**Assessing the heterogeneity of 27 Chinese scales for screening depression among children, adolescents, and young adults**

Haoyuan Wang 1, Mengzhen Hu 1, Liuqing Tian 1, Weibiao Liu 1, Yuanyuan An1, Ying Li 2, Hu Chuan-Peng 1#

(1School of Psychology, Nanjing Normal University, Nanjing, China; 2首都医科大学附属北京儿童医院精神科)

**CrediT Author Statement**

**Haoyuan Wang：**Conceptualization, Data curation, Investigation, Writing - Original Draft. **Mengzhen Hu：**Data curation, Visualization, Investigation. **Liuqing Tian:** Investigation. **Weibiao Liu:** Investigation. **Yuanyuan An:** Writing- Reviewing and Editing. **Ying Li:** Investigation, Writing- Reviewing, and Editing. **Chuan-Peng Hu：**Conceptualization, Supervision**,** Investigation, Project administration, Writing-Reviewing and Editing.

Corresponding author: Hu Chuan-Peng, email: [hcp4715@hotmail.com](mailto:hcp4715@hotmail.com)

# 1. Introduction

[儿童青少年抑郁障碍的严重性（患病率、社会成本等）]

Major depression, a highly prevalent mental disorder, imposes substantial personal, social, and economic burdens both in China (Fu and Zhang 2023) and globally (Herrman et al. 2022; McGrath et al. 2023). Adolescence is a critical period for the onset of depression: depression can occur as early as the age of 10 (Dattani, 2022), with a peak onset at 19.5 years (Solmi et al. 2022). Moreover, depression during the adolescence also accompanied with self-harm and suicide (Zeynep Başgöze et al. 2021), resulting in over 800,000 deaths annually among individuals aged 15-29 (Amaltinga and Mbinta 2020). The prevention and alleviation of depression are urgent issues in China (Citation) and globally (UNICEF n.d.).

While the effectiveness of interventions has garnered significant attention (Citations), an equally important issue is often overlooked: how to screen or measure the depression (Fried, Flake, and Robinaugh 2022). Choosing different measurement for depression often results in significantly different reported prevalence rates of depression. For example, when employing a strict, clinical interview-based standard, Deng et al. (2023) reported that the prevalence of major depression among Chinese children aged 6-16 years is approximately 2% to 3% (Deng et al., 2023; Li et al., 2022). In contrast, studies using self-reported scales for depression screening report much higher prevalence rates: 14.6% among elementary school students, 23.6% to 24.2% among middle school students, and 28.0% among high school students (Chen, Zhang, and Yu., 2022; Huang, Zhang, and Yu 2022; Yu, Zhang, and Yu 2022; Zhang, Jin, and Yu 2022). Also, recent meta-analyses reported the moderation of the scales used on the prevalence rates of depression among children and adolescents (Chen et al. 2022; Huang et al. 2022; Yu et al. 2022; Zhang et al. 2022).

These findings underscore the need to reconsider how depression is conceptualized, operationalized, and measured. The lack of clarity in measuring depression hinders the diagnosis and treatment of this mental disdorder (Fried, Flake, and Robinaugh 2022). Indeed, Fried (2017) analyzed the content of seven widely used depression scales, revealed a low overlap between scale in the symptoms. In a separate study, Fried and Nesse (2015) investigated the number of unique symptom profiles reported by 3703 depressed outpatients and found that nearly half of the symptom were endorsed by only one individual, further highlighting the variability in diagnostic criteria in clinincal settings. Similarly, Veal et al (2024) found that among 450 randomized controlled trials (RCTs) of unipolar and bipolar depression and identified 3888 different measures of outcomes. All these findings suggest it’s urgent to address the issue of measuring depression.

However, all previous studies focused on measuring depression among adults, few attention has been paid to the measurement of depression among children and adolescent. Previous studies have shown that depression in childred and adolescents are different from adults (). To fill the gap, we conduct a comprehensive content analysis of 27 Chinese depression scales to assess their heterogeneity in measuring depression. Our findings reveal a low overlapping among these scales. Additionally, our results provided a comprehensive list of symptoms used for screening depression. These results also offer insights into the potential psychopathology and social pathways of depression in Chinese society (and, by extension, East Asian societies) and will inspire future development of scales for measuring depression in clinical settings and public health policies.

# 2. Methods

We took three steps to extract symptoms from all scales that measure depressions among the Chinese student’s population (see Fig 1 for the flowchart). Firstly, we identified all scales that has been used for screening depression. Secondly, we identified unique symptoms of each scale. Thirdly, we compared the symptoms across all scales. The latter two steps followed Fried (2017) with minor modifications (see details below).

## 2.1 Identify and screen scales

We identified scales that measure depression from four recent meta-analyses which synthesized the prevalence of different mental health problems among four Chinese students populations: elementary school, middle school, high school, and college (于晓琪等, 2022; 黄潇潇等, 2022; 张亚利等, 2022; 陈雨濛等, 2022). We extracted all papers included in these meta-analyses and selected scales that were used for screening depression. In total, we identified 34 scales from 465 articles from all articles included in these four meta-analyses.

We then screened versions of scales and identified the most valid version for later analysis (See supplementary material for details). If a scale has multiple versions, we choose one of these version based on the following criteria: (1) If the reference(s) of the scale cited in articles was in English, we search for the Chinese versions because all participants in the above mentioned articles are Chinese students; (2) If the reference(s) of the scale included both English and Chinese versions of the scale, we included only the Chinese version; (3) If there were multiple Chinese versions and if the one(s) published later in time methologically improved the previous ones, we choose the later version; (4) If there were multiple Chinese versions and no obvious methodological advances were reported, we choose the one explicitly stated symptom names; (5) If all other condition were equal, we selected the most frequently cited one. For instance, CES-D was first translated by汪向东等(1999), which was used by 39 of all 465 papers included in the four menta-analyses. However, 章婕等(2010) improved the translation of the 20th item, thus, we chose the revised version by 章婕等(2010).

## 2.2 Identify symptoms within scales

In this phase, four trained coders focused on identifying items that assessed identical or similar symptoms within each scale. The procedure was as follows: First, the four coders independently identified and consolidated items within each scale. Then, they formed two pairs, with each pair reviewing their results and resolving any discrepancies within the pair. Subsequently, the results from two pairs of coders were cross-checked and any remaining inconsistencies were discussed and resolve with the corresponding author when necessary. The final consolidated version of the results underwent independent verification by a clinically certified psychiatrist (Y. L).

## 2.3 Compare symptoms across different scales

In this step, the same four coders compared symptoms across all scales. The procedure was the same as the identification of symptoms within scales: indepenedent individual coding, discussion by pairs, cross-checking between pairs, discussion with the corresponding author, and verification by a clinically certificated doctor.

The coding processes were as follow. We first pooled symptoms from all scales together and identified unique symptoms. In this step, we retained both compound symptoms and specific symptoms, as in Fried (2017). Compound symptoms are symptoms that include a range of related symptoms, whereas specific symptoms are more concrete and describe specific patterns. For example, “appetite changes” is a compound symptom, it includes two specific symptoms: “appetite increased” and “appetite decreased”, and all three of them were treated as unique symptoms. We employed an approach that maximize the amount of different symptoms. More specifically, if the items describe similar symptom using different words and that the words have significantly different meanings under the Chinese context, we treat them as belonging to the same compound symptom but are different specific symptoms. For instance, there are many different words for describe depressed mood in different scales, we used 'depressed moods' as the compound symptom but distinguish different specific symptoms such as: 'blue', 'low mood', 'sad', and 'anhedonia'. This approach is slightly different from Fried (2017), where he coded all these item as a specific symptom 'Sad moods'.

Then, we assigned score of all scales on all unique symptoms. More specifically, a scale was scored as “0” on a symptom if it does not have items that measures this symptom. For instance, the CDI has no item for 'Depressed mood', we assigned “0” for CDI on this symptom. If a scale has an item that directly measures a symptom, compound or specific, it was coded as 2 on that symptom. Note if a scale has an item measures a compound symptom, then, this scale not only had a score of 2 on that compound symptom but also has a score of 1 on all specific symptoms of this compound symptom. For example, CDI has an item directly measures the compound symptom “appetite change” and scored 2 on this compound symptom. Importantly, even CDI does not have items for 'appetite increased' and 'appetite decreased', it scored 1 on these two specific symptoms (see Fig 2 and Supplementary Materials for details). However, if the item measures a specific symptom under a compound symptom, this scale was coded “2” on that specific symptom but still “0” on the corresponding compound symptom.

[Insert Fig 2 later]

**Figure 2***. Number of articles that used each scale for screening depression.* Total number of articles were based on those included in four meta-analyses (XXX). Scales with \* were excluded from our analyses.

## 2.4 Data analyses

We reported the descriptive summary of scales as well as the symptoms within scale. We highlighted symptoms that are used in DSM-5 for diagonosis of depression. More specifically, there are 28 symptoms are overlapped with the DSM-5 symptoms for depression, where were derived from the nine symptoms in DSM-5 and all their specific symptoms, see Fried & Nesse, (2015) for details.

We used Jaccard index to quantify the degree of overlap between different scales (Fried, 2017). The formal of Jaccard index is s/(u1 + u2 + s), where "s" represents the number of items shared by two scales, and "u1" and "u2" denote the number of items that are exclusively present in each of the two scales. Jaccard index ranges from 0 (no overlap among scales) to 1 (complete overlap). We interpreted Jaccard index as in Fried (2017): very weak 0.00–0.19, weak 0.20–0.39, moderate 0.40–0.59, strong 0.60–0.79, and very strong 0.80–1.0. Moreover, we explored the relationship between the mean Jaccard coefficient and the scale length across 27 scales by employing Spearman correlation.

[Insert Fig later]

**Figure 1**

*Research flowchart*

# 3 Result

# 3.1 A summary of scales

Among all papers included in these four meta-analyses (citations), 465 reported depression. We identified 34 unique scales in these articles. Among all these scales, the items of four scales, the Mini International Neuropsychiatric Interview for children and adolescents (Mini-KID), WHO-CIDI 3.0, Psychological Health Inventory (PHI), and the Symptom Checklist 45, were not findable. The items of the other two scales, the Beck Depression Inventory (Zhang Yuxin Revised Edition) and Short Depression Scale, were not available either because of unidentifiable citations. These six scales were excluded from furthere analyses. The items used in 'Gu & Chen (2020) 'and 'Ji (2007)' were identifical but in different languages, thus we regarded these two studies used the same scale referred them as 'Ji (2007)'. Also, the boys’ and girls’ version of the Child Behavior Checklist (CBCL) were treated as one scale. In total, 27 scales were included in the current study. See Fig 2 for the frequency of citations of these scales among all 465 empirical papers in the meta-analyses (CITATIONS). The seven most frequently used scales in this study are SDS, SCL-90, CES-D, CDI, DSRSC, BDI-I, and MSSMHS, among which, the CES-D and SDS are consistent with the scales selected by Fried (2017).

# 3.2 Items and symptoms within scales

For the 27 scales included, there are 425 items in total. Among them, 73 items were merged into 31 symptoms. Also, there were two cases where one item measured two symptoms. The item from Ji (2007), "During the past 12 months, did you ever feel so sad or hopeless almost every day for two weeks or more in a row that you stopped doing your usual activities?" meausres both 'sad' and 'Sense of hopelessness'. And the 8th item from PHQ-9, "Actions or speech slowed down to a noticeable extent, or conversely—feeling restless or agitated, being unable to sit still, more than usual", measures both 'Agitation' and 'Retardation'. Thus, a total of 385 symptoms were identified across all 27 scales (See supplementary materials for number of items and symptoms of each included scale).

The comparison of 385 symptoms across 27 scales resulted in unique 84 symptoms (see Figure 3). Among these, eight are compound symptoms, including *Depressive mood, Irritability*, *Self-abasement*, *Interest/pleasure loss*, *Somatization*, *Appetite changes*, *Somnipathy*, and *Reduced socialization*. Among all 27 scales, 19 of them did not include any idiosyncratic symptoms. For the other eight scales the rate of idiosyncratic symptom varied from 3.9% to 22.2%. Interestingly, all scales include symptoms that were not covered in DSM-5. The DSI has the highest proportion of DSM-5 symptoms for depression, 71.42% of the total nine DSM-5 depression symptoms. Please see the supplementary matierals for detailed information.

[Insert Fig later]

**Figure 3***. Content Overlap Across* *Twenty-seven Depression Scales.* Each row represents a symptom, each column represents a scale.If a scale measures a symptom, then there is a dot or a circle on that row. The former represent compound symptoms and the latter represent specific symptoms. Symptoms in bold font are from DSM-5. ADI: Adolescent Depression Inventory, CDI: Children's Depression Inventory, HAMD: Hamilton Depression Rating Scale for Depression, DSI: Depression Status Inventory, SDS: Self- Rating Depression Scale, MFQ-C: Mood and Feelings Questionnaire, CBCL: Child Behavior Checklist, BDI-II：Beck Depression Inventory-II, DSRSC: Depression Self-rating Scale for Children, BDI-I: Beck Depression Inventory, KADS-11: Kutcher Adolescent Depression Scale, CES-D: The Center for Epidemiological Studies Depression Scale, PHQ-9: Patient Health Questionnaire-9 items, CSSDS: China Middle school students' depression scale, CES-D-C: Center for Epidemiologic Studies Depression Scale for Children, UPI: University Personality Inventory, SMFQ: Short Mood and Feelings Questionnaire, SCL-90: Symptom Checklist 90, CES-D-13: Short version of Center for Epidemiologic Studies Depression Scale, CCSMHS: Chinese College Student Mental Health Scale, DASS-21: The Depression Anxiety Stress Scale, BSRS-5: Brief Symptom Rating Scale, Sakuma\_2010: Sakuma et al.(2010) self-designed questionnaire, MSSMHS: Middle-school students Mental Health Scale, CEPS: China Education Panel Survey, HADS: Hospital Anxiety and Depression Scale, Ji\_2007: Comprehensive Survey Report on Health-Related/Risk Behaviors among Chinese Adolescents.The right side of the figure is *Number of symptoms that appear across combinations of scales.* There are 18 symptoms that appear only in one scale; these symptoms are referred to as idiosyncratic symptoms.

# 3.3 Symptoms across scales

Among the 84 symptoms, 18 (21.43%) were idiosyncratic symptoms and only appeared in one scale. None of the 84 symptoms were present in all scales. The most frequently presented symptom, appeared in 22 out of 27 scales, was *Sense of hopelessness*. The second most measured symptom, 18 out of 27, was *Interest loss*. Note that *markedly diminished interest or pleasure*, a key symptom of diagnosis of major depression in DSM-5, is splitted into two specific symptoms: *interest loss* and *pleasure loss* in this study*.* We found *Pleasure loss* was observed less frequently than *Interest loss*, being measured in 9 out of 27 scales. Another frequently measured symptom is the compound symptom *Depressed mood*, which was directly measured in 5 scales. However, this compound symptom includes several specific symptoms: *blue* appeared in 10 scales, *low mood* in 15 scales, *sad* in 13 scales, and *anhedonia* in 16 scales*.* Combined as a cluster, depressed mood and related specific symptom presented in 25 out of 27 scale and are the most frequently measured symptom.

The degree of overlap between scales was calculated using the Jaccard coefficient. The mean overlap across all scales was 0.19, range from 0.09 to 0.25, indicating a very low level of similarity between these scales (see Figure 5). CES-D has the highest mean degree of overlap with other scale. The highest overlap, 0.75, appeared between two versions of CES-D: CES-D for adults and CES-D-C for chidlren. The second highest overlap, 0.72, was between DSI and SDS. Many scales that have no overlap with each other. For example, there was no overlap between China Education Panel Survey with Middle-school students Mental Health Scale, PHQ-9, CSSDS, SMFQ, CSSMHS. Note that because Ji\_2005 has only one item that measures two symptoms, it has no overlap with PHQ-9, KADS-11.

We found a correlation between the mean Jaccard coefficient of each scale and the scale length, *r* = 0.54, 95% CI [0.14, 0.78]. Similarly, the correlation between mean overlap of scales and the number of captured symptoms is *r* = 0.70, 95% CI [0.39, 0.87]. These findings suggests that longer scales exhibit increased overlap with other scales, thus demonstrating enhanced representativeness.

[Insert Fig later]

**Figure 5.** *Overlap of item content of 27 depression scales.* ADI: Adolescent Depression Inventory, CDI: Children's Depression Inventory, HAMD: Hamilton Depression Rating Scale for Depression, DSI: Depression Status Inventory, SDS: Self- Rating Depression Scale, MFQ-C: Mood and Feelings Questionnaire, CBCL: Child Behavior Checklist, BDI-II：Beck Depression Inventory-II, DSRSC: Depression Self-rating Scale for Children, BDI-I: Beck Depression Inventory, KADS-11: Kutcher Adolescent Depression Scale, CES-D: The Center for Epidemiological Studies Depression Scale, PHQ-9: Patient Health Questionnaire-9 items, CSSDS: China Middle school students' depression scale, CES-D-C: Center for Epidemiologic Studies Depression Scale for Children, UPI: University Personality Inventory, SMFQ: Short Mood and Feelings Questionnaire, SCL-90: Symptom Checklist 90, CES-D-13: Short version of Center for Epidemiologic Studies Depression Scale, CCSMHS: Chinese College Student Mental Health Scale, DASS-21: The Depression Anxiety Stress Scale, BSRS-5: Brief Symptom Rating Scale, Sakuma\_2010: Sakuma et al.(2010) self-designed questionnaire, MSSMHS: Middle-school students Mental Health Scale, CEPS: China Education Panel Survey, HADS: Hospital Anxiety and Depression Scale, Ji\_2007: Comprehensive Survey Report on Health-Related/Risk Behaviors among Chinese Adolescents. Mean overlap is detailed in the supplementary materials.

# 4 讨论

[第一段: 总结结果，并强调本研究的特色（基于中国实际研究的情况）]

We identified 27 scales used in actual Chinese research for measuring depression from the most recent four meta-analyses, and derived 84 different symptoms from them. The mean overlap among all scales is 0.19, with 21.43% of the symptoms appearing only in a single scale, and no symptom presenting in all the scales. Our results are similar to those of Fried (2017), who evaluated the overlap among 52 depression symptoms across seven common depression scales. Fried (2017) demonstrated that 40% of symptoms appeared exclusively in one scale, while 12% were common across all instruments. We highlight that in the study of depression among Chinese students, the use of depression scales as interchangeable measurements is problematic. This could potentially jeopardize the universality and reproducibility of depression disorder research.

[第二段: 介绍具体量表的结果，并与已有研究进行比较]

Our results suggest that CES-D has the highest mean overlap with other scales(0.25), while in Fried (2017) research, CES-D has the lowest mean overlap with other scales(0.27). This might be because our study also included CES-D-C and CES-D-13. Moreover, in our study, as the scale of inclusion increased, CES-D had no idiosyncratic symptoms, but in Fried (2017), CES-D displayed 33% of idiosyncratic symptoms. In our study, the Ji\_2005 scale has the lowest mean overlap (0.09) with other scales, which is not surprising as this scale only includes one item.

较长与重叠率为正相关。

尽管先前研究发现问卷短比较好，但是过短的问卷可能损失了很多信息。

Dms-5的结果，有趣的是基于dsm-5编造的phq-9量表，只有xx。这是因为中文版的phq-9量表相对dsm-5失去了很多信息。例如第七条dsm-5是xx

则包括了feelings of worthlessness or inappropriate guilt;这两个症状，中文版对应的题目是6·觉得自己很糟，或很失败，或让自己或家人失望则没有内疚的意味了，因此没测到内疚这个症状。

[第三段: 分析异质性的原因，以及重新阐述问题的严重性]

抑郁量表的异质性来源可能是由于1构念的不清晰，2编造目的不一样，3文化背景的影响。例如在学习对抑郁的重要，因此中文抑郁量表adi和xx特别强调学习。

因此，在合并来自不同量表的研究结果时（例如元分析），特别是那些几乎没有重叠的量表时，可能不明智。

目前仍普遍认为量表可以互相替代，研究结果表明量表之间异质性非常强。因此在推广结果时，特别是大规模调查的结果，应该强调这是某个量表的结果。 这样可以帮助我们理解为什么大规模调查的抑郁检出率不一样。

[第四段：如何解读结果，call for better standards]

不能认为低重叠就是糟糕的量表，高重叠就是好的量表。它只意味着使用CES-D识别的结果更不可能推广到其他量表。

但这并不意味着它是一个特别好的量表，决定选择哪个量表时还需要考虑心理度量特性、治疗环境和文化考虑、管理和格式、性能特性（例如特异性和敏感性）以及评估的目的和目标（例如干预计划，筛查，结果评估）等多个因素。

选择测量工具时应考虑多种因素，并根据评估的目的和目标进行选择。有关哪些量表是适当的量表的问题，需要对工具在特定领域的表现和偏差进行实证比较。

本研究建议使用多个量表。

[第五段: 展望未来]

可以强调本土编制的症状。例如本土编制的才测到跟学习有关的内容。

**参 考 文 献**

Amaltinga, Awuni Prosper Mandela, and James Fenibe Mbinta. 2020. ‘Factors Associated with Depression among Young People Globally: A Narrative Review’. *International Journal Of Community Medicine And Public Health* 7(9):3711. doi: 10.18203/2394-6040.ijcmph20203949.

Chen Yumeng, Zhang Yali, and Yu Guoliang. 2022. ‘Prevalence of mental health problems among college students in mainland China from 2010 to 2020: A meta-analysis’. *Advances in Psychological Science* 30(5):991–1004. doi: 10.3724/SP.J.1042.2022.00991.

Dattani, Saloni. 2022. ‘At What Age Do People Experience Depression for the First Time?’ *Our World in Data*.

Deng, Hu, Fang Wen, Hui Xu, Hanxue Yang, Junjuan Yan, Yi Zheng, Yonghua Cui, and Ying Li. 2023. ‘Prevalence of Affective Disorders in Chinese School-Attending Children and Adolescents Aged 6–16 Based on a National Survey by MINI-Kid’. *Journal of Affective Disorders* 331:192–99. doi: 10.1016/j.jad.2023.03.060.

Eaton, Nicholas R., Laura F. Bringmann, Timon Elmer, Eiko I. Fried, Miriam K. Forbes, Ashley L. Greene, Robert F. Krueger, Roman Kotov, Patrick D. McGorry, Cristina Mei, and Monika A. Waszczuk. 2023. ‘A Review of Approaches and Models in Psychopathology Conceptualization Research’. *Nature Reviews Psychology* 2(10):622–36. doi: 10.1038/s44159-023-00218-4.

Fried, Eiko I. 2017. ‘The 52 Symptoms of Major Depression: Lack of Content Overlap among Seven Common Depression Scales’. *Journal of Affective Disorders* 208:191–97. doi: 10.1016/j.jad.2016.10.019.

Fried, Eiko I., Sacha Epskamp, Randolph M. Nesse, Francis Tuerlinckx, and Denny Borsboom. 2016. ‘What Are “good” Depression Symptoms? Comparing the Centrality of DSM and Non-DSM Symptoms of Depression in a Network Analysis’. *Journal of Affective Disorders* 189:314–20. doi: 10.1016/j.jad.2015.09.005.

Fried, Eiko I., Jessica K. Flake, and Donald J. Robinaugh. 2022. ‘Revisiting the Theoretical and Methodological Foundations of Depression Measurement’. *Nature Reviews Psychology* 1(6):358–68. doi: 10.1038/s44159-022-00050-2.

Fried, Eiko I., and Randolph M. Nesse. 2015. ‘Depression Sum-Scores Don’t Add up: Why Analyzing Specific Depression Symptoms Is Essential’. *BMC Medicine* 13(1):72. doi: 10.1186/s12916-015-0325-4.

Fu, Xiaolan, and kan Zhang. 2023. *心理健康蓝皮书 中国国民心理健康发展报告(2021-2022)*. 1st ed. 北京: 社会科学文献出版社.

Herrman, Helen, Vikram Patel, Christian Kieling, Michael Berk, Claudia Buchweitz, Pim Cuijpers, Toshiaki A. Furukawa, Ronald C. Kessler, Brandon A. Kohrt, Mario Maj, Patrick McGorry, Charles F. Reynolds, Myrna M. Weissman, Dixon Chibanda, Christopher Dowrick, Louise M. Howard, Christina W. Hoven, Martin Knapp, Helen S. Mayberg, Brenda W. J. H. Penninx, Shuiyuan Xiao, Madhukar Trivedi, Rudolf Uher, Lakshmi Vijayakumar, and Miranda Wolpert. 2022. ‘Time for United Action on Depression: A Lancet–World Psychiatric Association Commission’. *The Lancet* 399(10328):957–1022. doi: 10.1016/S0140-6736(21)02141-3.

Huang Xiaoxiao, Zhang Yali, and Yu Guoliang. 2022. ‘Prevalence of mental health problems among primary school students in Chinese mainland from 2010 to 2010:A meta-analysis’. *Advances in Psychological Science* 30(5):953–64. doi: 10.3724/SP.J.1042.2022.00953.

Li, Fenghua, Yonghua Cui, Ying Li, Lanting Guo, Xiaoyan Ke, Jing Liu, Xuerong Luo, Yi Zheng, and James F. Leckman. 2022. ‘Prevalence of Mental Disorders in School Children and Adolescents in China: Diagnostic Data from Detailed Clinical Assessments of 17,524 Individuals’. *Journal of Child Psychology and Psychiatry* 63(1):34–46. doi: 10.1111/jcpp.13445.

McGrath, John J., Ali Al-Hamzawi, Jordi Alonso, Yasmin Altwaijri, Laura H Andrade, Evelyn J. Bromet, Ronny Bruffaerts, José Miguel Caldas De Almeida, Stephanie Chardoul, Wai Tat Chiu, Louisa Degenhardt, Olga V. Demler, Finola Ferry, Oye Gureje, Josep Maria Haro, Elie G. Karam, Georges Karam, Salma M. Khaled, Viviane Kovess-Masfety, Marta Magno, Maria Elena Medina-Mora, Jacek Moskalewicz, Fernando Navarro-Mateu, Daisuke Nishi, Oleguer Plana-Ripoll, José Posada-Villa, Charlene Rapsey, Nancy A. Sampson, Juan Carlos Stagnaro, Dan J. Stein, Margreet Ten Have, Yolanda Torres, Cristian Vladescu, Peter W. Woodruff, Zahari Zarkov, Ronald C. Kessler, Sergio Aguilar-Gaxiola, Ali Al-Hamzawi, Jordi Alonso, Yasmin A. Altwaijri, Laura Helena Andrade, Lukoye Atwoli, Corina Benjet, Evelyn J. Bromet, Ronny Bruffaerts, Brendan Bunting, José Miguel Caldas-de-Almeida, Graça Cardoso, Stephanie Chardoul, Alfredo H. Cía, Louisa Degenhardt, Giovanni De Girolamo, Oye Gureje, Josep Maria Haro, Meredith G. Harris, Hristo Hinkov, Chi-yi Hu, Peter De Jonge, Aimee N. Karam, Elie G. Karam, Georges Karam, Alan E. Kazdin, Norito Kawakami, Ronald C. Kessler, Andrzej Kiejna, Viviane Kovess-Masfety, John J. McGrath, Maria Elena Medina-Mora, Jacek Moskalewicz, Fernando Navarro-Mateu, Daisuke Nishi, Marina Piazza, José Posada-Villa, Kate M. Scott, Juan Carlos Stagnaro, Dan J. Stein, Margreet Ten Have, Yolanda Torres, Maria Carmen Viana, Daniel V. Vigo, Cristian Vladescu, David R. Williams, Peter Woodruff, Bogdan Wojtyniak, Miguel Xavier, and Alan M. Zaslavsky. 2023. ‘Age of Onset and Cumulative Risk of Mental Disorders: A Cross-National Analysis of Population Surveys from 29 Countries’. *The Lancet Psychiatry* 10(9):668–81. doi: 10.1016/S2215-0366(23)00193-1.

Solmi, Marco, Joaquim Radua, Miriam Olivola, Enrico Croce, Livia Soardo, Gonzalo Salazar De Pablo, Jae Il Shin, James B. Kirkbride, Peter Jones, Jae Han Kim, Jong Yeob Kim, Andrè F. Carvalho, Mary V. Seeman, Christoph U. Correll, and Paolo Fusar-Poli. 2022. ‘Age at Onset of Mental Disorders Worldwide: Large-Scale Meta-Analysis of 192 Epidemiological Studies’. *Molecular Psychiatry* 27(1):281–95. doi: 10.1038/s41380-021-01161-7.

UNICEF. n.d. ‘ADOLESCENT MENTAL HEALTH’.

William Styron. 1990. *Darkness Visible: A Memoir of Madness.* New York: Random House.

Yu, Xiaoqi, Yali Zhang, and Guoliang Yu. 2022. ‘Prevalence of Mental Health Problems among Senior High School Students in Mainland of China from 2010 to 2020: A Meta-Analysis’. *Advances in Psychological Science* 30(5):978. doi: 10.3724/SP.J.1042.2022.00978.

Zeynep Başgöze, Andrea Wiglesworth, Katherine A. Carosella, Bonnie Klimes-Dougan, and Kathryn R. Cullen. 2021. ‘Depression, Non-Suicidal Self-Injury, and Suicidality in Adolescents: Common and Distinct Precursors, Correlates, and Outcomes’. *Journal of Psychiatry and Brain Science*. doi: 10.20900/jpbs.20210018.

Zhang Yali, Jin Juanjuan, and Yu Guoliang. 2022. ‘Prevalence of mental health problems among junior high school students in Chinese mainland from 2010 to 2020: A meta-analysis’. *Advances in Psychological Science* 30(5):965–77. doi: 10.3724/SP.J.1042.2022.00965.

汪向东, 王希林, and 马弘. 1999. 心理卫生评定量表手册. 中国心理卫生杂志社.

章婕, 吴振云, 方格, 李娟, 韩布新, and 陈祉妍. 2010. ‘流调中心抑郁量表全国城市常模的建立’. 中国心理卫生杂志 24(2):139–43.