# Psychiatry and Psychological Disorders

*Research Article*

# Psychometric Properties of the Child and Adolescent PsychProfiler v5: Teacher-Report Form

**Stephen Houghton[1]\*, Rapson Gomez[2], Shane Langsford[3]**

[1]Graduate School of Education, University of Western Australia, Perth, WA- 6009 Australia
[2]Institute of Health and Wellbeing, Federation University, Ballarat, VIC- 3353, Australia
[3]Psychological & Educational Consultancy Services, Perth, WA- 6008 Australia

**\*Corresponding Author:** Stephen Houghton, Graduate School of Education, University of Western Australia, Perth, WA- 6009, Australia.

## Abstract

The Child and Adolescent PsychProfiler version 5 (CAPP v5) is a measure for screening 14 common DSM-5-TR disorders in children and adolescents. The separation of Attention-Deficit/Hyperactivity Disorder (ADHD) and Specific Learning Disorder (SLD) by subtype results in 17 screening scales covering the 14 disorders. With separate self-report (SRF), parent-report (PRF), and teacher-report (TRF) versions, theoretically all three CAPP versions should have a 17-factor structure. Although recent studies with the CAPP-PRF and CAPP-SRF have confirmed this structure, the CAPP-TRF has not been formally tested. In the current study two groups of 6- to 11-year-old children had the CAPP-TRF completed by their teacher: Group 1 (N =1345, mean age 8.13 years, SD 1.61 years) and Group 2 (222 children, mean age 8.90 years, SD 2.19 years). Group 2 children also completed the self-report Beck Youth Inventories, while their teachers completed the Conners 3 T-L. Together, the results across the two samples supported a 13-factor model, and almost all factors in this model showed acceptable reliability (alpha and omega coefficients) and discriminant and criterion validity. These findings indicate acceptable psychometric properties for a slightly revised version of the CAPP-TRF and somewhat complements the parent-report and self-report PsychProfiler forms as an effective teacher-report screener for common DSM-5-TR disorders in children.

**Keywords:** Child and Adolescent PsychProfiler (CAPP), CAPP-TRF, Factor structure, Psychometric properties, Teacher-report ratings, Screening, DSM-5-TR.

## Introduction

According to the Australian Institute of Health and Welfare [1], in 2013–14 almost 14% of children experienced a mental disorder, with Attention-Deficit/Hyperactivity Disorder (ADHD) being the most common (8.2%), followed by anxiety disorders (6.9%). Similarly, The National Survey of Mental Disorders in Australian Children and Adolescents [2] also found that almost one in seven (13.9%) of Australian 4–17-year-olds had a mental health disorder in the previous 12 months (equivalent to 560,000 school aged students).

Prior to the COVID-19 pandemic up to 20% of young people globally reported mental health problems [3] and a meta-analysis by [4], indicated the global prevalence of clinically elevated anxiety (18%) and depression (23.8%) amongst N = 89,879 youth was significantly higher than the estimates provided pre-pandemic.

Contemporary evidence suggests the mental health of young people continues to decline [5], with the 2025 second Lancet Commission on child and adolescent health and

*J. Psychiatry. Psychiatr. Disord. Vol. 5 Iss. 1 (039)*

*Page-01*

wellbeing, commenting "globally the health and wellbeing of adolescents (ages 10–24 years) is at a tipping point" [6]. It is unsurprising therefore, that worldwide there is an increasing emphasis on mental health within the context of schools [7]. Given the Second Lancet Commission [6] highlighted that mental health from age 10 to 19 years "has pronounced, lasting effects on a range of functional outcomes in adulthood" (p. 34), obtaining an early and accurate assessment and diagnoses of mental disorders in children is critical. It is widely recommended that the assessment and diagnoses of mental disorders in children involve information derived from multiple informants, especially parents, teachers and the child [8,9]. In this context, there is a paucity of clinically reliable and valid teacher-report instruments that cover a broad range of childhood disorders, aligned with major clinical classification systems, such as the current Diagnostic and Statistical Manual of Mental Disorders, Fifth Edition, Text Revision [DSM-5-TR, the [10]. One potentially useful teacher-report measure that aligns with many common DSM-5-TR childhood mental disorders is the Child and Adolescent PsychProfiler-Teacher Report Form (CAPP-TRF; [11-13]. However, to date there is limited information on the psychometric properties for this teacher-report version of this instrument. As such, the goal of the present study was to examine the factor structure, reliability, and multiple forms of validity for the CAPP-TRF.

### The PsychProfiler [11-13].

The CAPP-TRF version 5 is part of the PsychProfiler [13] suite of instruments, a comprehensive global screening instrument oriented to the DSM-5-TR and developed for the simultaneous investigation of the 20 most common disorders (see www.psychprofiler.com and Supplementary Table S1). The DSM-5-TR [10] diagnostic criteria of all the disorders included within the PsychProfiler version 5 remain identical to its predecessor the DSM-5 [14], and therefore, our reference to the DSM-5-TR in this paper is also entirely applicable to the earlier DSM-5.

There are two parallel versions of the PsychProfiler, the Child and Adolescent PsychProfiler (CAPP) and the Adult PsychProfiler (APP), which between them cover children, adolescents, and adults.

The PsychProfiler Manual [13] presents the theoretical basis for the CAPP (pp.11–15), and this is in part reproduced in Supplementary Table S1. There are three versions within the CAPP, namely the self-report form (SRF), parent-report form (PRF), and teacher-report form (TRF). All three versions screen for 14 of the most common DSM-5-TR [10] disorders. The disorders are Attention-Deficit/Hyperactivity Disorder (ADHD); Oppositional Defiant Disorder (ODD); Conduct Disorder (CD); Specific Learning Disorder (SLD); Autism Spectrum Disorder (ASD); Language Disorder (LD); Speech Sound Disorder (SSD); Generalized Anxiety Disorder (GAD); Persistent Depressive Disorder (PDD); Separation Anxiety

Disorder (SAD); Obsessive Compulsive Disorder (OCD); Posttraumatic Stress Disorder (PTSD); Anorexia Nervosa (AN); and Bulimia Nervosa (BN). For ADHD, there are separate scales for ADHD: Predominantly Inattentive Presentation (ADHDI), ADHD: Predominantly Hyperactive–Impulsive Presentation (ADHDHI); and ADHD: Combined Presentation (ADHDC); and for SLD, there are separate scales for SLD with impairment in Reading (SLD-R), SLD with impairment in Written Expression (SLD-W), and SLD with impairment in Mathematics (SLD-M). Thus, there are 17 screening scales in total across the 14 disorders. For all three CAPP versions (i.e., CAPP-SRF, CAPP-TRF, and CAPP-PRF), the items comprising the different screening scales correspond directly to the DSM-5-TR (APA, 2022) disorder symptoms with the same name. Thus, a 17-factor structure is theoretically feasible for all forms of the CAPP, with the factors being the 17 screening scales corresponding to their DSM-5-TR equivalents.

Each of the three versions of the CAPP have 126 items that are rated on a 6-point Likert-type scale (never = 0, rarely = 1, sometimes = 2, regularly = 3, often = 4, and very often = 5). Three of the items are not used for screening purposes and are for rater-reliability purposes and therefore not assigned to any factor scale. Therefore, 123 items for clinical screening are used within each version of the CAPP. For all versions of the CAPP, the screening scales and the DSM-5-TR [10] disorders with the same names are highly congruent, thereby indicating strong face validity for the different screening scales in all three versions of the CAPP.

### Existing Psychometric Properties of the PsychProfiler

As stated in the PsychProfiler Manual, Second Edition, the initial CAPP "was subjected to a rigorous psychometric analysis and found to be reliable and valid" [13]. Although information supporting inter-rater reliability, clinical calibration, the use of a six-point scale for the ratings of the items, and suitable readability for use by children and adolescents can be found in the PsychProfiler Manual, to date there is only limited published information on the psychometric properties of the CAPP self-report (CAPP-SRF) and parent-report (CAPP-PR) versions [15,16].

Additionally, for the CAPP-PRF, for adolescents, [16] found support for the 17-factor model, with almost all factors showing acceptable reliability (alpha and omega coefficients) and acceptable discriminant validity. Support was also reported for criterion, concurrent, and discriminant validity of the scales in the CAPP-PRF. For the CAPP-SRF, similar findings were reported by [15] involving adolescents. These findings indicate support for the functionality of the CAPP-PRF and CAPP-SRF for screening common childhood and adolescent DSM-5 disorders, based on parent and self-ratings respectively.

In contrast to the CAPP-PRF and CAPP-SRF, there is currently limited psychometric information on the CAPP-TRF. A

preliminary confirmatory factor analytic investigation by [17] did find the individual screening scales in the CAPP-TRF were unidimensional. At present, we have very little of the psychometric properties commonly needed for sound use of a clinical measure, such as factor structure, various from of validity, and internal consistency.

Although the teacher-report (CAPP-TRF), self-report (CAPP-SRF), and parent-report (CAPP-PRF) versions are almost identical, it cannot be assumed that the CAPP-TRF will also show comparable acceptable psychometric properties to that found for the CAPP-PRF see and CAPP-SRF [16,15]. It is therefore prudent that these be independently demonstrated. Providing confirmation of their factor structure (e.g., the presumed 17-factor structure) and other psychometric properties such as reliability and validity would be helpful for their continuing development and clinical utility. Consequently, this present study set out to further investigate the 17-factor structural model for the CAPP-SRF, and relatedly, its validity (discriminant, criterion, and concurrent) and reliability (alpha and omega coefficients).

## Framework for Evaluating the Psychometric Properties of a New Clinical Measure

The American Educational Research Association, American Psychological Association, and National Association for Measurement in Education [18] Standards for Educational and Psychological Testing have proposed guidelines for the development and evaluation of tests and testing practices. They also include guidelines for assessing the validity of interpretations of test scores for the intended test uses. With regards to test validation, this document does not focus on distinct types of validity but rather focuses on different aspects of validity. As [19] summarized, these include: Evidence based on test content (i.e., the relationship of the item themes, wording and format with the intended construct, including administration process); response processes (the cognitive processes and interpretation of items by respondents and users, as measured against the intended construct); internal structure (the extent to which item interrelationships conform to the intended construct); relations to other variables (the pattern of relationships of test scores to external variables as predicted by the intended construct); and consequences of testing (intended and unintended consequences, as can be traced to a source of invalidity such as construct under-representation or construct-irrelevant variance) as necessary for interpreting and using test scores.

We evaluated the psychometric properties of the CAPP-TRF using the framework set out in these Standards. As can be deduced, these guidelines relate to various forms of test validity. Face validity can be inferred if the test measures what it claims to on the surface (e.g., an ADHD test looks like an ADHD test), and internal validity can be inferred if the proposed theoretical measurement model holds. Construct validity can be assumed if the test measures the theoretical

concept (construct) it is supposed to, like depression or anxiety. Content validity can be assumed if the test covers all relevant aspects of the area being measured (e.g., a test for DSM-R-TR symptoms for ADHD covers all the DSM-R-TR symptoms for ADHD). Criterion validity (or more specifically, concurrent validity) can be indicated if the test correlates with a current measure of the same thing (e.g., the ADHD scale matching an established one). For this, convergent validity can be indicated if a test shows a strong correlation, and discriminant (or divergent) validity can be indicated if there is no or low significant correlation.

For the CAPP-TRF, the evidence base for test content and response processes (i.e., use of a six-point scale for rating items, readability levels for children and adolescents, and clinical calibration) have already been established see [13]. Given that the items for the different disorder scales align with the corresponding disorders in DSM-5-TR, it can be assumed that it has face validity, construct validity, and content validity. In contrast, there is limited evidence for internal validity, criterion validity, including convergent validity and discriminant/divergent validity. Accordingly, these need to be established, and it is this that forms the primary objective of this paper.

## Aims of the Study

The aim of the current study was to use CFA to examine support for the 17-factor structure of the CAPP-TRF for children. Although it is generally recommended that exploratory factor analysis (EFA) or principal component analysis (PCA) be conducted initially to examine the expected causal connections between variables (i.e., factor structure of a measure), it is not mandatory. As noted by researchers [20,21], when there is strong theory underlying the measurement model or when the relationships of the factors and related items are known, the application of CFA prior to EFA/PFA is justified. Given that the CAPP-TR was developed to align fully with the DSM-5-TR disorders, it can be considered to have a well-developed underlying model, and a clear expectation of the patterns of factor loadings (i.e., the relevant items for each of the separate disorders loading on its specific disorders) - a requirement to conduct a CFA. In addition, the factor structure has support from previous studies involving parallel versions of the CAPP-TRF (i.e., CAPP-PRF, [16] and CAPP-SRF [15]. Taking these points into consideration, we decided to progress straight to the examination of the factor structure of the CAPP-TRF using CFA. However, we also decided that if the CFA failed to establish at least an acceptable model, EFA would then be used to explore potential reasons for the misfit, such as cross-loadings or problematic items. The factors in the CFA model were ADHD-HI, ADHD-I, CD, ODD, SLD-R, SLD-W, SDL-M, ASD, LD, SSD, GAD, PDD, SAD, PTSD, OCD, AN, and BN, with each scale corresponding to the symptoms of the DSM-5-TR disorder with the same name. However, given the complexity of the model tested

(123 items, with each item being rated on a six-point interval, and these items loading on 17 factors), we were apprehensive about the ability of the model to show admissible solution. Related to this, we decided that if this was not found, the output of the original 17-factor model would be re-examined and the model revised accordingly. Additionally, the factors in the model had to show clarity (salience and significance of the loadings on their designated factors), reliability (alpha and omega coefficients), and discriminant and criterion validity. However, as the data set used for the CFA did not include variables that could be used for examining criterion validity, we used another data set which included variables suitable for this purpose. For ease of communication, we refer to this sample as Sample 2, and the sample used in the CFA as Sample 1. Together, therefore, the study examined the factor structure (internal validity) of the CAPP-TRF, and the reliability (alpha and omega coefficients), and discriminant and criterion validities of the factors in the CAPP-TRF model. Based primarily on the recently completed study involving the CAPP-PRF [16] and CAPP-SRF [15], we expected be support for the 17-factor model, with the factors in the model showing clarity (items loading significantly and saliently on their respective designated factors), internal consistency, and validity (discriminant and criterion).

### Supplementary Material

The PsychProfiler v5 (Langsford, Houghton, & Douglas, 2022) is a comprehensive screening instrument for the simultaneous investigation of 20 of the most common disorders found in children, adolescents, and adults. The Child and Adolescent PsychProfiler (CAPP) allows for the simultaneous screening of the 14 most prevalent disorders in children and adolescents. The DSM-5-TR disorders in the CAPP are listed below in their respective DSM-5-TR disorder category.

| DSM-5-TR disorder category | DSM-5-TR disorder |
| --- | --- |
| Anxiety Disorders | Generalised Anxiety Disorder; Separation Anxiety Disorder |
| Attention-Deficit/Hyperactivity Disorder | Attention-Deficit/Hyperactivity Disorder (Predominantly Inattentive Presentation, Predominantly Hyperactive/Impulsive Presentation, Combined Presentation) |
| Autism Spectrum Disorder | Autism Spectrum Disorder |
| Communication Disorders | Language Disorder; Speech Sound Disorder |
| Depressive Disorders | Persistent Depressive Disorder |
| Disruptive, Impulse-Control, & Conduct Disorders | Conduct Disorder; Oppositional Defiant Disorder |
| Feeding and Eating Disorders | Anorexia Nervosa; Bulimia Nervosa |
| Obsessive-Compulsive and Related Disorders | Obsessive-Compulsive Disorder |
| Specific Learning Disorder | Specific Learning Disorder (with impairment in Reading, Written Expression, Mathematics) |
| Trauma and Stressor-Related Disorders | Posttraumatic Stress Disorder |

**Supplementary Table S1:** Brief Description of the PsychProfiler

# Method

### Participants

Sample 1 comprised children (N = 1345) whose teachers had completed either an online or paper version of the CAPP-TRF [16] on their behalf. The mean age (SD, range) of children rated was 8.13 years (1.61 years; 6 years to 11 years). For those with information on gender (N = 1336), there were 924 (69.2%) boys, (mean age =8.09 years, SD = 1.62 years) and 412 (30.8%) girls (mean age = 8.64 years, SD = 1.59 years). There was no significant difference for age across boys and girls, t (df = 1334) = 1.537, ns. Supplementary Table S2 shows the frequencies of those at-risk for the 17 disorders in the CAPP-TRF in Sample 1. (The method used to establish at-risk for the different disorders is described in the measures section.) As shown in Supplementary Table S2, participants were more at-risk for ADHD-I and the three SLDs (SLD-R, SLD-W, and SLD-M) than the other disorders. The risk for AN, BN, and OCD were extremely low at 2 (0.15%), 3 (0.22%) and 13 (0.97%), respectively. Considering this, we excluded these disorders in the CFA of the CAPP-TRF. This meant that the initial CAPP-TRF model tested was a 14-factor model, and not the originally intended 17-factor model.

Sample 2 comprised clinic-referred children (N = 222) whose teachers completed a paper version of the CAPP-TRF [13], as part of the overall assessment requested at clinic they were attending. The mean age (SD, range) of children rated was 8.90 years (SD = 2.10 years, range = 6 years to 11 years), of which 139 (62.6%) were boys, (mean age = 8.82

*J. Psychiatry. Psychiatr. Disord. Vol. 5 Iss. 1 (039)*

*Page-04*

years, SD = 2.30 years) and 83 (37.3%) were girls (mean age = 8.91 years, SD = 2.31 years). There was no significant difference for age across these groups, t (df = 220) = .282, ns.

| Risk for Disorder | Screened | | |
| --- | --- | --- | --- |
| | Negative | Positive | Rank Order |
| GAD | 1195 | 150 | 12 |
| SAD | 1284 | 61 | 14 |
| ADHDHI | 977 | 368 | 6 |
| ADHDI | 683 | 662 | 2 |
| ASD | 1180 | 165 | 11 |
| LD | 884 | 461 | 5 |
| SSD | 1108 | 237 | 10 |
| PDD | 1054 | 291 | 9 |
| CD | 999 | 346 | 7 |
| ODD | 1011 | 334 | 8 |
| AN | 1343 | 2 | 17 |
| BN | 1338 | 7 | 16 |
| OCD | 1332 | 13 | 15 |
| SLD-R | 763 | 582 | 3 |
| SLD-W | 553 | 792 | 1 |
| SLD-M | 768 | 577 | 4 |
| PTSD | 1242 | 103 | 13 |

**Note**. GAD/PDD = Generalised Anxiety Disorder/Persistent Depressive Disorder; SAD = Separation Anxiety Disorder, ADHDHI = Attention-Deficit/Hyperactivity Disorder: Hyperactive/Impulsive; ADHDI = Attention Deficit/Hyperactivity Disorder: Inattentive; ASD = Autism Spectrum Disorder; LD = Language Disorder; SSD =Speech Sound Disorder; CD = Conduct Disorder; ODD = Oppositional Defiant Disorder; SLD-R = Specific Learning Disorder: Reading; SLD-W = Specific Learning Disorder: Written Expression; SLD-M = Specific Learning Disorder: Mathematics; PTSD = Posttraumatic Stress Disorder.

*Supplementary Table S2*: Risk for Disorders in Sample 1

**Measures**
For both Samples 1 and 2, teachers completed the CAPP-TRF [16]. Additionally for Sample 2, children completed self-ratings for the Beck Youth Inventories, Second Edition BYI-2; [22], and their teachers completed the Conners 3–Long Teacher Report [23]. Some children in Sample 2 were also administered the Wechsler Individual Achievement Test, Third Edition WIAT-III; [24].

**Teacher-Report Form Child and Adolescent PsychProfiler CAPP-TRF; [13]**
Further to the information provided in the introduction about the CAPP-TRF, all of its items are rated on a six-point Likert scale (never = 0, rarely = 2, sometimes =2, regularly = 3, often = 4, and very often = 5). For calculating whether an item endorses the symptom, the item scores were recoded as follows: never, rarely, sometimes = 0 (symptom not present); and regularly, often, and very often = 1 (symptom present).

For the CFA that examined support for the factor structure of the CAPP-TRF, we used these item rating scores.
For establishing if a child was at-risk for a disorder, the total recoded scores for all the items with the scales for the disorders is computed. When the total recoded score for a scale exceeds the screening cut-off score, the disorder that it corresponds to is considered present. Currently, the cut-off scores for each screening scale are identical with the symptom threshold cut-off scores for them in the DSM-5-TR (e.g., if the CAPP ODD scale has a score of 4 or more for ODD, then ODD is considered to have a positive screen, as the symptom threshold cut-off score for ODD in the DSM-5-TR is 4). Given the close alignment of the items in the CAPP-TRF screening scales with the symptoms in the corresponding DSM-5-TR disorders, this practice seems intuitively prudent. For the study that examined the criterion validity of the CAPP-TRF scales, the disorders were established using these screening cut-off scores. Since they were not formally

diagnosed, we refer to children as being "at-risk" for the disorders.

## The Beck Youth Inventories, Second Edition BYI-2;[22]

The Beck Youth Inventories, Second Edition BYI-2; [22] is for children and adolescents 7 to 18 years of age. It comprises five self-report inventories, measuring depression (Beck's Depression Inventory for Youths; BDI-Y), anxiety (Beck's Anxiety Inventory for Youths; BAI-Y), anger (Beck's Anger Inventory for Youths; BANI-Y), disruptive behaviour (Beck's Disruptive Behaviour Inventory for Youths; BDBI-Y), and self-concept (Beck's Self-Concept Inventory for Youths; BCSI-Y).

Each inventory has 20 items, resulting in 100 items in total. Individuals rate all 100 items in terms of the extent to which each statement describes them on a 4-point Likert scale (i.e., "0 = never," "1 = sometimes," "2 = often," "3 = very often"). For each inventory, item scores are summed and converted to T-Scores, with higher scores indicating greater severity. All five inventories have good construct validity, high reliability (coefficient alpha ranging from .86 to .96), and high test-retest reliability (coefficients ranging from .74 to .93 [22], thereby supporting their psychometric properties (factor structure, reliability, and validity), and use. The scores for all five inventories have shown adequate convergent validity with scales measuring related constructs [22], and ability to distinguish between clinical and nonclinical samples [25].

For the present study, the ratings of only the BAI-Y, BDI-Y and the BDBI-Y were used. The BDI-Y includes items covering negative thoughts, feelings of guilt and sadness, and sleep issues; the BAI-Y includes items covering concerns and apprehension regarding school, the future, reactions from others, losing control, and physiological anxiety symptoms. Thus, the scores for the BDI-Y can be considered suitable for evaluation of the criterion validity of CAPP-TRF PDD, and the BAI-Y can be considered suitable for evaluation of the criterion validity of anxiety-related disorders of the CAPP-TRF screening scales (i.e., GAD, SAD, OCD, and PTSD). The BDBI-Y includes items relating to behaviours and attitude associated with ODD and CD (e.g., "I hurt people"). Thus, it can be considered suitable for evaluation of the criterion validity of CAPP-TRF disruptive behavior disorders (CD and ODD). The scales measured by BANI-Y and BCSI-Y (anger and self-concept) subscales were not used as they could not be directly linked to the disorders measured in the CAPP-TRF.

## The Conners 3rd Edition Teacher-Report Long Form [23]

The Conners 3-T-L is a teacher-report screening measure used in the diagnosis of ADHD and disorders commonly co-morbid with ADHD in children and adolescents aged 6 to 18 years. While it has scores for a number of correlates associated with ADHD (e.g., learning problems/executive functioning), it also includes four DSM-5-TR Symptom Scales. These are ADHD Inattentive (C-ADHD-I), ADHD Hyperactive-Impulsive (C-ADHD-HI), Conduct Disorder (C-CD), and Oppositional Defiant Disorder [23]. The scores for these scales are highly correlated with the DSM-5-TR disorders with the same name [23], meaning they are appropriate for evaluating the criterion validity of the CAPP-TRF screening scales for ADHD-I, ADHD-HI, CD, and ODD as they correspond appropriately with DSM-5-TR symptoms.

## Wechsler Individual Achievement Test, Third Edition: Australian and New Zealand Standardised Edition WIAT-III; [24]

The WIAT-III is an individually administered instrument designed to measure the academic achievement of students who are in kindergarten through to the final year of secondary school (i.e. Year 12), or ages 4 years 0 months to 50 years 11 months. It consists of a total of 16 subtests grouped into listening, speaking, reading, writing, spelling, and mathematical skills. The Australian and New Zealand version of the test was standardized on a sample of 1360 Australian and New Zealand students and features comprehensive normative information. A SLD for reading, written expression, and mathematics can be diagnosed if an individual scores 1.0 standard deviations below the mean (which equates to a standard score of 85/16th percentile) on any of the subtests. We used this cut-off (i.e., below 16th percentile) for the reading, written expression, and mathematics subtests to identify those potentially with a SLD with impairment in reading, written expression, and mathematics, respectively. Thus, they can be considered suitable for evaluation of the criterion validity of CAPP-TRF SLD-R, SLD-W and SLD-M, respectively.

## Procedure

The PsychProfiler measures (including the CAPP-TRF) have a designated website (www.psychprofiler.com) that can be used by those interested in the online screening of DSM-5 disorders using any of the PsychProfiler forms. The primary users are psychologists, psychiatrists, paediatricians, and the general public. The participants involved in Study 1 provided data through the website. On completion of the Psych-Profiler, individuals and parents were requested, if they so wished, to check a statement consenting to their data being used for future research and instrument validation purposes. Only adolescents with CAPP-TRF ratings and consent from the teacher were included in Sample 1.

Sample 2 included children seen in a clinic setting in Perth, Western Australia that they were attending to complete an ADHD, SLD, and/or ASD assessment. These children were referred to the clinic from a variety of sources (e.g., privately by their parents, through their school, or from a general practitioner, paediatrician, or child and adolescent psychiatrist). The teachers of the children in this sample completed the CAPP-TRF and Conners 3-T-L [22], as part of the as-

sessment that their children were undergoing, and the children also completed the self-report BYI-2 [22]. The children were also administered the WIAT-III [24] to collect academic achievement results. Parents were provided with a Consent Form to sign should they consent to their child's de-identified data being used for future research and instrument validation purposes. Only information of children whose parents had signed the Consent Form were included in Sample 2. Thus, informed consent was obtained from the of all child participants involved in the study.

## Statistical Analysis

As previously mentioned, there were two different samples involved in the current study. Sample 1 comprised 1345 ratings of children provided by their teachers for the CAPP-TRF. The intended primary goal of the analysis for this sample was to use CFA to examine the fit of the 17-factor CAPP model. Initially, we used the descriptive module in Jeffreys' Amazing Statistics Program [26] version 0.16.6.0 statistical software to compute the mean and standard deviation scores, and the dispersion statistics of the 123 items of the CAPP-TRF. [27] has suggested that data can be considered to have normal univariate distribution if skewness is between -3 to +3 and kurtosis is between -10 to +10. [28] assert nonnormality can be seen as problematic if ≥ 80% of responses are at one end of the scale.

Mplus Version 7 [29] was used to analyse the 17-factor CAPP-TRF model. As each item was dichotomous, we applied WLSMV extraction as there was nonnormality problems in the data set (details presented below). WLSMV is a robust estimator, recommended for CFA with ordered-categorical scores, as is the case in the current study. This method does not assume normally distributed variables. According to [30-32], relative to other estimators, the WLSMV estimator provides the best option for modeling categorical data, including binary scored items. Despite this, when the model has a large number of items and latent factors (as is the case in the current study with 123 items and 14 latent factors), it is not uncommon for researchers to use normal theory estimators, such as those based on maximum likelihood (ML), to conduct CFA even when ordinal data are present (e.g., [33]. This is especially valuable when it is desired to include all the missing values in the data set using full information ML estimation, and when the major focus of interest is not on the metric level of the analyses, as is the case in this present study. Therefore, given nonnormality in the data set, we examine model fit using both WLSMV and robust ML (MLR). At the statistical level, the global fit of this model was examined using the chi-square test, i.e., WLSMV$\chi$2 when the WLSMV extraction was applied, and MLR when MLR extraction was applied. However, as the chi-square statistic is inflated by large sample sizes, several approximate fit indices have been proposed. Among others, commonly used fit indices have included the relative chi square (rela-

tive $\chi$2) (sometime called normed $\chi$2) which is a ratio of the chi-square statistic to the respective degrees of freedom ($\chi$2/df), Root Mean Squared Error of Approximation (RMSEA), Comparative Fit Index (CFI), Tucker Lewis Index (TLI), and the Standardized Root Mean Residual (SRMR) for MLR extraction and Weighted Root Mean Square Residual (WRMR) for WLSMV extraction.

[34] suggested that for model fit, chi-square, RMSEA, CFI, and SRMR/WRMR be examined in combination. Although RMSEA, CFI and TLI are frequently used to evaluate model fit, we did not use these in the current study because as these values are derived from the chi-square value, they would also be compromised [35]. In contrast, SRMR/WRMR is not derived from the chi-square value [36]. Taking all of this into consideration, the current study used relative $\chi$2 and SRMR/WRMR in combination to evaluate model fit. For relative $\chi$2, acceptable values have ranged from less than 2 [37] to less than 5 [38] to be deemed acceptable. [39] have proposed that for SRMR, values ≤ .08 = acceptable. For the WRMR, we used ≤1 for establishing acceptable fit [40]. Despite not relying on the CFI, TLI and RMSEA values for evaluating model fit, we report these values for those interested. For these indices, [39] have proposed that for RMSEA, values < 0.06 = good fit, <0.08 = acceptable fit, and > 0.08 to .10 = marginal fit. For CFI and TLI, values ≥.95 = good fit, and ≥ .90 = acceptable fit.

Additionally, for model acceptance in this study, it was necessary for the loadings of the indicators in the model to be significant and salient (>.30; [41], and for the factors to demonstrate acceptable discriminant validity (r <. .85; [27], and acceptable reliability omega coefficients [42]. Although there are no universally accepted guidelines at present for interpreting omega coefficients, [43] proposed that omega values should meet the same standards as alpha coefficients. For alpha coefficients, guidelines for acceptability ranged from .70 and above [34]. For the current study we used omega values of at least .70 as acceptable. Also, for those interested, we report on the alpha reliability coefficients of the CAPP-TRF factors. Finally, given model complexity, we were apprehensive about the ability of the 17-factor CAPP-TRF model to show admissible solution. Related to this, we decided that if this occurred, the model would be revised to achieve a model that had acceptable solution.

Sample 2 comprised 222 clinic-referred children with CARR-TRF ratings. For this sample, the data used for the disorders were the total recoded scores for the 17 scales in the CAPP-TRF [13]. The primary analysis for this sample was to examine the criterion validity of the relevant CAPP-TRF factors. For this, we used SPSS. We examined Pearson's correlations of the relevant CAPP-TRF factors with the total scores in the BDI-Y, BAI-Y, and BDBI-Y BYI-2; [22] DSM-5 symptom scales for C-ADHD-I, C-ADHD-HI, C-CD, and C-ODD of the Conners 3 T-L. Correlations were also computed for the factors in the CAPP-TRF scales with the diagnosis of specific

learning disorders for reading, written expression, and mathematics, based on the WIAT-III scores. For all correlations, effects sizes were examined using [44] guidelines (small = .10, medium = .30, and large = .50).

### Data availability statement

The data supporting this study's findings are available from the corresponding author upon reasonable request.

## Results

### Sample Size Requirements

[45] software for computing sample size requirements for the CFA models was used to evaluate the sample size required for the study involving Sample 1. The anticipated effect size was set at 0.3, power at 0.8, the number of latent variables at 17, the number of observed variables at 123, and probability at 0.05. The analysis recommended a minimum sample size of 218 to detect effect, and a minimum sample size of 462 for model structure. Therefore, with N = 1345, our sample size was more than adequate for the CFA in the current study.

### General Comments Relating to the Tables in this Study

Considering the CFA involved 123 items and 17 factors, the summaries of our tables are lengthy. In view of this, all tables for the study are presented as supplementary materials.

### Descriptive and Dispersion Statistics of the CAPP-TRF Items

Supplementary Table S3 shows the mean and standard deviation scores, missing values, and the dispersion statistics of the 123 items of the CAPP-TRF. The overall mean (SD) score for the 123 items was 1.651 (0.682). As all items were rated on a six-point Likert scales (never = 0, rarely = 1, sometimes = 2, regularly = 3, often = 4, and very often = 5), the overall mean score suggests that in general, individuals were endorsing either rarely or sometimes as their responses. Overall, therefore, the participants in the study had relatively low pathology. In terms of missing values for the items, they ranged from 6 (for item #67) to 810 (for item #51).

The skewness scores ranged from -0.683 to 9.409, with 18 items having values outside -3 to +3. For kurtosis, the scores ranged from -1.599 to -93.85, with 17 items having values outside -10 to +10. As suggested, data can be considered to have normal univariate distribution if skewness is between -3 to +3 and kurtosis is between -10 to +10 [27], and that nonnormality can be considered problematic if ≥ 80% of responses are at one end of the scale [28]. Considering this, our skewness and kurtosis findings can be interpreted as reflecting normality for most, but not all, items. Consequently, we used robust estimations in the CFA, i.e., WLSMV when the WLSMV extraction was applied, and the MLR when the MLR extraction was applied. It is worth noting that with the MLSMV, missing values are deleted listwise. Thus, the analysis does not include all participants in the study. In contrast, when MLR is applied, full-information maximum likelihood FIML; e.g., [46] can be applied to handle missing values in the CFA. Thus, the analysis involving MLR included all participants in the study

| Items | Valid | Missing | Mean | Standard Deviation | Skewness | Kurtosis |
|---|---|---|---|---|---|---|
| S112 | 781 | 9 | 2.73 | 1.69 | -0.07 | -1.22 |
| S1 | 784 | 6 | 2.72 | 1.48 | 0.1 | -1.08 |
| S94 | 778 | 12 | 2.02 | 1.58 | 0.39 | -0.93 |
| S54 | 779 | 11 | 1.68 | 1.52 | 0.7 | -0.5 |
| S15 | 786 | 4 | 1.88 | 1.47 | 0.41 | -0.78 |
| S24 | 782 | 8 | 2.67 | 1.76 | -0.05 | -1.33 |
| S118 | 781 | 9 | 2.77 | 1.55 | 0.01 | -1.08 |
| S69 | 782 | 8 | 2.31 | 1.6 | 0.25 | -1.07 |
| S64 | 783 | 7 | 0.82 | 1.19 | 1.67 | 2.38 |
| S23 | 781 | 9 | 1.63 | 1.6 | 0.78 | -0.44 |
| S111 | 781 | 9 | 0.89 | 1.24 | 1.61 | 2.27 |
| S70 | 783 | 7 | 0.82 | 1.34 | 1.79 | 2.41 |
| S36 | 785 | 5 | 1.13 | 1.35 | 1.26 | 0.97 |
| S43 | 785 | 5 | 0.93 | 1.33 | 1.55 | 1.66 |
| S51 | 785 | 5 | 0.76 | 1.27 | 1.83 | 2.65 |
| S18 | 786 | 4 | 0.92 | 1.31 | 1.53 | 1.62 |
| S46 | 783 | 7 | 2.61 | 1.69 | -0.03 | -1.22 |
| S79 | 784 | 6 | 2.14 | 1.68 | 0.34 | -1.09 |

| | | | | | | |
|---|---|---|---|---|---|---|
| S30 | 787 | 3 | 1.31 | 1.47 | 1.07 | 0.22 |
| S74 | 785 | 5 | 1.59 | 1.43 | 0.77 | -0.19 |
| S72 | 784 | 6 | 2.04 | 1.47 | 0.42 | -0.71 |
| S40 | 786 | 4 | 1.83 | 1.66 | 0.57 | -0.88 |
| S13 | 784 | 6 | 1.49 | 1.4 | 0.82 | -0.06 |
| S87 | 783 | 7 | 1.6 | 1.4 | 0.86 | 0.11 |
| S50 | 783 | 7 | 1.62 | 1.37 | 0.82 | 0.07 |
| S57 | 785 | 5 | 2.96 | 1.49 | -0.16 | -0.96 |
| S81 | 782 | 8 | 2.04 | 1.35 | 0.54 | -0.28 |
| S104 | 782 | 8 | 2.35 | 1.45 | 0.25 | -0.79 |
| S75 | 783 | 7 | 2.43 | 1.43 | 0.23 | -0.8 |
| S100 | 785 | 5 | 2.54 | 1.58 | 0.05 | -1.11 |
| S3 | 779 | 11 | 2.88 | 1.53 | -0.16 | -1.07 |
| S42 | 785 | 5 | 2.19 | 1.49 | 0.39 | -0.8 |
| S97 | 781 | 9 | 3.04 | 1.53 | -0.3 | -0.95 |
| S6 | 782 | 8 | 2.81 | 1.46 | -0.02 | -0.95 |
| S35 | 778 | 12 | 1.74 | 1.41 | 0.49 | -0.6 |
| S5 | 781 | 9 | 1.61 | 1.54 | 0.76 | -0.46 |
| S116 | 782 | 8 | 1.26 | 1.37 | 1.03 | 0.26 |
| S71 | 784 | 6 | 0.97 | 1.13 | 1.44 | 2.08 |
| S8 | 782 | 8 | 2.2 | 1.68 | 0.34 | -1.08 |
| S123 | 781 | 9 | 1.76 | 1.47 | 0.61 | -0.46 |
| S59 | 781 | 9 | 1.14 | 1.3 | 1.22 | 0.87 |
| S124 | 775 | 15 | 1.06 | 1.29 | 1.32 | 1.18 |
| S126 | 782 | 8 | 1.17 | 1.45 | 1.19 | 0.49 |
| S41 | 781 | 9 | 1.82 | 1.51 | 0.6 | -0.55 |
| S82 | 783 | 7 | 1.08 | 1.45 | 1.32 | 0.75 |
| S62 | 775 | 15 | 1.9 | 1.61 | 0.49 | -0.86 |
| S11 | 781 | 9 | 1.58 | 1.46 | 0.72 | -0.43 |
| S115 | 780 | 10 | 1.71 | 1.43 | 0.66 | -0.33 |
| S114 | 782 | 8 | 2.13 | 1.45 | 0.37 | -0.69 |
| S32 | 781 | 9 | 2.22 | 1.46 | 0.36 | -0.78 |
| S78 | 785 | 5 | 1.36 | 1.31 | 0.99 | 0.56 |
| S109 | 784 | 6 | 1.47 | 1.26 | 0.89 | 0.46 |
| S125 | 783 | 7 | 2.24 | 1.4 | 0.33 | -0.62 |
| S84 | 784 | 6 | 1.72 | 1.37 | 0.64 | -0.23 |
| S68 | 780 | 10 | 1.13 | 1.32 | 1.18 | 0.71 |
| S48 | 759 | 31 | 0.99 | 1.2 | 1.29 | 1.21 |
| S122 | 783 | 7 | 1.36 | 1.34 | 0.99 | 0.34 |
| S83 | 780 | 10 | 0.93 | 1.23 | 1.45 | 1.66 |
| S91 | 782 | 8 | 2.01 | 1.59 | 0.43 | -0.9 |
| S107 | 780 | 10 | 2.29 | 1.55 | 0.27 | -0.98 |

| | | | | | | |
|---|---|---|---|---|---|---|
| S28 | 783 | 7 | 2.7 | 1.62 | -0.03 | -1.17 |
| S38 | 785 | 5 | 2.64 | 1.61 | 0.05 | -1.17 |
| S27 | 784 | 6 | 2.12 | 1.54 | 0.37 | -0.89 |
| S19 | 778 | 12 | 2.15 | 1.67 | 0.25 | -1.14 |
| S37 | 783 | 7 | 2.98 | 1.54 | -0.22 | -1.02 |
| S99 | 780 | 10 | 2.68 | 1.71 | -0.07 | -1.28 |
| S63 | 783 | 7 | 0.75 | 1.33 | 1.97 | 3.04 |
| S44 | 784 | 6 | 0.41 | 0.93 | 2.73 | 7.84 |
| S53 | 783 | 7 | 1.57 | 1.35 | 0.82 | 0.01 |
| S4 | 783 | 7 | 0.78 | 1.08 | 1.69 | 2.92 |
| S90 | 785 | 5 | 0.42 | 0.83 | 2.51 | 7.38 |
| S17 | 783 | 7 | 0.6 | 1.09 | 2.2 | 4.73 |
| S73 | 786 | 4 | 0.38 | 0.79 | 2.72 | 9.11 |
| S16 | 788 | 2 | 0.42 | 0.85 | 2.7 | 8.57 |
| S77 | 785 | 5 | 0.13 | 0.45 | 4.75 | 30.77 |
| S89 | 782 | 8 | 0.17 | 0.55 | 4.78 | 30.08 |
| S31 | 786 | 4 | 0.08 | 0.34 | 5.3 | 35.54 |
| S95 | 785 | 5 | 0.37 | 0.81 | 2.89 | 9.74 |
| S113 | 784 | 6 | 2.07 | 1.51 | 0.52 | -0.64 |
| S102 | 784 | 6 | 1.92 | 1.43 | 0.58 | -0.43 |
| S26 | 785 | 5 | 1.63 | 1.23 | 0.81 | 0.44 |
| S80 | 783 | 7 | 1.02 | 1.15 | 1.32 | 1.72 |
| S110 | 785 | 5 | 1.14 | 1.17 | 1.23 | 1.55 |
| S92 | 781 | 9 | 2.4 | 1.53 | 0.21 | -0.99 |
| S20 | 781 | 9 | 1.89 | 1.43 | 0.57 | -0.48 |
| S14 | 784 | 6 | 1.86 | 1.47 | 0.53 | -0.6 |
| S86 | 783 | 7 | 0.89 | 1.29 | 1.58 | 1.86 |
| S105 | 782 | 8 | 1.59 | 1.79 | 0.77 | -0.83 |
| S67 | 780 | 10 | 0.85 | 1.42 | 1.68 | 1.74 |
| S61 | 780 | 10 | 1.87 | 1.8 | 0.51 | -1.15 |
| S21 | 786 | 4 | 0.88 | 1.24 | 1.52 | 1.74 |
| S29 | 782 | 8 | 1.9 | 1.49 | 0.61 | -0.52 |
| S45 | 783 | 7 | 0.93 | 1.25 | 1.53 | 1.93 |
| S9 | 785 | 5 | 1.95 | 1.83 | 0.49 | -1.19 |
| S93 | 782 | 8 | 1.64 | 1.79 | 0.73 | -0.89 |
| S52 | 786 | 4 | 1.21 | 1.44 | 1.19 | 0.53 |
| S55 | 783 | 7 | 2.04 | 1.74 | 0.4 | -1.15 |
| S22 | 782 | 8 | 2.03 | 1.73 | 0.39 | -1.11 |
| S60 | 780 | 10 | 1.98 | 1.67 | 0.41 | -1.03 |
| S76 | 782 | 8 | 1.46 | 1.63 | 0.97 | -0.3 |
| S119 | 775 | 15 | 1.64 | 1.53 | 0.69 | -0.51 |
| S65 | 780 | 10 | 1.54 | 1.58 | 0.83 | -0.42 |

| S66 | 786 | 4 | 1.46 | 1.56 | 0.94 | -0.16 |
|-----|-----|---|------|------|------|-------|
| S88 | 784 | 6 | 1.75 | 1.49 | 0.64 | -0.44 |
| S25 | 787 | 3 | 1.98 | 1.71 | 0.42 | -1.11 |
| S121 | 779 | 11 | 0.78 | 1.4 | 1.84 | 2.29 |
| S34 | 786 | 4 | 2.49 | 1.62 | 0.1 | -1.09 |
| S106 | 779 | 11 | 1.92 | 1.61 | 0.51 | -0.82 |
| S58 | 784 | 6 | 2.06 | 1.63 | 0.52 | -0.9 |
| S96 | 783 | 7 | 2.18 | 1.71 | 0.36 | -1.13 |
| S108 | 780 | 10 | 1.05 | 1.51 | 1.36 | 0.69 |
| S47 | 783 | 7 | 2.41 | 1.77 | 0.18 | -1.27 |
| S103 | 780 | 10 | 1.92 | 1.65 | 0.54 | -0.85 |
| S12 | 784 | 6 | 0.87 | 1.21 | 1.56 | 2.07 |
| S56 | 785 | 5 | 1.97 | 1.73 | 0.49 | -1.04 |
| S33 | 780 | 10 | 1.62 | 1.74 | 0.7 | -0.88 |
| S2 | 780 | 10 | 1.96 | 1.67 | 0.46 | -1.05 |
| S98 | 781 | 9 | 1.57 | 1.64 | 0.78 | -0.59 |
| S7 | 782 | 8 | 1.34 | 1.49 | 0.99 | 0.02 |
| S85 | 780 | 10 | 1.58 | 1.62 | 0.76 | -0.58 |
| S10 | 777 | 13 | 2.08 | 1.69 | 0.33 | -1.12 |
| S120 | 778 | 12 | 1.36 | 1.59 | 0.97 | -0.25 |
| S49 | 782 | 8 | 2.08 | 1.55 | 0.38 | -0.89 |

**Supplementary Table S3:** Mean, Standard Deviation, and Dispersion Statistics of the 126 items of the PsychProfiler

**Fit indices of the 17-factor and Revised CAPP-TRF Models**

Using the WLSMV estimator, the fit values for the 14-factor CAPP-TRF model (N = 1345) were: WLSMV $X^2$ (df = 5473) = 35904.595; relative $X^2$ = 6.60; WRMR = 3.616; CFI = 0.866; TLI = 0.861; RMSEA = 0.064 CI [.064, .065]. Based on current guidelines, with the exception of RMSEA, all other fit values can be interpreted as indicating unacceptable model fit. According to [33], the RMSEA is inappropriate for evaluating model fit when the model has a large number of items and estimated with WLSMV extraction. When MLR extraction was used, the fit values for the 14-factor CAPP-TRF model (N = 1345) were: MLR $X^2$ (df = 5473) = 26794.425; relative $X^2$ = 4.89; SRMR = 0.092 CFI = 0.765; TLI = 0.757; RMSEA = .054, 90% CI [.053, .054]. Based on current guidelines, the relative $X^2$ and the SRMR values (used here for evaluating global model fit) can be interpreted as indicating acceptable model fit. Although not used here for evaluating model fit, it should be noted that the RMSEA value (0.054) also indicated good fit. Despite these positive findings, our output for both the WLSMV and MLR extraction models indicated that our model was inadmissible as the correlations between GAD and PDD was above 1, i.e., 1.046 and 1.151 for the MLR and WLSMV extractions, respectively. Considering this, this problem was resolved by combining the items in these two factors into a single entity that we called factor GAD/PDD. This is acceptable as there is robust evidence that for children and adolescents, general anxiety and depression items are highly correlated [47], and load together on a single factor, as is the case for the parent, teacher, and self-report questionnaires of the Achenbach System of Empirically Based Assessment (ASEBA; [48]. For this revised 13-factor model, the WLSMV extraction results indicated the following fit values: WLSMV$X^2$ (df = 5486) = 36199.973; relative $X^2$ = 6.60; WRMR = 3.637; CFI = 0.865; TLI = 0.860; RMSEA = 0.065 CI [.064, .065]. Thus, there was unacceptable fit when the WLSMV was applied. Additionally, the output indicated an inadmissible solution, and we were unable to resolve this issue with this model, using WLSMV extraction.

When MLR extraction was applied to the revised 13-factor model, there was an acceptable solution, and the fit values were: MLR $X^2$ (df = 5486) = 26920.375; relative $X^2$ = 4.90; SRMR = 0.093; CFI = 0.764; TLI = 0.756; RMSEA = .054, 90% CI [.053, .055]. Therefore, based on the relative $X^2$ and the SRMR values (used here for evaluating global model fit) as well as the RMSEA, the findings can be interpreted as indicating acceptable model fit for the this revised 13-factor model. Considering this, we opted to examine model fit using the results from the MLR analysis. This approach ensured the utilization of full information ML to account for missing

values. As mentioned previously, when the model has a large number of items and latent factors (as is the case in the current study with 123 items and 14/13 latent factors), it is not uncommon for researchers to use ML-based extraction methods to conduct CFA even when ordinal data are present (e.g., Cinamon, 2016; Lau et al., 2016).

### Pattern of factor loadings in the revised 13-factor PP model

Supplementary Table S4 shows the pattern of factor loadings in the revised 13-factor CAPP-TRF model. As shown in the table, all items loaded significantly and saliently (≥ .30) on their respective designated factors. Thus, the factors in the model were clearly defined.

### Correlations of the factors in the revised 13-factor CAPP-TRF model

The correlations of the factors in the revised 13-factor CAPP-TRF model are provided in Supplementary Table S5. As can be seen, the correlations for CD with LD, SSD, SLD-R, SLD-W and SLD-M; GAD with LD, SSD and SLD-W; ASD with SLD-R; ADHDHI with SLD-M; and PTSD with SLD-R were all nonsignificant. All other correlations were significant and positive. Of these, only SLD-R with SLD-W had a correlation > .85. These findings indicate that apart from SLD-R with SLD-W (known to commonly co-occur), there was good support for discriminant validity across the other latent factors.

As can be ascertained from Supplementary Table S5, for the disorders traditionally considered neurodevelopmental disorders screening scales (ASD, LD, SSD, SLD-R, SLD-W an SLD-M), all intercorrelations were of large effect sizes, and generally higher than their intercorrelations with other disorders. This was also the case for all the screening scales for the disorders traditionally considered internalizing disorders (GAD/PDD, SAD, and PTSD). For the intercorrelations involving the screening scales for the disorders traditionally considered externalizing disorders (ADHD-I, ADHD-HI, CD, and ODD), all intercorrelations involving ADHDHI, CD, and ODD were of large effect, and generally higher than their intercorrelations with other disorders. While ADHD-I correlated with large effect size with ADHD-HI, it connected with only moderate effects sizes with CD, and ODD. ADHD-I correlated with large effect sizes with ASD, LD and SLD-W, and medium effects sizes with SSD, SLD-R, and SRD-M. Thus, ADHD-I correlated relatively strongly with a number of neurodevelopmental disorders. Notwithstanding these groupings, there were several large effect size correlations across these groupings. These were ASD with ADHD-I; CD with GAD and ASD; SLD-W with ADHD-I; and PTSD with ASD and ODD.

### Internal consistency reliability coefficients for the 13 factors in CAPP-TRF model

Supplementary Table S4 includes the factor-based internal consistency reliability omega coefficients of the 13 latent factors in the revised CAPP-TRF model. As shown, they ranged from .858 to .950. The internal consistency reliability alpha coefficients ranged from .851 to .949. These values indicated good internal consistency reliability for all 13 CAPP-TRF factors.

### Criterion Validity of the 17 CAPP-TRF scales

Supplementary Table S7 shows the Pearson's correlation coefficients of the total scores of the 14 CAPP-TRF scales (PDD and GAD treated separately for these analysis) with the total scores for Conners 3 T-L DSM-5 Symptom Scales, BYI-2 scales, and the WIAT-III ability scores. For clarity we focus only on those associations that we earlier speculated could be used to establish the criterion validity of the different CAPP scales.

As shown in Supplementary Table S7, CAPP-TRF ADHDH-I correlated positively at close to large effect size with Conners 3 T-L HI, and this correlation was higher than other correlation pairs. This was also the case for CAPP-TRF ADHD-I with Conners 3 T-L IA, and for CAPP-TRF CD with Conners 3 T-L CD. The CAPP-TRF ODD scores correlated with small effect size with the Conners 3 T-L ODD. The CAPP-TRF ODD correlated with large effect size with Conners 3 T-L CD. Overall, these findings can be interpreted as indicating support for the criterion validity for the CAPP-TRS scales for AHDH-HI, ADHD-I, CD and ODD.

Supplementary Table S7 also shows that CAPP-TRF SLD-R correlated positively at close to large effect size with WIAT-III Reading, and this correlation was higher than other correlation pairs. The CAPP-TRF SLD-W correlated positively at moderate effect size with WIAT-III written scores. It also showed a similar magnitude of correlation with WIAT-III Reading scores. The CAPP-TRF scores for SLD-M correlated positively at close to small effect size with WIAT-III Maths. Despite the low correlation, this correlation was higher than other correlation pairs. Overall, therefore, these findings can be interpreted as indicating support for the criterion validity for the Reading, Written Expression, and Mathematics scales of the CAPP-TRF.

All of the CAPP-TRF anxiety-related disorders (GAD, SAD, and PTSD) were not associated with the BYI-2 anxiety scale; and CAPP-TRF PDD was not associated with the BYI-2 depression scale. ADHD-I and SSD were correlated positively with moderate effect sizes with the BYI-2 depression scale. Neither CAPP-TRF ODD or CAPP-TRF CD were associated with the BYI-2 disruptive behavior scale. CAPP-TRF SAD, ADHD, PDD were correlated positively with BYI-2 disruptive behavior scale. Thus, these findings did not indicate support for the criterion validity for the anxiety-related and depression disorder scales in in CAPP-TRF.

| | GAD/PDD | SAD | ADHDHI | ADHDI | ASD | LD | SSD | CD | ODD | SLD-R | SLD-W | SLD-M | PTSD |
|------|---------|-----|--------|-------|-----|-----|-----|-----|-----|-------|-------|-------|------|
| 36 | .57 | | | | | | | | | | | | |
| S1 | .64 | | | | | | | | | | | | |
| S94 | .58 | | | | | | | | | | | | |
| S54 | .83 | | | | | | | | | | | | |
| S15 | .60 | | | | | | | | | | | | |
| S24 | .53 | | | | | | | | | | | | |
| S118 | .32 | | | | | | | | | | | | |
| S69 | .50 | | | | | | | | | | | | |
| S91 | .76 | | | | | | | | | | | | |
| S107 | .81 | | | | | | | | | | | | |
| S28 | .59 | | | | | | | | | | | | |
| S38 | .63 | | | | | | | | | | | | |
| S27 | .80 | | | | | | | | | | | | |
| S19 | .51 | | | | | | | | | | | | |
| S37 | .42 | | | | | | | | | | | | |
| S99 | .61 | | | | | | | | | | | | |
| S64 | | .81 | | | | | | | | | | | |
| S23 | | .79 | | | | | | | | | | | |
| S111 | | .86 | | | | | | | | | | | |
| S70 | | .71 | | | | | | | | | | | |
| S36 | | .87 | | | | | | | | | | | |
| S43 | | .79 | | | | | | | | | | | |
| S51 | | .86 | | | | | | | | | | | |
| S18 | | .76 | | | | | | | | | | | |
| S46 | | | .78 | | | | | | | | | | |
| S79 | | | .86 | | | | | | | | | | |
| S30 | | | .79 | | | | | | | | | | |
| S74 | | | .86 | | | | | | | | | | |
| S72 | | | .71 | | | | | | | | | | |

| | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| S40 | | | .76 | | | | | | | | |
| S13 | | | .74 | | | | | | | | |
| S87 | | | .72 | | | | | | | | |
| S50 | | | .85 | | | | | | | | |
| S57 | | | | .87 | | | | | | | |
| S81 | | | | .64 | | | | | | | |
| S104 | | | | .82 | | | | | | | |
| S75 | | | | .77 | | | | | | | |
| S100 | | | | .83 | | | | | | | |
| S3 | | | | .73 | | | | | | | |
| S42 | | | | .72 | | | | | | | |
| S97 | | | | .83 | | | | | | | |
| S6 | | | | .62 | | | | | | | |
| S35 | | | | | .83 | | | | | | |
| S5 | | | | | .70 | | | | | | |
| S116 | | | | | .65 | | | | | | |
| S71 | | | | | .75 | | | | | | |
| S8 | | | | | .71 | | | | | | |
| S123 | | | | | .69 | | | | | | |
| S59 | | | | | .68 | | | | | | |
| S124 | | | | | .42 | | | | | | |
| S126 | | | | | .53 | | | | | | |
| S41 | | | | | .65 | | | | | | |
| S82 | | | | | .49 | | | | | | |
| S62 | | | | | .69 | | | | | | |
| S11 | | | | | | .78 | | | | | |
| S115 | | | | | | .81 | | | | | |
| S114 | | | | | | .77 | | | | | |
| S32 | | | | | | .70 | | | | | |
| S78 | | | | | | .78 | | | | | |

| | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| S109 | | | | | | .78 | | | | | | | |
| S125 | | | | | | .78 | | | | | | | |
| S84 | | | | | | .75 | | | | | | | |
| S68 | | | | | | | .88 | | | | | | |
| S48 | | | | | | | .80 | | | | | | |
| S122 | | | | | | | .87 | | | | | | |
| S83 | | | | | | | .75 | | | | | | |
| S63 | | | | | | | | .48 | | | | | |
| S44 | | | | | | | | .27 | | | | | |
| S53 | | | | | | | | .70 | | | | | |
| S4 | | | | | | | | .87 | | | | | |
| S90 | | | | | | | | .82 | | | | | |
| S17 | | | | | | | | .47 | | | | | |
| S73 | | | | | | | | .62 | | | | | |
| S16 | | | | | | | | .80 | | | | | |
| S77 | | | | | | | | .56 | | | | | |
| S89 | | | | | | | | .63 | | | | | |
| S31 | | | | | | | | .27 | | | | | |
| S95 | | | | | | | | .88 | | | | | |
| S113 | | | | | | | | | .90 | | | | |
| S102 | | | | | | | | | .83 | | | | |
| S26 | | | | | | | | | .72 | | | | |
| S80 | | | | | | | | | .77 | | | | |
| S110 | | | | | | | | | .80 | | | | |
| S92 | | | | | | | | | .84 | | | | |
| S20 | | | | | | | | | .89 | | | | |
| S14 | | | | | | | | | .91 | | | | |
| S66 | | | | | | | | | | .94 | | | |
| S88 | | | | | | | | | | .80 | | | |

| | GAD/PDD | SAD | ADHDHI | ADHDI | ASD | LD | SSD | CD | ODD | SLD-R | SLD-W | SLD-M | PTSD |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| S25 | | | | | | | | | | .91 | | | |
| S121 | | | | | | | | | | .84 | | | |
| S34 | | | | | | | | | | | .83 | | |
| S106 | | | | | | | | | | | .77 | | |
| S58 | | | | | | | | | | | .83 | | |
| S96 | | | | | | | | | | | .55 | | |
| S108 | | | | | | | | | | | .77 | | |
| S47 | | | | | | | | | | | | .94 | |
| S103 | | | | | | | | | | | | .93 | |
| S12 | | | | | | | | | | | | .76 | |
| S56 | | | | | | | | | | | | .94 | |
| S33 | | | | | | | | | | | | .83 | |
| S2 | | | | | | | | | | | | | .76 |
| S98 | | | | | | | | | | | | | .69 |
| S7 | | | | | | | | | | | | | .76 |
| S85 | | | | | | | | | | | | | .69 |
| S10 | | | | | | | | | | | | | .71 |
| S120 | | | | | | | | | | | | | .66 |
| S49 | | | | | | | | | | | | | .69 |
| Reliability | | | | | | | | | | | | | |
| Omega | .910 | .907 | .937 | .926 | .907 | .920 | .899 | .913 | .950 | .928 | .869 | .948 | .858 |
| Alpha | .908 | .902 | .936 | .925 | .901 | .920 | .897 | .888 | .949 | .923 | .866 | .945 | .851 |

**Note.** GAD/PDD = Generalised Anxiety Disorder/Persistent Depressive Disorder; SAD = Separation Anxiety Disorder, ADHDHI = Attention-Deficit/Hyperactivity Disorder: Hyperactive/Impulsive; ADHDI = Attention Deficit/Hyperactivity Disorder: Inattentive; ASD = Autism Spectrum Disorder; LD = Language Disorder; SSD =Speech Sound Disorder; CD = Conduct Disorder; ODD = Oppositional Defiant Disorder; SLD-R = Specific Learning Disorder: Reading; SLD-W = Specific Learning Disorder: Written Expression; SLD-M = Specific Learning Disorder: Mathematics; PTSD = Posttraumatic Stress Disorder.

*Supplementary Table S4:* Factor Loadings and Omega Coefficient of 13-Factor PsychProfiler Model

| | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| GAD/PDD (1) | 1.00 | .58 | .52 | .51 | .74 | .19 | .13 | .65 | .82 | .05 | .20 | .09 | .83 |
| SAD (2) | | 1.00 | .14 | .20 | .41 | .18 | .21 | .26 | .32 | .10 | .13 | .14 | .79 |
| ADHDHI (3)) | | | 1.00 | .67 | .50 | .13 | .15 | .59 | .66 | .03 | .20 | .02 | .26 |
| ADHDI (4) | | | | 1.00 | .55 | .57 | .32 | .34 | .39 | .34 | .58 | .37 | .26 |
| ASD (5) | | | | | 1.00 | .41 | .34 | .53 | .63 | .05 | .22 | .13 | .60 |
| LD (6) | | | | | | 1.00 | .75 | .00 | -.03 | .73 | .78 | .71 | .15 |
| SSD (7) | | | | | | | 1.00 | .02 | .00 | .60 | .57 | .50 | .13 |
| CD (8) | | | | | | | | 1.00 | .89 | -.04 | .05 | -.04 | .48 |
| ODD (9) | | | | | | | | | 1.00 | -.08 | .05 | -.09 | .58 |
| SLD-R (10) | | | | | | | | | | 1.00 | .86 | .71 | .02 |
| SLD-W (11) | | | | | | | | | | | 1.00 | .71 | .11 |
| SLD-M (12) | | | | | | | | | | | | 1.00 | .08 |
| PTSD (13) | | | | | | | | | | | | | 1.00 |

**Note.** GAD/PDD = Generalised Anxiety Disorder/Persistent Depressive Disorder; SAD = Separation Anxiety Disorder, ADHDHI = Attention-Deficit/Hyperactivity Disorder: Hyperactive/Impulsive; ADHDI = Attention Deficit/Hyperactivity Disorder: Inattentive; ASD = Autism Spectrum Disorder; LD = Language Disorder; SSD =Speech Sound Disorder; CD = Conduct Disorder; ODD = Oppositional Defiant Disorder; SLD-R = Specific Learning Disorder: Reading; SLD-W = Specific Learning Disorder: Written Expression; SLD-M = Specific Learning Disorder: Mathematics; PTSD = Posttraumatic Stress Disorder.

**Supplementary Table S5:** Correlations of the Factors in the Seventeen-Factor PsychProfiler Model

| | Large | Medium | Small | Nil |
|---|---|---|---|---|
| ADHDHI (1) | ADHD-I, CD, ODD, ASD, LD, SSD, GAD | SLD-R, SLD-W, SLD-M, SAD, OCD, PTSD, PDD | AN, BN | - |
| ADHDI (2) | ADHD-HI, ODD, ASD, LD, SSD, SLD-W, GAD | SLD-R, SLD-M, OCD, PTSD, PDD | SAD, AN, BN | - |
| CD (3) | ADHD-HI, ODD | ADHD-I, ASD, SSD, GAD, PTSD, PDD | LD, SLD-R, SLD-W, SLD-M, SAD, OCD, AN, BN | |
| ODD (4) | ADHD-HI, ADHD-I, CD, ASD, GAD, PDD | LD, SSD, SAD, OCD, PTSD, BN | SLD-R, SLD-W, SLD-M, AN | |
| ASD (5) | ADHD-HI, ADHD-I, CD, ODD, LD, SSD, GAD, SAD, OCD, PTSD, PDD | SLD-R, SLD-W, SLD-M, AN, BN | | |
| LD (6) | ADHD-HI, ADHD-I, ASD, SSD, SLD-R, SLD-W, SLD-M, GAD | ODD, SAD, OCD, PTSD, PDD | CD, AN, BN | |
| SSD (7) | ADHD-HI, ADHD-I, ASD, LD, SLD-R, SLD-W | CD, ODD, SLD-M, GAD, SAD, OCD, PTSD, PDD | AN, BN | |
| SLDR (8) | LD, SSR, SLD-W | ADHD-HI, ADHD-I, ASD, SLD-M | CD, ODD, GAD, SAD, OCD, PTSD, PDD, AN | BN |
| SLDW (9) | ADHD-I, LD, SSD, SLD-R | ADHD-HI, ASD, SLD-M | CD, ODD, GAD, SAD, OCD, PTSD, PDD | AN, BN |
| SLDM (10) | LD | ADHD-HI, ADHD-I, ASD, SSD, SLD-R, SLD-W, GAD, SAD, PDD | CD, ODD, OCD, PTSD | AN, BN |
| GAD (11) | ADHD-HI, ADHD-I, ODD, ASD, LD, SSD, SAD, OCD, PTSD, PDD, AN, BN | CD, SLD-M | SLD-R, SLD-W | |
| SAD (12) | ASD, GAD, OCD, PTSD, PDD | ADHD-HI, ODD, LD, SSD, SLD-M, AN, BN | ADHD-I, CD, SLD-R, SLD-W | |
| OCD (13) | ASD, GAD, SAD | ADHD-HI, ADHD-I, ODD, LD, SSD, PTSD, PDD, BN | CD, SLD-R, SLD-W, SLD-M, AN | |
| PTSD (14) | ASD, GAD, SAD, OCD, PDD, AN, BN | ADHD-HI, ADHD-I, CD, ODD, LD, SSD | SLD-R, SLD-W, SLD-M | |
| PDD (15) | ADHD-I, ODD, ASD, GAD, SAD, OCD, PTSD, AN, BN | ADHD-HI, CD, LD, SSD, SLD-M | SLD-R, SLD-W | |
| AN (16) | GAD, SAD, PTSD, PDD, BN | ASD, LD, SDD, SLD-R, SLD-M, SLD-M, SAD, OCD | ADHD-HI, ADHD-I. CD | |
| BN (17) | GAD, OCD, PTSD, PDD, AN | ODD, ASD, LD, SSD, SAD | | |
| | ADHD-HI, ADHD-I. CD, SLD-R, SLD-M | SLD-W | | |

GAD = Generalised Anxiety Disorder; SAD = Separation Anxiety Disorder, ADHDHI = Attention-Deficit/Hyperactivity Disorder: Hyperactive/Impulsive; ADHDI = Attention Deficit/Hyperactivity Disorder: Inattentive; ASD = Autism Spectrum Disorder; LD = Language Disorder; SSD =Speech Sound Disorder; PDD = Persistent Depressive Disorder; CD = Conduct Disorder; ODD = Oppositional Defiant Disorder; AN = Anorexia Nervosa; BN = Bulimia Nervosa; OCD = Obsessive-Compulsive Disorder; SLD-R = Specific Learning Disorder: Reading; SLD-W = Specific Learning Disorder: Written Expression; SLD-M = Specific Learning Disorder: Mathematics; PTSD = Posttraumatic Stress Disorder

***Supplementary Table S6:*** Effect sizes of the correlations of for each CAPP-SRF factor with all the other CAPP-SRF factors

| PP Scales | Conners DSM-5 symptom scales (N = 193) | | | | BYI-2 Scales (N = 56) | | | | | WISC-V Composite Score (N = 173) | | | | | | WIAT-III (N =173) | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | CADH-DI | CADH-DHI | CCD | CODD | SC | Anx | Dep | Anger | DB | VCI | VSI | FRI | WMI | PSI | FSIQ | Read | Written | Maths |
| GAD | .10 | .27*** | .34*** | .37*** | -.016 | .13 | .02 | .25 | .20 | -.08 | -.01 | -.05 | -.03 | -.02 | -.05 | -.21** | -.065 | -.133 |
| SAD | -.01 | .10 | .13 | .09 | -.24 | .22 | .19 | .37** | .40** | .00 | -.08 | .01 | -.02 | .07 | .02 | -.140 | .046 | -.143 |
| ADHDHI | .26*** | .68*** | .36*** | .37*** | .00 | .12 | .06 | .21 | .23 | -.05 | -.02 | -.04 | .09 | 0.16* | .02 | -.003 | -.041 | -.044 |
| ADHDI | .63*** | .56*** | .26*** | .27*** | -.28* | .23 | .29* | .19 | .28* | -.07 | -.06 | -.10 | .08 | .06 | -.02 | -.085 | .009 | -.021 |
| ASD | .05 | .20** | .26*** | .21** | -.10 | .25 | .13 | .09 | -.14 | -.07 | .03 | -.01 | .00 | .08 | .00 | -.045 | -.035 | -.090 |
| LD | .27*** | -.02 | -.05 | -.05 | -.20 | .13 | .28* | .06 | .20 | -.24*** | -.11 | -.15* | -.21** | -.07 | -.26*** | .165* | .111 | .103 |
| SSD | .13 | .01 | -.04 | -.02 | -.01 | .16 | .13 | .19 | .15 | -.14* | -.09 | -.04 | -.07 | -.02 | -.12 | .084 | .035 | -.039 |
| PDD | .14 | .21** | .26*** | .32*** | -.26 | .01 | .15 | .24 | .37** | -.12 | -.03 | -.02 | .07 | .11 | .01 | -.239** | -.061 | -.099 |
| CD | .12 | .42*** | .66*** | .52*** | .01 | .09 | -.06 | .11 | .13 | -.12 | .05 | -.02 | .06 | 0.14* | .01 | -.043 | -.035 | -.079 |
| ODD | .02 | .15* | .12* | .19** | NaNª | NaNª | NaNª | NaNª | NaNª | .06 | .09 | -.02 | .07 | .06 | .10 | -.045 | -.035 | -.048 |
| SLD-R | .06 | .00 | -.07 | -.09 | -.08 | .14 | .10 | -.13 | -.08 | -.30*** | -.08 | -.17* | -.25*** | -.11 | -.30*** | .260** | .074 | .140 |
| SLD-W | .28*** | .14 | .01 | .02 | -0.167 | .21 | .19 | .11 | .18 | -.29*** | -.11 | -.18** | -.30*** | -.18** | -.31*** | .254** | .105 | .220** |
| SLD-M | .15* | -.11 | -.13 | -.12 | -0.07 | .13 | -.09 | -.13 | -13 | -.32*** | -.26*** | -.25*** | -.29*** | -.27*** | -.46*** | .057 | .351** | .101 |
| PTSD | .04 | .13 | .13 | .11 | NaNᵇ | NaNᵇ | NaNᵇ | NaNᵇ | NaNᵇ | .02 | -.05 | -.04 | .01 | -.08 | .01 | -.129 | -.016 | -.072 |

**p < .01; *p < .05.

*Note.* The shaded boxes are the expected associations for determining criterion validity of the CAPP scales. The darker shaded boxes indicate expected negative correlations, and the lighter shaded boxes indicate expected positive correlations.

GAD = Generalised Anxiety Disorder; SAD = Separation Anxiety Disorder, ADHDHI = Attention-Deficit/Hyperactivity Disorder: Hyperactive/Impulsive; ADHDI = Attention Deficit/Hyperactivity Disorder: Inattentive; ASD = Autism Spectrum Disorder; LD = Language Disorder; SSD =Speech Sound Disorder; PDD = Persistent Depressive Disorder; CD = Conduct Disorder; ODD = Oppositional Defiant Disorder; AN = Anorexia Nervosa; BN = Bulimia Nervosa; OCD = Obsessive-Compulsive Disorder; SLD-R = Specific Learning Disorder: Reading; SLD-W = Specific Learning Disorder: Written Expression; SLD-M = Specific Learning Disorder: Mathematics; PTSD = Posttraumatic Stress Disorder; .CADHD-I = Conners's ADHD-I; CADHD-HI = Conner's ADHD-HI, CDS = Conner's CD; CODD = Conner's ODD; SC = Self-Concept; Anx =Anxiety; Dep = Depression; DB = Disruptive Behaviour; BYI-2 = Beck Youth Inventories, Second Edition; WISC-V = Wechsler Intelligence Scale for Children, Fifth Edition; VCI = Verbal Comprehension Index; VSI = Visual Spatial Index; FRI = Fluid Reasoning Index; WMI = Working Memory Index, PSI = Processing Speed Index; FSIQ = Full-Scale Intelligence Quotient.

**Supplementary Table S7:** Test of Criterion Validity of the Factors in the Seventeen-Factor PsychProfiler Model

# Discussion

The primary aim of the current study was to use CFA to examine support for its 17-factor structure in a group of children. It also examined the reliability (alpha and omega coefficients), and validity (discriminant and criterion validity) for the factors in this model. Based on extremely low frequencies for AN, BN, and OCD these factors were excluded, thereby leaving a 14-factor structure to test, instead of the originally proposed 17-factor structure. Findings supported a 13-factor model, in which the items for GAD and PDD were combined into a single factor. For this model, all items loaded significantly and saliently on their respectively designated factors (i.e., the factors in the model were clearly defined). In addition, virtually all of the factors showed good discriminant validity, and good internal consistency reliability (i.e., omega coefficients ranged from .858 to .950 and their alpha coefficients ranged from .851 to .949). Most of the factors were correlated with each other. The externalizing and neurodevelopmental disorders demonstrated adequate discriminant and concurrent validity. This was not the case for the internalizing disorders. These findings can be interpreting as indicating somewhat acceptable psychometric properties for a slightly revised version of the CAPP-TRF. These findings have practical and clinical implications in relation to the utilisation of the CAPP-TRF, and additional implications for understanding the co-morbidity of common childhood disorders, and revision for the CAPP-TRF.

## Practical Implications

Overall, our findings can be interpreted as generally supportive of the psychometric properties of the CAPP-TRF. As such, it can be concluded that the CAPP-TRF has sufficient satisfactory psychometric properties for use as a screening measure of mental disorders in children. Importantly, as the different screening scales in the CAPP-TRF correspond to the symptoms proposed for the DSM-5-TR disorders with the same name, it could be assumed that for children, the CAPP-TRF screening scales are suitable for teacher screening of the DSM-5-TR clinical disorders with the same name. Notwithstanding the argument here that the CAPP-TRF is a useful scale for screening DSM-5-TR child psychopathologies, there are limitations to keep in mind when using this measure. First, there was no support for the discriminant validity between GAD and PDD. Second, important other psychometric properties are still missing for the clinical use of the CAPP-TRF, most importantly of which is a lack data on its predictive validity. Third, is that for calculating screening scores for a disorder, the item ratings are recoded as follows: never, rarely, sometimes = 0 (symptom absent); and regularly, often, and very often = 1 (symptom present), and the total recoded scores of items within each CAPP scale is used to ascertain if the disorder is present. Although the cut-off scores for each screening scale are identical with the symptom threshold cut-off scores in the DSM-5-TR, estab-

lishing cut-off scores for clinical disorders in screening scales is more complex. Best practice standards require application of empirically derived diagnostic utility statistics (such as, sensitivity, specificity, positive predictive power and negative predictive power). It is important therefore that clinicians keep this in mind when using the cut-off scores recommended for the CAPP-TRF.

## Implications for Comorbidity

Although all the 13 factors were correlated significantly with each other, there were noticeably strong correlations of the disorders within the traditionally considered neurodevelopmental disorders, and internalizing disorders. Except for ADHD-I, the disorders within the traditionally considered externally disorders also showed noticeably strong correlations with each other. ADHD-I was relatively associated more strongly with neurodevelopmental disorders than with externalizing disorders. These associations dictate that it is plausible that the CAPP-TRF screening scales can be clustered together into externalizing, internalizing, and neurodevelopmental groups, with ADHD-I being part of the neurodevelopmental cluster rather than the externalizing cluster.

Given the present study found significant correlations between all the factors in the CAPP-TRF it can be assumed that the factors have some degree of shared variance. As the CAPP-TRF screening scales are aligned to the DSM-5-TR disorders with the same names, it can be argued that our findings indicate considerable support for a high degree of comorbidity among the disorders that are screened by the CAPP-TRF. Taking into consideration large effect sizes between the CAPP-TRF factors, it can be speculated that there would be a possibility for more comorbidity within the internalizing disorder; with the exception of ADHD-I, within the externalizing disorders, and within the neurodevelopmental disorders, and for ADHD-I with neurodevelopmental disorders.

## Implication for the HiTop Model

Given the patterns of associations between the disorders in the current study, our findings have implications for the Hierarchical Taxonomy of Psychopathology model (HiTOP; [49,50]. The HiTOP model is a dimensional model of psychopathology, moving upwards from narrow to broader constructs of psychopathology [49,50]. The problems/syndromes/disorders in HiTOP are at five different hierarchical levels. Of relevance to the current study, at the very top is the superspectra or general p-factor (Kotov et al., 2021). This factor corresponds to the finding in this present study that all the CAPP-TRF disorders were correlated with each other. Just below the superspectra there are different spectra, followed by the subfactors. The spectra (6 in all) include internalising subfactors for anxiety, depression and eating

disorders; externalising subfactors for ADHD, CD and ODD. Although the initially proposed HiTOP model [49] did not include a spectrum for neurodevelopmental problems, a relatively recent study supports its inclusion [51]. Considering this, our findings can not only be accommodated in the HiTOP model, but they also suggest an important extension of that model. More specifically, they support a neurodevelopmental spectrum [51], that includes at the very least ASD, LD, SSD, SLD-R, SLD-W, and SLD-M. Our findings appear to indicate that the location of ADHD-HI and ADHD-I may not be in the same spectrum. Rather, ADHD-I is likely to be in the neurodevelopmental spectrum, and ADHD-HI in the externalizing spectrum. These findings also mean that the CAPP-TRF may potentially be a useful tool for research involving the HiTOP model [52].

### Implications for the Alternate Use and Revisions of the CAPP-TRF

As is evident, we found support for a revised 13-factor model in which the indicators for GAD with PDD were merged into a single factor mixed anxiety/depression disorder (GAD/DPP) factor. Thus, our findings suggest that at present the current CAPP-TRF is better viewed as comprising 13 scales comprising ADHD-HI, ADHD-I, ODD, CD, SLD-R, SLD-W, SLD-M, ASD, LD, SSD, GAD/PDD, SAD, PTSD. Related to this, it is important to note, our exclusion of factors for OCD, AN and BN because they were not included in the analysis due to their very low frequencies in our data set. Thus, our support for a 13-factor model for the CAPP-TRF, needs to be considered with this in mind. Our findings for a combined factor for GAD with PDD (GAD/DPP) and lack of support for the discriminant validity for SLD-R and SLD-W would also mean the need for revisions of these scales (GAD, PDD, SLD-R and SLD-W).

### Limitations

Although a strength of the present study is its large sample and the data generated from this sample came from children located in different parts of Australia, there are limitations to be considered. The current study examined the CAPP-TRF in children and since this measure was developed for use with children and adolescents aged 2 to 18 years of age our findings cannot be generalised to teacher ratings of the CAPP-TRF for adolescents. Furthermore, the study did not include a wider range of external variables that would have allowed for more comprehensive evaluation of the criterion validity of all the CAPP-TRF screening scales.

### Summary and concluding remarks

In summary, the overall findings of this study illustrate acceptable psychometric properties for a slightly revised 13-factor version of the CAPP-TRF and indicate that the CAPP-TRF is a suitable global screening instrument for teacher screening of DSM-5-TR disorders in children.

### Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

### Supplementary material

The Supplementary material for this article can be found online at: www.psychprofiler.com

# References

1. Australian Institute of Health and Welfare (2000). Australia's health 2000: the seventh biennial health report of the Australian Institute of Health and Welfare. Canberra: AIHW.
2. Lawrence, David, Jennifer Hafekost, Sarah E. Johnson, Suzy Saw, William J. Buckingham, Michael G. Sawyer, John Ainley, and Stephen R. Zubrick. "Key findings from the second Australian child and adolescent survey of mental health and wellbeing." Australian & *New Zealand Journal of Psychiatry* 50, no. 9 (2016): 876-886.
3. Hafekost, Jennifer, Sarah Johnson, David Lawrence, Michael Sawyer, John Ainley, Cathrine Mihalopoulos, and Stephen R. Zubrick. "Introducing 'young minds matter'." *Australian Economic Review* 49, no. 4 (2016): 503-514.

4. Racine, Nicole, Brae Anne McArthur, Jessica E. Cooke, Rachel Eirich, Jenney Zhu, and Sheri Madigan. "Global prevalence of depressive and anxiety symptoms in children and adolescents during COVID-19: a meta-analysis." *JAMA pediatrics* 175, no. 11 (2021): 1142-1150.

5. Jullien, Sophie, Ivelina Borisova, Joao Breda, Susanne Carai, Gabriele Fontana, Aleksandra Jovic, Martin M. Weber, Octavian Bivol, and Natasha Azzopardi Muscat. "Refocusing on the foundations: strategy for child and adolescent health in Europe and Central Asia 2026–2030–A healthy start for a healthy life." *Journal of Global Health* 15 (2025): 03046.

6. Baird, Sarah, Shakira Choonara, Peter S. Azzopardi, Prerna Banati, Judith Bessant, Olivia Biermann, Anthony Capon et al. "A call to action: the second Lancet Commission on adolescent health and wellbeing." *The Lancet* 405, no. 10493 (2025): 1945-2022.

7. Paton, K., S. Darling, C. Nowell, S. Gandhi, A. Jorm, L. M. Hart, M. B. H. Yap, and F. Oberklaid. "Development of the children's wellbeing continuum: Fostering conversation in child mental health." *Mental Health & Prevention* 37 (2025): 200385.

8. De Los Reyes, Andres, Tara M. Augenstein, Mo Wang, Sarah A. Thomas, Deborah AG Drabick, Darcy E. Burgers, and Jill Rabinowitz. "The validity of the multi-informant approach to assessing child and adolescent mental health." *Psychological bulletin* 141, no. 4 (2015): 858.

9. Srinath, Shoba, Preeti Jacob, Eesha Sharma, and Anita Gautam. "Clinical practice guidelines for assessment of children and adolescents." *Indian journal of psychiatry* 61, no. Suppl 2 (2019): 158-175.

10. American Psychiatric Association. (2022). Diagnostic and statistical manual of mental disorders (5th ed., text revision). Arlington, VA: *American Psychiatric Publishing*. https://doi.org/10.1176/appi.books.9780890425787.

11. Langsford, S., Houghton, S., & Douglas, G. (2007). PsychProfiler v4 Manual. Melbourne: ACER Press

12. Langsford, S., Houghton, S., & Douglas, G. (2014). PsychProfiler v5 Manual. Psychological and Educational Consultancy Services.

13. Langsford, S., Houghton, S., & Douglas, G. (2022). PsychProfiler v5 Manual, Second Edition. Psychological and Educational Consultancy Services.

14. American Psychiatric Association. (2013). Diagnostic and statistical manual of mental disorders (5th ed.). Washington, DC: Author. https:// doi.org/10.1176/appi.books.9780890425596.

15. Gomez, Rapson, Shane Langsford, Stephen Houghton, and Leila Karimi. "Psychometric Properties of the Child and Adolescent PsychProfiler: Self-Report Form." Assessment (2025): 10731911251398037.

16. Langsford, Shane, Rapson Gomez, Stephen Houghton, and Leila Karimi. "Psychometric properties of the Child and Adolescent PsychProfiler v5: a measure for screening 14 of the most common DSM-5 disorders." *Frontiers in Psychology* 15 (2024): 1267711.

17. Lawrence, D., Langsford, S. & Houghton, S. (2020, unpublished paper). Investigating the reliability and validity of the PsychProfiler v5.

18. American Educational Research Association, American Psychological Association, National Council on Measurement in Education, & Joint Committee on Standards for Educational and Psychological Testing. (2014) Standards for educational and psychological testing.

19. Hawkins, Melanie, Gerald R. Elsworth, Elizabeth Hoban, and Richard H. Osborne. "Questionnaire validation practice within a theoretical framework: a systematic descriptive literature review of health literacy assessments." *BMJ open* 10, no. 6 (2020): e035974.

20. Bandalos, Deborah L., and Sara J. Finney. "Factor analysis: Exploratory and confirmatory." In The reviewer's guide to quantitative methods in the social sciences, pp. 98-122. Routledge, 2018.

21. Hurley, Amy E., Terri A. Scandura, Chester A. Schriesheim, Michael T. Brannick, Anson Seers, Robert J. Vandenberg, and Larry J. Williams. "Exploratory and confirmatory factor analysis: Guidelines, issues, and alternatives." *Journal of organizational behavior* (1997): 667-683.

22. Beck, J. S., Beck, A. T., Jolly, J. B., & Steer, R. (2005). Beck Youth Inventories Second Edition for children and adolescents (BYI-II). San Antonio, Texas: *Harcourt Assessment*, Inc.

23. Conners, C. K. (2008). Conners 3 manual (3rd ed.). Toronto, Ontario, Canada: Multi-Health Systems Inc.

24. Wechsler, D. (2016). Wechsler individual achievement test, Third Edition: Australian and *New Zealand Standardised Edition*. Bloomington, MN: Pearson.

25. Thastum, Mikael, Kristine Ravn, Søren Sommer, and Anegen Trillingsgaard. "Reliability, validity and normative data for the Danish Beck Youth Inventories." *Scandinavian Journal of Psychology* 50, no. 1 (2009): 47-54.

26. JASP Team (2023). JASP (Version 0.16.6.0) [Computer Software].

27. Brown, T. A. (2006). Confirmatory factor analysis for applied research. New York: The Guilford Press.

28. Streiner, D. L., & Norman, G. R. 1995. Health measurement scales: A practical guide to their development and use (2nd ed). New York: *Oxford University Press*.

29. Muthen, L. K. & Muthen, B. O. (1998-2012). Mplus User's Guide. Seventh Edition. Los Angeles, CA: Muthen & Muthen.

30. Beauducel, Andre, and Philipp Yorck Herzberg. "On the performance of maximum likelihood versus means and variance adjusted weighted least squares estimation in CFA." Structural Equation Modeling 13, no. 2 (2006): 186-203.

31. Lubke, Gitta H., and Bengt O. Muthén. "Applying multigroup confirmatory factor models for continuous outcomes to Likert scale data complicates meaningful group comparisons." Structural equation modeling 11, no. 4 (2004): 514-534.

*J. Psychiatry. Psychiatr. Disord. Vol. 5 Iss. 1 (039)*

*Page-22*

32. Millsap, Roger E., and Jenn Yun-Tein. "Assessing factorial invariance in ordered-categorical measures." *Multivariate behavioral research* 39, no. 3 (2004): 479-515.

33. DiStefano, Christine, Heather L. McDaniel, Liyun Zhang, Dexin Shi, and Zhehan Jiang. "Fitting large factor analysis models with ordinal data." *Educational and Psychological Measurement* 79, no. 3 (2019): 417-436.

34. Kline, R. B. (2011). Principles and practice of structural equation modelling (3rd ed.). The Guilford Press.

35. Shi, Dexin, Taehun Lee, and Alberto Maydeu-Olivares. "Understanding the model size effect on SEM fit indices." *Educational and psychological measurement* 79, no. 2 (2019): 310-334.

36. Pavlov, Goran, Alberto Maydeu-Olivares, and Dexin Shi. "Using the standardized root mean squared residual (SRMR) to assess exact fit in structural equation models." *Educational and Psychological Measurement* 81, no. 1 (2021): 110-130.

37. Ullman, J. B. (2001). Structural equation modeling. In: B. G. Tabachnick, & L. S. Fidell (Eds.), Using multivariate statistics. Boston, MA: Pearson Education.

38. Schumacker, R. E., & Lomax, R. G. (2004). A beginner's guide to structural equation modeling. Psychology Press. Lawrence Erlbaum Associates, Publishers Mahwah New Jersey.

39. Hu, Li-tze, and Peter M. Bentler. "Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives." Structural equation modeling: a multidisciplinary journal 6, no. 1 (1999): 1-55.

40. Byrne, Barbara M. Structural equation modeling with Mplus: Basic concepts, applications, and programming. routledge, 2013.

41. Field, A. (2013). Discovering statistics using IBM SPSS statistics. Sage.

42. Zinbarg, Richard E., William Revelle, Iftah Yovel, and Wen Li. "Cronbach's α, Revelle's β, and McDonald's ωH: Their relations with each other and two alternative conceptualizations of reliability." psychometrika 70, no. 1 (2005): 123-133.

43. Watkins, Marley W. "The reliability of multidimensional neuropsychological measures: From alpha to omega." The clinical neuropsychologist 31, no. 6-7 (2017): 1113-1126.

44. Cohen, Jacob. "Statistical power analysis." Current directions in psychological science 1, no. 3 (1992): 98-101.

45. Soper, D. S. (2022). A-priori Sample Size Calculator for Structural Equation Models [Software]. Available from https://www.danielsoper.com/statcalc

46. Graham, John W. "Missing data analysis: Making it work in the real world." *Annual review of psychology* 60, no. 1 (2009): 549-576.

47. Fairburn, Christopher G., and Kristin Bohn. "Eating disorder NOS (EDNOS): an example of the troublesome "not otherwise specified"(NOS) category in DSM-IV." *Behaviour research and therapy* 43, no. 6 (2005): 691-701.

48. Achenbach, T. M., & Rescorla, L. A. (2001). Manual for ASEBA school age forms & profiles. Burlington: Research Centre for Children, Youth and Families, University of Vermont; 2001.

49. Kotov, Roman, Robert F. Krueger, David Watson, Thomas M. Achenbach, Robert R. Althoff, R. Michael Bagby, Timothy A. Brown et al. "The Hierarchical Taxonomy of Psychopathology (HiTOP): A dimensional alternative to traditional nosologies." *Journal of abnormal psychology* 126, no. 4 (2017): 454.

50. Ruggero, Camilo J., Roman Kotov, Christopher J. Hopwood, Michael First, Lee Anna Clark, Andrew E. Skodol, Stephanie N. Mullins-Sweatt et al. "Integrating the Hierarchical Taxonomy of Psychopathology (HiTOP) into clinical practice." *Journal of consulting and clinical psychology* 87, no. 12 (2019): 1069.

51. Michelini, Giorgia, Deanna M. Barch, Yuan Tian, David Watson, Daniel N. Klein, and Roman Kotov. "Delineating and validating higher-order dimensions of psychopathology in the Adolescent Brain Cognitive Development (ABCD) study." *Translational psychiatry* 9, no. 1 (2019): 261.

52. Simms, Leonard J., Aidan GC Wright, David Cicero, Roman Kotov, Stephanie N. Mullins-Sweatt, Martin Sellbom, David Watson, Thomas A. Widiger, and Johannes Zimmermann. "Development of measures for the Hierarchical Taxonomy of Psychopathology (HiTOP): A collaborative scale development project." Assessment 29, no. 1 (2022): 3-16.

53. World Health Organization. (2022). ICD-11: International classification of diseases (11th revision). https://icd.who.int/.