.* **11**, 46 (2017).
35. Haywood, D., Baughman, F. D., Mullan, B. A. & Heslop, K. R. What accounts for the factors of psychopathology? An investigation of the neurocognitive correlates of internalising, externalising, and the *p*-factor. *Brain Sci.* **12**, 421 (2022).
36. 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 (2021).
37. Flora, D. B., Finkel, E. J. & Foshee, V. A. Higher order factor structure of a self-control test: evidence from confirmatory factor analysis with polychoric correlations. *Educ. Psychol. Meas.* **63**, 112–127 (2003).
38. Lahey, B. B. et al. Measuring the hierarchical general factor model of psychopathology in young adults. *Int. J. Methods Psychiatr. Res.* **27**, e1593 (2018).
39. Arrindell, W. A. et al. SCL-90-R emotional distress ratings in substance use and impulse control disorders: one-factor, oblique first-order, higher-order, and bi-factor models compared. *Psychiat Res.* **255**, 173–185 (2017).
40. 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 (2024).
41. 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).
42. Wright, B. D. Comparing Rasch measurement and factor analysis. *Struct. Equ. Model.* **3**, 3–24 (1996).
43. Boone, W. J. Rasch analysis for instrument development: Why, when, and how? *CBE Life Sci. Educ.* **15**, rm4 (2016).
44. Chen, F. F. & Zhang, Z. Bifactor models in psychometric test development. In *The Wiley Handbook of Psychometric Testing* 325–345 (Wiley Blackwell, 2018).
45. Moore, T. M. et al. Criterion validity and relationships between alternative hierarchical dimensional models of general and specific psychopathology. *J. Abnorm. Psychol.* **129**, 677–688 (2020).
46. Tiego, J. et al. Measurement of the problematic usage of the Internet unidimensional quasitrait continuum with item response theory. *Psychol. Assess.* **33**, 652–671 (2021).
47. Wagner, J. L. et al. Development and validation of the NDDI-E-Y: a screening tool for depressive symptoms in pediatric epilepsy. *Epilepsia* **57**, 1265–1270 (2016).
48. Sibeoni, J. et al. Development and validation of a patient-reported outcome in systemic sclerosis: the hand scleroderma lived experience (HANDE) scale. *Br. J. Dermatol.* **186**, 96–105 (2022).
49. Wu, J. et al. Development and validation of a youth climate anxiety scale for the youth development instrument survey. *Int. J. Ment. Health Nurs.* **32**, 1473–1483 (2023).
50. Bjornsen, H. N., Eilertsen, M.-E. B., Ringdal, R., Espnes, G. A. & Moksnes, U. K. Positive mental health literacy: development and validation of a measure among Norwegian adolescents. *BMC Public Health.* **17**, 717 (2017).
51. Xie, C. et al. A shared neural basis underlying psychiatric comorbidity. *Nat. Med.* **29**, 1232–1242 (2023).
52. Murray, A. L., Eisner, M. & Ribeaud, D. The development of the general factor of psychopathology 'p factor' through childhood and adolescence. *J. Abnorm. Child Psychol.* **44**, 1573–1586 (2016).
53. Furber, G., Leach, M., Guy, S. & Segal, L. Developing a broad categorisation scheme to describe risk factors for mental illness, for use in prevention policy and planning. *Aust. N. Z. J. Psychiatry.* **51**, 230–240 (2016).
54. Satici, S. A. Psychological vulnerability, resilience, and subjective well-being: the mediating role of hope. *Personal. Individ. Differ.* **102**, 68–73 (2016).
55. Wright, M. O. D., Masten, A. S. & Narayan, A. J. Resilience processes in development: four waves of research on positive adaptation in the context of adversity. In *Handbook of Resilience in Children* (eds. Goldstein, S. & Brooks, R. B.) 15–37 (Springer US, Boston, 2013).
56. Zubin, J. & Spring, B. Vulnerability: a new view of schizophrenia. *J. Abnorm. Psychol.* **86**, 103–126 (1977).
57. Moullin, J. C., Sklar, M., Ehrhart, M. G., Green, A. & Aarons, G. A. Provider REport of sustainment scale (PRESS): development and validation of a brief measure of inner context sustainment. *Implement. Sci.* **16**, 86 (2021).
58. Deng, J. et al. The prevalence of depressive symptoms, anxiety symptoms and sleep disturbance in higher education students during the COVID-19 pandemic: a systematic review and meta-analysis. *Psychiat. Res.* **301**, 113863 (2021).
59. Yu, Y. et al. Do preschool teachers in Southwest China need more mental health education? An online cross-sectional survey 1 year after the COVID-19 pandemic. *Front Psychol.* **13**, 907838–907838 (2022).
60. Schmitz, N., Kruse, J., Heckrath, C., Alberti, L. & Tress, W. Diagnosing mental disorders in primary care: the general health questionnaire (GHQ) and the symptom check list (SCL-90-R) as screening instruments. *Soc. Psychiatry Psychiatr. Epidemiol.* **34**, 360–366 (1999).
61. Olsen, L. R., Mortensen, E. L. & Bech, P. The SCL-90 and SCL-90R versions validated by item response models in a Danish community sample. *Acta Psychiatr. Scand.* **110**, 225–229 (2004).
62. Carrozzino, D., Patierno, C., Pignolo, C. & Christensen, K. S. The concept of psychological distress and its assessment: a clinimetric analysis of the SCL-90-R. *Int. J. Stress Manage.* **30**, 235–248 (2022).

63. Cross-Disorder Group of the Psychiatric Genomics Consortium. Identification of risk loci with shared effects on five major psychiatric disorders: a genome-wide analysis. *Lancet* **381**, 1371–1379 (2013).
64. Adam, D. “P factor” could open an important window on core attributes of mental health maladies. *Proc. Natl. Acad. Sci. USA* **120**, e2316297120 (2023).
65. Shi, H., Zhu, H. & Ni, Y. COVID-19 in China: a rapid review of the impacts on the mental health of undergraduate students. *Front. Public Health*. <https://doi.org/10.3389/fpubh.2022.940285> (2022).
66. Mak, K.-K. et al. Epidemiology of internet behaviors and addiction among adolescents in six Asian countries. *Cyberpsychol. Behav. Social Netw.* **17**, 720–728 (2014).
67. Li, L. et al. Prevalence of sleep disturbances in Chinese university students: a comprehensive meta-analysis. *J. Sleep Res.* <https://doi.org/10.1111/jsr.12648> (2018).
68. Newman, I., Ding, L. Y. & Feng, Y. H. Estimate of undergraduate university student alcohol use in China: a systematic review and meta-analysis. *Arch. Public Health*. <https://doi.org/10.1186/s13690-017-0220-x> (2017).
69. Fung, A. L. C. Adolescent reactive and proactive aggression, and bullying in Hong Kong: prevalence, psychosocial correlates, and prevention. *J. Adolesc. Health*. **64**, S65–S72 (2019).
70. Wolfe, F. & Kong, S. X. Rasch analysis of the Western Ontario MacMaster questionnaire (WOMAC) in 2205 patients with osteoarthritis, rheumatoid arthritis, and fibromyalgia. *Ann. Rheum. Dis.* **58**, 563–568 (1999).
71. Regehr, C., Glancy, D. & Pitts, A. Interventions to reduce stress in university students: a review and meta-analysis. *J. Affect. Disord.* **148**, 1–11 (2013).
72. Mahmoud, A. B., Hack-polay, D., Fuxman, L., Massetti, B. & Al Samarh, A. Z. Developing and validating a new multi-dimensional scale for anti-social behavior in a higher education setting. *Scand. J. Psychol.* **61**, 502–512 (2020).
73. Graves, B. S., Hall, M. E., Dias-Karch, C., Haischer, M. H. & Apter, C. Gender differences in perceived stress and coping among college students. *PLoS One* **16**, e0255634 (2021).
74. Grossman, M. & Wood, W. Sex differences in intensity of emotional experience: a social role interpretation. *J. Pers. Soc. Psychol.* **65**, 1010–1022 (1993).
75. Chaplin, T. M., Hong, K., Bergquist, K. & Sinha, R. Gender differences in response to emotional stress: an assessment across subjective, behavioral, and physiological domains and relations to alcohol craving. *Alcohol. Clin. Exp. Res.* **32**, 1242–1250 (2008).
76. McGee, R. & Kypri, K. Alcohol-related problems experienced by university students in New Zealand. *Aust. N. Z. J. Public Health*. **28**, 321–323 (2004).
77. Yaremko, S. K. & Lawson, K. L. Gender, internalization of expressive traits, and expectations of parenting. *Sex Roles*. **57**, 675–687 (2007).
78. Yeh, Y.-C., Park, D., Yang, S.-Y. & Kuo, C.-C. Developing and validating the Taiwan version of the meaningful activity participation assessment (T-MAPA) with Rasch analysis. *BMC Geriatr.* <https://doi.org/10.1186/s12877-023-03839-9> (2023).
79. Huang, X.-J. et al. Equating the PHQ-9 and GAD-7 to the HADS depression and anxiety subscales in patients with major depressive disorder. *J. Affect. Disord.* **311**, 327–335 (2022).
80. 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. Integ.* **28**, 253–268 (2018).
81. Southward, M. W., Cheavens, J. S. & Coccaro, E. F. Defining the p-factor: an empirical test of five leading theories. *Psychol. Med.* **53**, 2732–2743 (2022).
82. Newby, J. M., McKinnon, A., Kuyken, W., Gilbody, S. & Dalgleish, T. Systematic review and meta-analysis of transdiagnostic psychological treatments for anxiety and depressive disorders in adulthood. *Clin. Psychol. Rev.* **40**, 91–110 (2015).
83. Van Dijk, S., Jeffrey, J. & Katz, M. R. A randomized, controlled, pilot study of dialectical behavior therapy skills in a psychoeducational group for individuals with bipolar disorder. *J. Affect. Disord.* **145**, 386–393 (2013).
84. Barlow, D. H. et al. The unified protocol for transdiagnostic treatment of emotional disorders compared with diagnosis-specific protocols for anxiety disorders: a randomized clinical trial. *JAMA Psychiatry* **74**, 875–884 (2017).
85. Kaslow, N. J., Stark, K. D., Printz, B., Livingston, R. & Ling Tsai, S. Cognitive triad inventory for children: development and relation to depression and anxiety. *J. Clin. Child Psychol.* **21**, 339–347 (1992).
86. Morey, M. E., Arora, P. & Stark, K. D. MULTIPLE-STAGE SCREENING OF YOUTH DEPRESSION IN SCHOOLS. *Psychol. Schools*. **52**, 800–814 (2015).
87. Hilton, R. A. et al. Transdiagnostic neurocognitive dysfunction in children and adolescents with mental illness. *Nature Mental Health*. **2**, 299–309 (2024).
88. Duan, W. & Wang, Y. Latent profile analysis of the three-dimensional model of character strengths to distinguish at-strengths and at-risk populations. *Qual. Life Res.* **27**, 2983–2990 (2018).
89. Chorpita, B. F., Daleiden, E. L. & Weisz, J. R. Identifying and selecting the common elements of evidence based interventions: a distillation and matching model. *Ment. Health Serv. Res.* **7**, 5–20 (2005).
90. Murray, L. K. et al. A common elements treatment approach for adult mental health problems in low- and middle-income countries. *Cogn. Behav. Pract.* **21**, 111–123 (2014).
91. Eid, M. Multi-faceted constructs in abnormal psychology: implications of the bifactor S-1 model for individual clinical assessment. *J. Abnorm. Child Psychol.* **48**, 895–900 (2020).
92. Li, T. Financial decentralization and geographical stratification of access to higher education in china: the case of shanghai. *Chin. Sociol. Rev.* **49**, 212–238 (2017).
93. Thompson, R., Hornigold, R., Page, L. & Waite, T. Associations between high ambient temperatures and heat waves with mental health outcomes: a systematic review. *Public Health*. **161**, 171–191 (2018).
94. Schatzberg, A. F. et al. Neuropsychological deficits in psychotic versus nonpsychotic major depression and no mental illness. *Am. J. Psychiatry*. **157**, 1095–1100 (2000).
95. Walsh, A. E. L. et al. A collaborative realist review of remote measurement technologies for depression in young people. *Nat. Hum. Behav.* **8**, 480–492 (2024).
96. Ali, A. et al. A single-channel wireless EEG headset enabled neural activities analysis for mental healthcare applications. *Wirel. Personal Commun.* **125**, 3699–3713 (2022).
97. Yasin, S. et al. EEG based major depressive disorder and bipolar disorder detection using neural networks: a review. *Comput. Methods Programs Biomed.* **202**, 106007 (2021).
98. Ramp, M., Khan, F., Misajon, R. A. & Pallant, J. F. Rasch analysis of the multiple sclerosis impact scale (MSIS-29). *Health Qual. Life Outcomes* **7**, 58 (2009).
99. Dziak, J. J., Lanza, S. T. & Tan, X. Effect size, statistical power, and sample size requirements for the bootstrap likelihood ratio test in latent class analysis. *Struct. Equ. Model.* **21**, 534–552 (2014).
100. Faul, F., Erdfelder, E., Buchner, A. & Lang, A.-G. Statistical power analyses using G*Power 3.1: Tests for correlation and regression analyses. *Beh. Res. Meth.* **41**, 1149–1160 (2009).
101. Kircaburun, K. et al. Development and validation of the Mukbang addiction scale. *Int J Mental Health Ad.* **19**, 1031–1044 (2021).
102. Granello, D. H. & Wheaton, J. E. Online data collection: strategies for research. *J. Couns. Dev.* **82**, 387–393 (2004).

103. Stantcheva, S. How to run surveys: a guide to creating your own identifying variation and revealing the invisible. *Annu. Rev. Econ.* **15**, 205–234 (2023).
104. Fischer, R. & Milfont, T. L. Standardization in psychological research. *Int. J. Psychol. Res.* **3**, 88–96 (2010).
105. Berinsky, A. J., Margolis, M. F. & Sances, M. W. Separating the shirkers from the workers? Making sure respondents pay attention on self-administered surveys. *Am. J. Polit. Sci.* **58**, 739–753 (2014).
106. Zickar, M. J. & Keith, M. G. Innovations in sampling: improving the appropriateness and quality of samples in organizational research. *Annu. Rev. Organ. Psychol. Organ. Behav.* **10**, 315–337 (2023).
107. Rios, J. A. & Deng, J. Does the choice of response time threshold procedure substantially affect inferences concerning the identification and exclusion of rapid guessing responses? A meta-analysis. *Large Scale Assess. Educ.* **9**, 18 (2021).
108. Yi, H., Shin, K. & Shin, C. Development of the sleep quality scale. *J. Sleep Res.* **15**, 309–316 (2006).
109. Carini, E. et al. Assessing hospital performance indicators. What dimensions? Evidence from an umbrella review. *BMC Health Serv. Res.* **20**, 1038 (2020).
110. Wang, Q., She, Z., Xi, J., Ding, F. & Xu, Z. Psychometric evaluation of the Chinese version of the revised spontaneity assessment inventory (SAI-R-C). *Arts Psychother.* **80**, 101935 (2022).
111. Zhang, B. et al. The big five inventory–2 in China: a comprehensive psychometric evaluation in four diverse samples. *Assessment.* **29**, 1262–1284 (2021).
112. Antony, M. M., Bieling, P. J., Cox, B. J., Enns, M. W. & Swinson, R. P. Psychometric properties of the 42-item and 21-item versions of the depression anxiety stress scales in clinical groups and a community sample. *Psychol. Assess.* **10**, 176–181 (1998).
113. Dwight, A. R., Briesch, A. M., Hoffman, J. A. & Rutt, C. Systematic review of the psychometric evidence supporting use of the depression anxiety stress scales, short form (DASS-21) with youth. *Child Youth Care Forum.* **53**, 1235–1250 (2024).
114. Qiu, B., Zhou, K. & Xia, Y. Daxuesheng xuexijuandailiangbiao de bianzhi ji xinxiadujianyan [Development and validation of the college student learning burnout scale]. *Adv. Social Sci.* **9**, 1950–1957 (2020).
115. Huang, M. et al. Assessing psychometric properties and measurement invariance of the sleep quality questionnaire among healthcare students. *BMC Psychol.* **12**, 41 (2024).
116. Babor, T. F., Higgins-Biddle, J. C., Saunders, J. B. & Monteiro, M. G. *The alcohol use disorders identification test* (World Health Organization, Geneva, 2001).
117. de Meneses-Gaya, C., Zuardi, A. W., Loureiro, S. R. & Crippa, J. A. S. Alcohol use disorders identification test (AUDIT): an updated systematic review of psychometric properties. *Psychol. Neurosci.* **2**, 83–97 (2009).
118. Dhir, A., Chen, S. & Nieminen, M. Psychometric validation of the compulsive internet use scale: relationship with adolescents' demographics, ICT accessibility, and problematic ICT use. *Soc. Sci. Comput. Rev.* **34**, 197–214 (2015).
119. Buss, A. H. & Perry, M. The aggression questionnaire. *J Pers Soc Psychol.* **63**, 452–459 (1992).
120. Franco, C., Amutio, A., López-González, L., Oriol, X. & Martínez-Taboada, C. Effect of a mindfulness training program on the impulsivity and aggression levels of adolescents with behavioral problems in the classroom. *Front Psychol.* **7**, 1385 (2016).
121. Denollet, J. & Duijndam, S. The multidimensional nature of adult social inhibition: inhibition, sensitivity and withdrawal facets of the SIQ15. *J. Affect. Disord.* **245**, 569–579 (2019).
122. Duijndam, S., Karreman, A., Denollet, J. & Kupper, N. Physiological and emotional responses to evaluative stress in socially inhibited young adults. *Biol. Psychol.* **149**, 107811 (2020).
123. Siddaway, A. P., Wood, A. M., O'Carroll, R. E. & O'Connor, R. C. Characterizing self-injurious cognitions: development and validation of the suicide attempt beliefs scale (SABS) and the nonsuicidal self-injury beliefs scale (NSIBS). *Psychol. Assess.* **31**, 592–608 (2019).
124. Duan, W. & Li, Y. Convergent, discriminant, and incremental validities of person–environment fit scale for creativity in predicting innovative behavior. *J. Creat. Behav.* **54**, 525–534 (2020).
125. Duan, W. & Mu, W. Validation of a Chinese version of the stress overload scale-short and its use as a screening tool for mental health status. *Qual. Life Res.* **27**, 411–421 (2018).
126. Liu, R. & Duan, W. Assessing perceptions of resilience: the understanding from network analysis. *Front. Public Health.* <https://doi.org/10.3389/fpubh.2023.1017871> (2023).
127. Bu, H. & Duan, W. Strength-based flourishing intervention to promote resilience in individuals with physical disabilities in disadvantaged communities: a randomized controlled trial. *Res. Soc. Work. Pract.* **31**, 53–64 (2020).
128. Duan, W. et al. The online strength-informed acceptance and commitment therapy among COVID-19-affected adolescents. *Res. Soc. Work. Pract.* **32**, 465–474 (2022).
129. Duan, W., Yu, X. & Tang, X. “Humor ABC” program: specific strength intervention in facilitating the positive development of left-behind children. *J. Happiness. Stud.* **24**, 1–20 (2023).
130. Marsh, H. W., Hau, K.-T., Balla, J. R. & Grayson, D. Is more ever too much? The number of indicators per factor in confirmatory factor analysis. *Multivar. Behav. Res.* **33**, 181–220 (1998).
131. Ho, S. M. Y. et al. A brief strengths scale for individuals with mental health issues. *Psychol. Assess.* **28**, 147–157 (2016).
132. Duan, W. & Bu, H. Development and initial validation of a short three-dimensional inventory of character strengths. *Qual. Life Res.* **26**, 2519–2531 (2017).
133. Derogatis, L. R., Lipman, R. S. & Covi, L. SCL-90: an outpatient psychiatric rating scale—preliminary report. *Psychopharmacol. Bull.* **9**, 13–28 (1973).
134. Chen, F. et al. Identification and characterization of college students with attenuated psychosis syndrome in China. *Psychiat. Res.* **216**, 346–350 (2014).
135. Pedersen, G. & Karterud, S. Is SCL-90R helpful for the clinician in assessing DSM-IV symptom disorders? *Acta Psychiatr. Scand.* **110**, 215–224 (2004).
136. Lai, J.-S., Cook, K., Stone, A., Beaumont, J. & Cella, D. Classical test theory and item response theory/Rasch model to assess differences between patient-reported fatigue using 7-day and 4-week recall periods. *J. Clin. Epidemiol.* **62**, 991–997 (2009).
137. Robinson, M. D. & Clore, G. L. Belief and feeling: evidence for an accessibility model of emotional self-report. *Psychol. Bull.* **128**, 934–960 (2002).
138. Keeley, J. W. et al. Psychotic disorder symptom rating scales: are dichotomous or multi-point scales more clinically useful?—an ICD-11 field study. *Schizophr. Res.* **202**, 254–259 (2018).
139. Lecomte, T., Wallace, C. J., Caron, J., Perreault, M. & Lecomte, J. Further validation of the client assessment of strengths interests and goals. *Schizophr. Res.* **66**, 59–70 (2004).
140. Donner, A. & Eliasziw, M. Statistical implications of the choice between a dichotomous or continuous trait in studies of interobserver agreement. *Biometrics.* **50**, 550–555 (1994).
141. Franco, F., Gabriele, M., Andrea, G., Paolo, V. & Adriano, C. Evidence of multidimensionality in the ALSFRS-R scale: a critical appraisal on its measurement properties using Rasch analysis. *J. Neurol. Neurosurg. Psychiatry.* **84**, 1340 (2013).
142. Tennant, A. & Conaghan, P. G. The Rasch measurement model in rheumatology: What is it and why use it? When should it be applied, and what should one look for in a Rasch paper? *Arthritis Care Res.* **57**, 1358–1362 (2007).

143. Reckase, M. D. Unifactor latent trait models applied to multifactor tests: results and implications. *J. Educ. Statistics* **4**, 207–230 (1979).
144. Catley, M. J., O'Connell, N. E. & Moseley, G. L. How good is the neurophysiology of pain questionnaire? A Rasch analysis of psychometric properties. *J. Pain*. **14**, 818–827 (2013).
145. Bauer, D. J. A more general model for testing measurement invariance and differential item functioning. *Psychol. Methods*. **22**, 507–526 (2017).
146. Rouquette, A., Hardouin, J.-B., Vanhaesebrouck, A., Sébille, V. & Coste, J. Differential item functioning (DIF) in composite health measurement scale: recommendations for characterizing DIF with meaningful consequences within the Rasch model framework. *PLoS One* **14**, e0215073 (2019).
147. Martinez, H. S., Klanecky, A. K. & McChargue, D. E. Problem drinking among at-risk college students: the examination of Greek involvement, freshman status, and history of mental health problems. *J. Am. Coll. Health*. **66**, 579–587 (2018).
148. Skogen, J. C., Thørrisen, M. M., Olsen, E., Hesse, M. & Aas, R. W. Evidence for essential unidimensionality of AUDIT and measurement invariance across gender, age and education. Results from the WIRUS study. *Drug Alcohol Depend.* **202**, 87–92 (2019).
149. Wang, J., Li, C., Meng, X. & Liu, D. Validation of the Chinese version of the procrastination at work scale. *Front. Psychol.* **12**, 726595 (2021).
150. Cordier, R. et al. Using Rasch analysis to evaluate the reliability and validity of the swallowing quality of life questionnaire: an item response theory approach. *Dysphagia* **33**, 441–456 (2018).
151. Pallant, J. F. & Tennant, A. An introduction to the Rasch measurement model: an example using the hospital anxiety and depression scale (HADS). *Br. J. Clin. Psychol.* **46**, 1–18 (2007).
152. Graham, J. W. Missing data analysis: making it work in the real world. *Annu. Rev. Psychol.* **60**, 549–576 (2009).
153. Thullen, M. J., Taliaferro, L. A. & Muehlenkamp, J. J. Suicide ideation and attempts among adolescents engaged in risk behaviors: a latent class analysis. *J. Res. Adolesc.* **26**, 587–594 (2016).
154. Zheng, D. D. et al. Multimorbidity patterns and their relationship to mortality in the US older adult population. *PLoS One* **16**, e0245053 (2021).
155. Garrett, E. S., Eaton, W. W. & Zeger, S. Methods for evaluating the performance of diagnostic tests in the absence of a gold standard: a latent class model approach. *Stat. Med.* **21**, 1289–1307 (2002).
156. Bányai, F. et al. Problematic social media use: results from a large-scale nationally representative adolescent sample. *PLoS One*. **12**, e0169839 (2017).
157. Hanna, A. & Rounds, J. How accurate are interest inventories? A quantitative review of career choice hit rates. *Psychol. Bull.* **146**, 765–796 (2020).
158. Hansen, J.-I. C. & Dik, B. J. Evidence of 12-year predictive and concurrent validity for SII occupational scale scores. *J. Vocat. Behav.* **67**, 365–378 (2005).
159. Pacheco, J. P. et al. Mental health problems among medical students in Brazil: a systematic review and meta-analysis. *Rev. Bras. Psiquiatr.* **39**, 369–378 (2017).
160. Savaglio, M., O'Donnell, R., Hatzikiiriakidis, K., Vicary, D. & Skouteris, H. The impact of community mental health programs for Australian youth: a systematic review. *Clin. Child Fam. Psychol. Rev.* **25**, 573–590 (2022).
161. Raccanello, D., Rocca, E., Vicentini, G. & Brondino, M. Eighteen months of COVID-19 pandemic through the lenses of self or others: a meta-analysis on children and adolescents' mental health. *Child Youth Care Forum*. **52**, 737–760 (2023).
162. Chan, M. F. et al. Child and adolescent mental health disorders in the GCC: a systematic review and meta-analysis. *Int. J. Pediatr. Adolesc. Med.* **8**, 134–145 (2021).
163. Schwarz, S., Grasmann, D., Schreiber, F. & Stangier, U. Mental imagery and its relevance for psychopathology and psychological treatment in children and adolescents: a systematic review. *Int. J. Cogn. Ther.* **13**, 303–327 (2020).
164. Xie, Q.-W., Dai, X., Tang, X., Chan, C. H. Y. & Chan, C. L. W. Risk of mental disorders in children and adolescents with atopic dermatitis: a systematic review and meta-analysis. *Front Psychol.* **10**, 1773 (2019).
165. Ma, L. et al. Prevalence of mental health problems among children and adolescents during the COVID-19 pandemic: a systematic review and meta-analysis. *J. Affect. Disord.* **293**, 78–89 (2021).
166. Li, Y., Wang, A., Wu, Y., Han, N. & Huang, H. Impact of the COVID-19 pandemic on the mental health of college students: a systematic review and meta-analysis. *Front. Psychol.* **12**, 846789 (2021).
167. Jia, Q. et al. Mental health among medical students during COVID-19: a systematic review and meta-analysis. *Front Psychol.* **13**, 846789 (2022).
168. Dunning, D. L. et al. Research review: the effects of mindfulness-based interventions on cognition and mental health in children and adolescents—a meta-analysis of randomized controlled trials. *J. Child Psychol. Psychiatry* **60**, 244–258 (2019).
169. Cielo, F., Ulberg, R. & Di Giacomo, D. Psychological impact of the COVID-19 outbreak on mental health outcomes among youth: a rapid narrative review. *Int. J. Environ. Res. Public Health*. **18**, 6067 (2021).
170. Bronsard, G. et al. The prevalence of mental disorders among children and adolescents in the child welfare system: a systematic review and meta-analysis. *Medicine* **95**, 2622 (2016).
171. Becker, S. P., Luebke, A. M. & Langberg, J. M. Co-occurring mental health problems and peer functioning among youth with attention-deficit/hyperactivity disorder: a review and recommendations for future research. *Clin. Child Fam. Psychol. Rev.* **15**, 279–302 (2012).
172. Blackmore, R. et al. Systematic review and meta-analysis: the prevalence of mental illness in child and adolescent refugees and asylum seekers. *J. Am. Acad. Child Adolesc. Psychiatry* **59**, 705–714 (2020).
173. Marino, C., Gini, G., Vieno, A. & Spada, M. M. The associations between problematic Facebook use, psychological distress and well-being among adolescents and young adults: a systematic review and meta-analysis. *J. Affect. Disord.* **226**, 274–281 (2018).
174. Rodriguez-Ayllon, M. et al. Role of physical activity and sedentary behavior in the mental health of preschoolers, children and adolescents: a systematic review and meta-analysis. *Sports Med.* **49**, 1383–1410 (2019).
175. Polanczyk, G. V., Salum, G. A., Sugaya, L. S., Caye, A. & Rohde, L. A. Annual research review: a meta-analysis of the worldwide prevalence of mental disorders in children and adolescents. *J. Child Psychol. Psychiatry*. **56**, 345–365 (2015).
176. Cuijpers, P. et al. Prevention and treatment of mental health and psychosocial problems in college students: an umbrella review of meta-analyses. *Clin. Psychol. Sci. Pract.* **28**, 229–244 (2021).
177. Martinez, K. et al. Psychological factors associated with diabetes self-management among adolescents with type 1 diabetes: a systematic review. *J. Health Psychol.* **23**, 1749–1765 (2018).
178. Buizza, C., Bazzoli, L. & Ghilardi, A. Changes in college students mental health and lifestyle during the COVID-19 pandemic: a systematic review of longitudinal studies. *Adolesc. Res. Rev.* **7**, 537–550 (2022).
179. Kuder, S. J., Accardo, A. L. & Bomgardner, E. M. Mental health and university students on the autism spectrum: a literature review. *Rev. J. Autism Dev. Disord.* **8**, 421–435 (2021).
180. Zhang, J., Yang, S. X., Wang, L., Han, L. H. & Wu, X. Y. The influence of sedentary behaviour on mental health among children and adolescents: a systematic review and meta-analysis of longitudinal studies. *J. Affect. Disord.* **306**, 90–114 (2022).

181. Chawla, N., Tom, A., Sen, M. S. & Sagar, R. Psychological impact of COVID-19 on children and adolescents: a systematic review. *Indian J. Psychol. Med.* **43**, 294–299 (2021).

Acknowledgements

We would like to thank Yao Li from the Student Counseling and Mental Health Center at East China University of Science and Technology for his assistance in collecting the survey data. This study was funded by the Natural Science Foundation of Shanghai (grant 23ZR1415000). The funder played no role in the study design, data collection, analysis, and interpretation of data, or the writing of this manuscript.

Author contributions

Y.L., H.S., and W.D. drafted the manuscript. Y.L. and W.D. conducted the data analyses. All authors (S.C., C.Z., R.L., and Y.J.) provided valuable interpretations of the data, made significant edits to the manuscript, and approved the final version of the manuscript for submission.

Competing interests

The authors declare no competing interests.

Additional information

Correspondence and requests for materials should be addressed to Wenjie Duan.

Reprints and permissions information is available at <http://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-NonCommercial-NoDerivatives 4.0 International License, which permits any non-commercial use, sharing, 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 you modified the licensed material. You do not have permission under this licence to share adapted material derived from this article or parts of it. 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-nc-nd/4.0/>.

© The Author(s) 2024