



ORIGINAL ARTICLE

Internalizing problems can differ in boys and girls since early childhood: findings from the Child Behavioral Checklist 1.5-5

Adriana Maldonado-Martinez,^{1,2} Sheila C. Caetano,³ Marcos V. Ribeiro,³
Alexandra Restrepo-Henao,^{1,2} Paola M. Okuda,³ Thiago M. Fidalgo,³ Pamela J. Surkan,⁴
Luis Manuel Silva,⁵ Silvia S. Martins¹

¹Department of Epidemiology, Mailman School of Public Health, Columbia University, New York, NY, USA. ²Facultad Nacional de Salud Pública, Universidad de Antioquia, Medellín, Colombia. ³Departamento de Psiquiatria, Escola Paulista de Medicina, Universidade Federal de São Paulo, São Paulo, SP, Brazil. ⁴Department of International Health, Bloomberg School of Public Health, Johns Hopkins University, Baltimore, MD, USA. ⁵Facultad de Psicología, Pontificia Universidad Javeriana, Bogotá, Colombia.

Objective: Internalizing problems disproportionately affect females in adolescence and adulthood, but research at earlier ages is limited due to a focus on disruptive behaviors. Our study addresses this gap by exploring the structure of internalizing problems and gender differences in Brazilian preschoolers.

Methods: We analyzed data from the Child Behavioral Checklist 1.5-5 (CBCL 1.5-5) as administered in the Preschool Mental Health Study (PreK Survey), involving 1,292 children aged 4 to 5 in Embu das Artes, state of São Paulo, Brazil. Confirmatory factor analysis and comparisons of means explored internalizing problems and gender variations.

Results: A two-factor model best fit both internalizing and externalizing problems. A hierarchical model with four factors (emotionally reactive, anxiety/depression, somatic complaints, and withdrawn) best fit internalizing problems, achieving partial invariance between boys and girls. Boys scored higher in the withdrawn syndrome, while girls scored higher in the somatic complaint syndrome.

Conclusion: Preschoolers' internalizing problems warrant attention beyond their link to externalizing problems. While the overall construct is similar in boys and girls, divergent syndrome scores indicate potential distinct risk patterns requiring further exploration.

Keywords: Preschool mental health; anxiety; depression; gender differences; confirmatory factor analysis

Introduction

According to a 2017 World Health Organization (WHO) report,¹ Brazil has the highest global prevalence of anxiety (9.3%) and associated disability, and ranks fifth in depression prevalence (5.8%) among individuals aged 15 or older. In young children, such as preschoolers, internalizing problems (clinically significant symptoms) receive limited attention,² mainly because they are not as overtly disruptive as externalizing problems (aggressive behaviors or attention problems). Additionally, detection of internalizing problems in preschoolers relies on third-party observations, typically from parents or caregivers,³ due to children's inability to accurately report their symptoms at such an early age.

Based on a study comparing data from two birth cohorts (1993 and 2000) in Pelotas,⁴ Brazil, preschoolers' internalizing problems showed no significant increase in

prevalence over this 11-year period. However, there was a 10% rise in scores, indicating an escalation in severity. This is worrisome considering the potential for continuity into adolescence and progression to mental disorders in later years.^{5,6} Unfortunately, there is insufficient research on preschoolers' internalizing problems in Brazil, and, consequently, little evidence to inform preventive and therapeutic interventions for young children.

In our previous systematic review of the literature, we found no research on the factor structure of internalizing problems or validated behavioral problem scales for preschoolers in Brazil. Nevertheless, three⁷⁻⁹ of the six included Brazilian studies administered adapted versions of the Child Behavioral Checklist (CBCL 1.5-5),¹⁰ the most widely used scale for preschoolers' behavioral problems, while two studies used the CBCL 4-18.^{4,11}

In the CBCL 1.5-5,¹⁰ behavioral problems in preschoolers are categorized into a seven-factor structure: there

Correspondence: Adriana Maldonado Martinez, 722 West 168th Street, Mailman School of Public Health, New York, NY, USA, 10032.

E-mail: am3615@columbia.edu

Submitted Mar 07 2024, accepted Aug 03 2024.

How to cite this article: Maldonado-Martinez A, Caetano SC, Ribeiro MV, Restrepo-Henao A, Okuda PM, et al. Internalizing problems can differ in boys and girls since early childhood: findings from the Child Behavioral Checklist 1.5-5. Braz J Psychiatry. 2025;47:e20243616. <http://doi.org/10.47626/1516-4446-2024-3616>

are two factors for externalizing problems, four factors for internalizing problems, and a single factor for sleep problems. The two first-order factors or externalizing syndromes are “aggressive behavior” and “attention problems.” The four first-order internalizing syndromes are “anxious/depressed,” “somatic complaints,” “withdrawn,” and “emotionally reactive.” A study tested the seven-factor model of the CBCL 1.5-5 with data from 23 countries in Asia, Australasia, Europe, the Middle East, and South America (not including Brazil).¹² The seven-factor model was a good fit for the data in each of these countries. In Latin America, this included a report of the construct validity of the CBCL 1.5-5 in Chile¹³ – a high-income country (HIC) – and a report of criterion validity in Mexico.¹⁴

To date, there are no reports of validation of the CBCL 1.5-5 for measuring internalizing problems in Brazilian preschoolers. Additionally, the potential for misspecification of the four-factor model, including issues such as incorrect factor number specification or omitting cross-loading of items,¹⁵ remains uncertain for this population. A study¹⁶ of older Brazilian children revealed a dominant general psychopathology factor (“p factor”¹⁷) in the Development and Well-Being Assessment (DAWBA), accounting for most variance in a community sample of 2,512 children aged 6 to 12 years in São Paulo and Porto Alegre. Although the CBCL was not used and preschoolers were not included in the sample, three specific factors (fear, distress, and externalizing problems) accounted for little residual variance compared to a single factor, mirroring findings in adolescents and adults in HIC.^{17,18} Whether the CBCL’s hierarchical model or a general p factor is the best fit for Brazilian preschoolers remains unexplored. Brazilian studies on the CBCL 4-18 (for children aged 4 to 18 years)¹⁹ validated a hierarchical model and affirmed the multicultural robustness of the CBCL 4-18 in the country.²⁰

Another aspect heretofore unexplored in Brazil is the difference in internalizing problems by sex. In Brazil, as in several countries,²¹ adolescent and adult females have a significantly higher prevalence of internalizing disorders than males, at least twice as high. Adolescent and adult males are less likely than females to seek treatment for internalizing disorders or discuss their symptoms²² (being withdrawn). Females are more likely than males to develop conditions usually linked to internalizing disorders, such as irritable bowel syndrome²³ (having somatic complaints). One plausible explanation for these dissimilarities involves different gender roles and societal expectations,²⁴ which expose females and males to different risk factors (such as males being the providers and women being the caregivers). An exploration of differences in the structure and characteristics of internalizing problems in females and males in the preschool years, when children are still developing their gender identity,²⁵ could help elucidate the mechanisms that lead to this disproportionality in prevalence estimates of internalizing disorders in adolescent and adult females versus males.

Considering that i) research on internalizing problems in preschoolers in Brazil is scarce, ii) internalizing

problems are rarely the focus of research in this age group compared to externalizing problems, iii) there is a lack of valid measures of internalizing problems in this population, iv) there are potential differences by sex on internalizing problems at an early age, v) the CBCL 1.5-5 is the most widely used scale to measure internalizing problems, and vi) understanding the construct of internalizing problems in preschoolers could guide prevention and treatment efforts to reduce the incidence of internalizing disorders in Brazil, the present study focused on analyzing the structure and characteristics of internalizing problems in the CBCL 1.5-5 differentially by sex in a representative sample of preschoolers from the Brazilian city of Embu das Artes, state of São Paulo. This study does not intend to validate the full CBCL 1.5-5 questionnaire for use in Brazil, but rather to gain a better understanding of preschoolers’ internalizing problems in this setting.

Methods

Data were obtained from the Preschool Mental Health Study (PreK Survey).²⁶

The Preschool Mental Health Study

The PreK Survey^{26,27} was a 2016 cross-sectional study of a sample of 1,292 preschool children (666 boys and 626 girls) aged 4-5 years and their caregivers in Embu das Artes, a city in the metropolitan area of São Paulo, Brazil. The study was developed by researchers at Departamento de Psiquiatria and Departamento de Medicina Preventiva, Universidade Federal de São Paulo (UNIFESP), and collaborators at Columbia University Mailman School of Public Health and Johns Hopkins Bloomberg School of Public Health. Given that mental illness starts early in life and disproportionately affects children from low- and middle-income countries (LMIC), the main objective of the PreK Survey was to assess the prevalence rates of internalizing and externalizing problems (clinically significant symptoms) and socio-emotional development delays among a representative sample of public preschool children aged 4-5 years in a poor urban region of Brazil.²⁷

Using an online database from the Brazilian Instituto Nacional de Estudos e Pesquisas Educacionais, all public preschools in Embu das Artes were identified. A representative sample of the 4-to-5-year-old population was recruited by stratified random sampling of public preschools/childcare centers, selected with a probability proportional to the number of 4-to-5-year-olds in each school within census tracts. After obtaining informed consent, trained health professionals from UNIFESP collected data by face-to-face interviews conducted in Portuguese with a parent or legal guardian of each 4-to-5-year-old from each preschool. The study included the Brazilian Portuguese version of the CBCL 1.5-5.¹⁰ The recruitment process and sociodemographic characteristics of the sample have been thoroughly described and documented previously.^{26,27} Based on the CBCL 1.5-5 scores obtained from the PreK Survey, the 6-month

prevalence was 25.4% for internalizing problems and 12.1% for externalizing problems.

Instrument

The CBCL 1.5-5¹⁰ is applied to primary caregivers of children aged 1.5 to 5. It examines three dimensions of problematic behaviors in preschoolers as observed by their parents: internalizing, externalizing, and sleep problems. It consists of 99 items that are rated 0 for “not true” of the child, 1 for “somewhat or sometimes” true, and 2 for “very true or often” true, each based on the preceding 2 months.

The four internalizing syndromes (somatic complaints, emotionally reactive, withdrawn, and anxious/depressed) and the corresponding problems are described in Box 1. The terms “syndromes” and “first-order factors” are synonymous in the CBCL and will be used interchangeably

throughout this text. Syndrome is the name that the authors give to the first-order factors in the CBCL 1.5-5.

Data analysis

The analyses were conducted in four stages to examine the structure of internalizing items in the CBCL 1.5-5. Analyses were based on confirmatory factor analysis (CFA), which postulates certain relationships among the observed and the latent variables, assuming a pre-specified pattern for the model parameters (factor loadings, structural parameters, residual variances).²⁸ Findings at each stage guided the analyses in the subsequent stages. All are summarized in Figure 1.

In stage 1, all 60 internalizing and externalizing questions of the original CBCL 1.5-5 scales adapted to Portuguese were included to test if a two-factor model (internalizing and externalizing, model 2a) was a better fit for behavioral problems than a model with only one factor

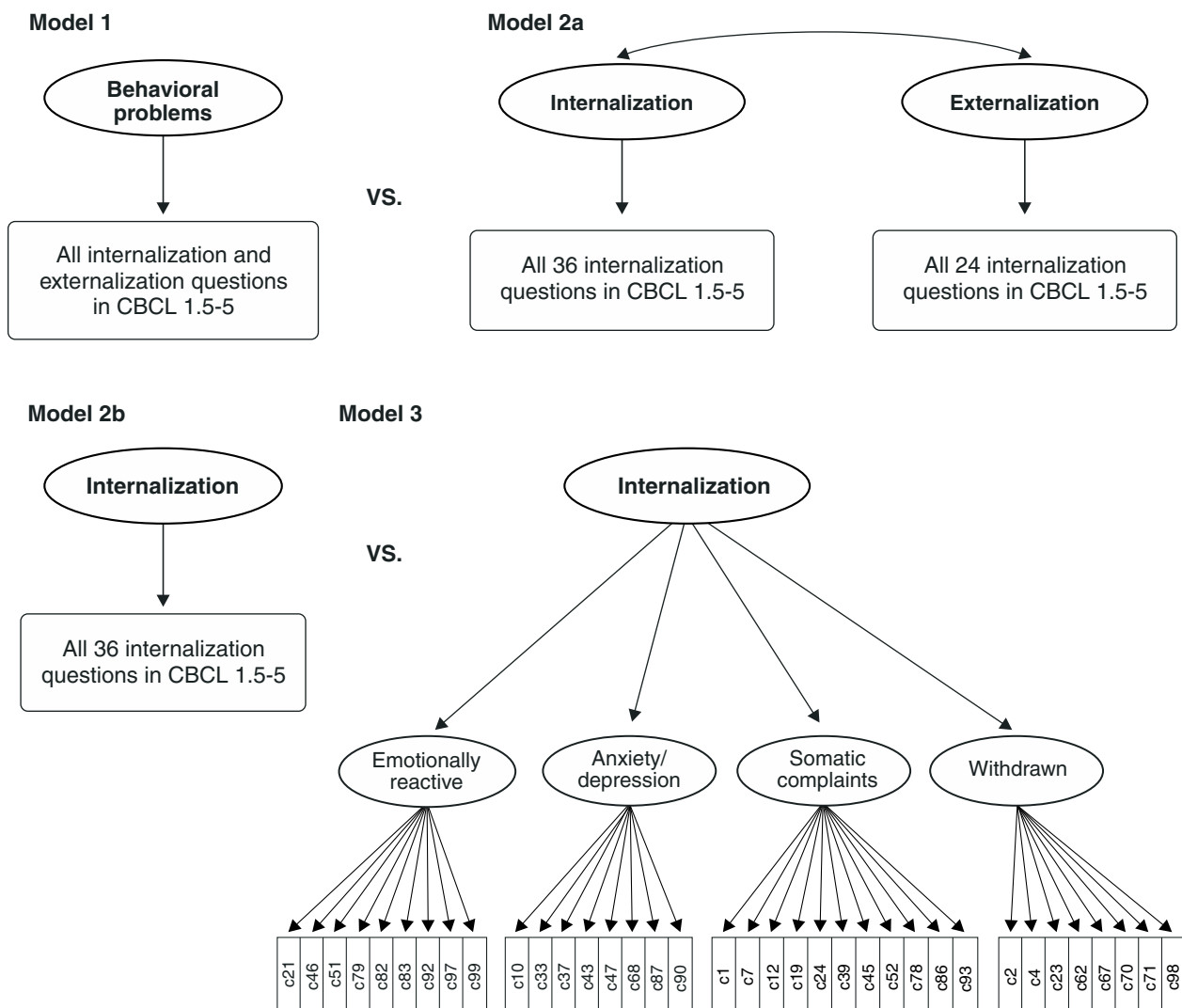


Figure 1 Models and stages of the analyses. Stage 1 = model 1 vs. model 2a; stage 2 = model 2b vs. model 3; stage 3 = model 3 for boys vs. girls (multigroup confirmatory factor analyses [CFA]); stage 4 = calculation of scores for model 3 separately for boys and girls. CBCL 1.5-5 = Child Behavioral Checklist 1.5-5.

Box 1 Internalizing questions in the CBCL

Internalizing questions in the CBCL (3-level responses: never true, sometimes true, always true)			
Somatic complaints	Emotionally reactive	Withdrawn	Anxiety/depression
1. Aches	21. Disturbed by change	2. Acts too young	10. Clings
7. Can't stand things out of place	46. Twitches	4. Avoids eye contact	33. Feelings hurt
12. Constipated	51. Panics	23. Doesn't answer	37. Upset by separation
19. Diarrhea	79. Shifts between sad-excited	62. Refuses active games	43. Looks unhappy
24. Doesn't eat well	82. Moody	67. Unresponsive to affection	47. Nervous
39. Headaches	83. Sulks	70. Little affection	68. Self-conscious
45. Nausea	92. Upset by new	71. Little interest	87. Fearful
52. Painful bowel movements	97. Whining	98. Withdrawn	90. Sad
78. Stomach aches	99. Worries		
86. Too concerned with neatness or cleanliness			
93. Vomits			

(model 1). The hypothesis for this first stage was based on the estimates of the chi-square difference test (at a 5% level of significance) and comparisons of other fit indices; model 2a would have a significantly better fit than model 1.

Model 2a having a better fit than model 1 indicated that internalizing problems was a distinct factor from externalizing problems. Since the focus of the study was internalizing problems, externalizing questions were excluded in stage 2, and only the 36 questions on internalization were included. The fit of two models (models 2b and 3) was compared. Model 2b consisted of only one internalizing factor, while model 3 was a hierarchical model of four first-order factors/syndromes factor (anxious/depressed, somatic complaints, withdrawn, and emotionally reactive) and one second-order factor for all internalizing problems. The hypothesis for this stage was based on the estimates of the chi-square difference test (at a 5% level of significance) and comparisons of other fit indices; model 3 would have a significantly better fit than model 2b.

In stage 3, since model 3 was a better fit than model 2b, multigroup CFA was conducted to assess measurement invariance by sex, to test how “stable” a given measurement model is across groups. Since this study does not aim to validate the internalizing symptoms scale and knowing that full measurement invariance in all steps (configural, metric, scalar, and residual) is often not supported in the literature, only configural and metric invariance were analyzed, providing partial measure invariance discussion. Configural invariance (equal factor configurations) means that the items measure the same construct in boys and girls (i.e., are best explained by the same model: four first-order factors and one-first-order factor). Metric invariance (equal loadings) means that the factor loadings on the items are equivalent across groups.²⁹ The hypotheses were that configural and metric invariance by sex would not be rejected based

on non-significant chi-square tests (at the 5% level) and on the examination of model fit indices such as the comparative fit index (CFI), Tucker-Lewis index (TLI), standardized root mean square residual (SRMR), and the root mean square error approximation (RMSEA).³⁰

Finally, in stage 4, although partial metric invariance was not rejected, estimates for each syndrome were obtained separately for boys and girls based on factor loadings, followed by a description of the main factors and problems in boys and girls. For this analytic stage, the hypothesis was that the main syndrome (the factor most endorsed and with the highest loadings) for both boys and girls would be anxious/depressed. Other hypotheses were that in girls, the factor somatic complaints would have a greater correlation with the other factors than in boys, and, in boys, the factor withdrawn would have a greater correlation with the other factors than in girls.

All CFA analyses were conducted in Mplus version 8.3,³⁰ which considers the survey’s design effects (data were clustered by school), including parameter as well as standard error estimation and model fit calculations. The default estimator for the analyses was the variance-adjusted weighted least squares (WLSMV). This robust estimator does not assume normally distributed variables and provides the best option for modeling categorical or ordered data. Standard indices such as chi-square, CFI, TLI, RMSEA, and SRMR were used to assess the goodness of fit in CFA. A CFI and TLI > 0.95 and RMSEA < 0.07 would indicate a good model fit, and an SRMR < 0.08 would indicate an acceptable model fit. Scores in stage 4 were calculated using SAS 9.4.³¹

According to Kline,²⁸ there should be at least 10 people per item in CFA. Therefore, the sample size of 1,292 children (666 boys and 636 girls) was appropriate for the analyses. In stage 1, there were approximately 22 people per item (60 internalizing and externalizing questions). In stage 2, there were approximately 35 people per item

(36 internalizing questions). In stages 3 and 4, there were approximately 18 people per item in the male and female groups. There were no missing data in these analyses. The CBCL 1.5-5 questionnaire was not validated as a whole due to the following reasons: i) in the multigroup analyses, the study's sample size was not large enough for the 99 total items (fewer than seven people per item); ii) sleep problems are all symptoms related to one particular event and are therefore more specific than internalizing and externalizing symptoms; and iii) with this study design, we can place greater emphasis on internalizing problems than on externalizing problems, leading to an analytical and potentially practical discussion addressing internalizing issues and covering the gap in the literature, as described above.

Results

As previously stated, throughout the text, the terms syndrome and factor will be used interchangeably. Syndrome is how the authors refer to the first-order factors in the CBCL 1.5-53.

Stage 1

As seen in Table 1, chi-square tests of model fit were significant for model 1 (one factor for all items) and model 2a (two factors),²⁹ which is expected in large samples. However, there was a significant decrease in the chi-square between models 1 and 2a (chi-square difference test = 147.052, degrees of freedom [df] = 1, $p < 0.001$), indicating a better fit of model 2a. Considering the other fit indices, both models had an acceptable fit based on the CFI/TLI (≥ 0.90) and the SRMR (< 0.08), and a good fit based on the RMSEA (< 0.07). Since model 2a had better fit indices than model 1 (Table 1), it was selected in

this stage, indicating that internalizing and externalizing problems can be seen as two factors.

Stage 2

Model 2b (one factor including all internalizing problems) and model 3 (four first-order factors and one second-order factor) were compared, as shown in Table 1. After an examination of the R-square, items 7 ("Can't stand things out of place") and 86 ("Too concerned with cleanliness and neatness") explained little of the variance of the factor of internalization (each explained less than 7%). Hence, these two items were excluded from the analyses, improving the fit of model 2b and model 3, as seen in Table 1.

Models 2b and 3 were significantly different as per the chi-square test (420.149, $df = 4$, $p < 0.001$), and model 3 had better fit indices (Table 1); as such, model 3 (the hierarchical model – internalizing second-order factor and four syndromes first-order factors) was selected. This finding indicates that there is one second-order internalizing factor that the four first-order factors can explain: anxiety/depression, emotionally reactive, withdrawn, and somatic complaints, endorsing the structure for internalizing problems in the CBCL 1.5-5.

Stage 3

Upon exploration of all the items of the anxious/depressed factor, the question "Is the child nervous or tense" (item 47) had a higher correlation with the factor emotionally reactive and a higher factor loading than with the factor anxious/depressed ($r = 0.754$ vs. $r = 0.643$, factor loadings: 0.748 vs. 0.734) in boys. Since the question "Is nervous or tense" also had a higher correlation with the factor emotionally reactive ($r = 0.683$ vs. $r = 0.646$) in girls, the question was moved to that factor.

Table 1 Fit indices of models

Fit indices after CFA	Internalizing and externalizing		Internalizing only	
	Model 1	Model 2a	Model 2b	Model 3
CFI	0.917	0.928	0.925	0.962
TLI	0.914	0.926	0.920	0.959
RMSEA	0.021	0.020	0.025	0.018
AIC	114187.877	113622.208	60132.129	59510.940
BIC	115117.387	114556.883	60658.851	60058.318
SRMR	0.079	0.075	0.076	0.062
Chi-square difference test	147.052, $df = 1$, $p < 0.001$		420.149, $df = 4$, $p < 0.001$	
Fit indices for model 3 in multigroup CFA [†]	Both		Boys	
			Girls	
CFI	0.963		0.957	
TLI	0.961		0.954	
RMSEA	0.018		0.019	
AIC	59297.989		28226.432	
BIC	59829.876		28683.685	
SRMR	0.060		0.076	

AIC = Akaike information criterion; BIC = Bayesian information criterion; CFA = confirmatory factor analyses; CFI = comparative fit index; df = degrees of freedom; model 1 = one factor for all; model 2a = two factors (internalization and externalization); model 2b = one factor for internalization; model 3 = four first-order factors, one second-order factor for internalization; RMSEA = root mean square error approximation; SRMR = standardized root mean square residual; TLI = Tucker-Lewis index.

[†] Removing item 45 and changing item 47 from anxiety/depression to emotionally reactive.

Item 45 (“Nausea, feels sick without a medical cause”) was endorsed by only 5% of the respondents (6% of the boys and 4% of the girls). Therefore, this item was removed, which improved model fit, decreasing from an estimated SRMR of 0.062 to 0.060 (Table 2).

The multigroup analyses led to the non-rejection of configural invariance and the non-rejection of metric invariance. Regarding configural invariance, despite a significant chi-square test, the model had a good fit, with good CFI (0.958), TLI (0.955), and RMSEA (0.019) as well as an acceptable SRMR (0.072). Therefore, configural invariance was not rejected. The factor loadings of the model combining boys and girls can be seen in Figure 2.

Regarding metric invariance, the assessment of the metric model against the configural model yielded a

Table 2 Factor loadings for the selected model (model 3) in boys and girls

Syndrome/item	Boys	Mean	Girls	Mean
Anxiety/depression				
CBCL10	0.413	0.588	0.491	0.603
CBCL33	0.572		0.675	
CBCL37	0.589		0.491	
CBCL43	0.857		0.801	
CBCL68	0.454		0.476	
CBCL87	0.480		0.533	
CBCL90	0.749		0.755	
Emotionally reactive				
CBCL21	0.686	0.635	0.668	0.627
CBCL46	0.526		0.377	
CBCL47	0.748		0.685	
CBCL51	0.698		0.785	
CBCL79	0.561		0.605	
CBCL82	0.803		0.743	
CBCL83	0.655		0.720	
CBCL92	0.596		0.455	
CBCL97	0.624		0.694	
CBCL99	0.448		0.535	
Withdrawn				
CBCL2	0.476	0.604	0.375	0.549
CBCL4	0.460		0.413	
CBCL23	0.572		0.497	
CBCL62	0.584		0.465	
CBCL67	0.687		0.716	
CBCL70	0.659		0.624	
CBCL71	0.732		0.616	
CBCL98	0.665		0.688	
Somatic complaints				
CBCL1	0.480	0.509	0.633	0.572
CBCL12	0.331		0.454	
CBCL19	0.462		0.495	
CBCL24	0.402		0.375	
CBCL39	0.474		0.625	
CBCL52	0.674		0.726	
CBCL78	0.731		0.805	
CBCL93	0.515		0.462	
Internalization				
Anxiety/depression	0.956		0.953	
Emotionally reactive	0.952		0.993	
Withdraw	0.802		0.794	
Somatic complaints	0.631		0.635	

CBCL = Child Behavioral Checklist.

significant chi-square (44.045, $df = 29$, $p = 0.036$) and an acceptable fit (SRMR = 0.074). Despite the significant chi-square, upon exploring the factor loadings for each item in boys and girls, they were considerably similar (Table 2), and metric invariance was not rejected. However, due to the removal of items throughout the analyses, only partial measurement invariance was attained.²⁹ Then, fit tests for the final model were estimated. Although the differences in factor loadings on boys and girls were nonsignificant, the scores were calculated separately for boys and girls to give the most accurate results by sex.

We hypothesized that the factor somatic complaints would have a higher correlation with the other factors in girls than in boys, and the factor withdrawn would have a higher correlation with the other factors in boys than in girls. However, although the correlation estimate of somatic complaints with the factors was indeed higher in girls than in boys ($r = 0.635$ vs. $r = 0.631$) and the correlation estimate of withdrawn with the factors was higher in boys than in girls ($r = 0.802$ vs. $r = 0.794$), the estimates were only marginally different. Thus, there is not sufficient evidence to support the hypotheses.

Regarding the frequency of the problems, in both boys and girls, the most frequently endorsed items (i.e., answered with “sometimes” or “always”) were “Feelings are easily hurt” (74% of girls and 70% of boys) and “Sulks easily” (64% of boys and girls). The least endorsed symptom was “Is unhappy, sad or depressed” (4.5% of girls and 6.4% of boys). The most frequently endorsed factor in boys and girls was anxious/depressed; items for this factor were endorsed on average by 43% of the girls and 42% of the boys.

In both boys and girls, the factors anxious/depressed and emotionally reactive were highly correlated ($r = 0.95$). In girls, the syndrome emotionally reactive had the highest mean loadings (0.63), the highest mean correlation with all factors ($r = 0.84$), the highest correlation ($r = 0.99$) with the second-order factor, and explained the most variance in the model. In boys, the highest mean loadings were for emotionally reactive and for withdrawn (Table 2), but correlation estimates with the other factors were higher for the anxious/depressed syndrome.

The scores were calculated based on factor loadings separately for boys and girls to give the most accurate estimates by sex. If all items were endorsed with always true, the maximum possible score would have been 33.30 among girls and 32.84 among boys. To obtain comparable scores for both boys and girls, percentages of the maximum possible scores were obtained. The average in percentage points for internalizing problems was 19.42 in girls (Q1: 10.3, median: 16.9, Q3: 26.23) and 19.18 in boys (Q1: 9.81, median: 15.70, Q3: 25.09). Based on the third quartile (Q3) estimates, a score of 26.2 points or higher was considered high for girls and a score of 25.1 points or higher was considered high for boys.

As seen in Table 3, regarding each syndrome, girls had significantly higher estimates on somatic complaints (13.2 vs. 11.2) and anxious/depressed (26 vs. 23.9) than boys, whereas boys had significantly higher estimates on the withdrawn (13.2 vs. 10.8) syndrome.

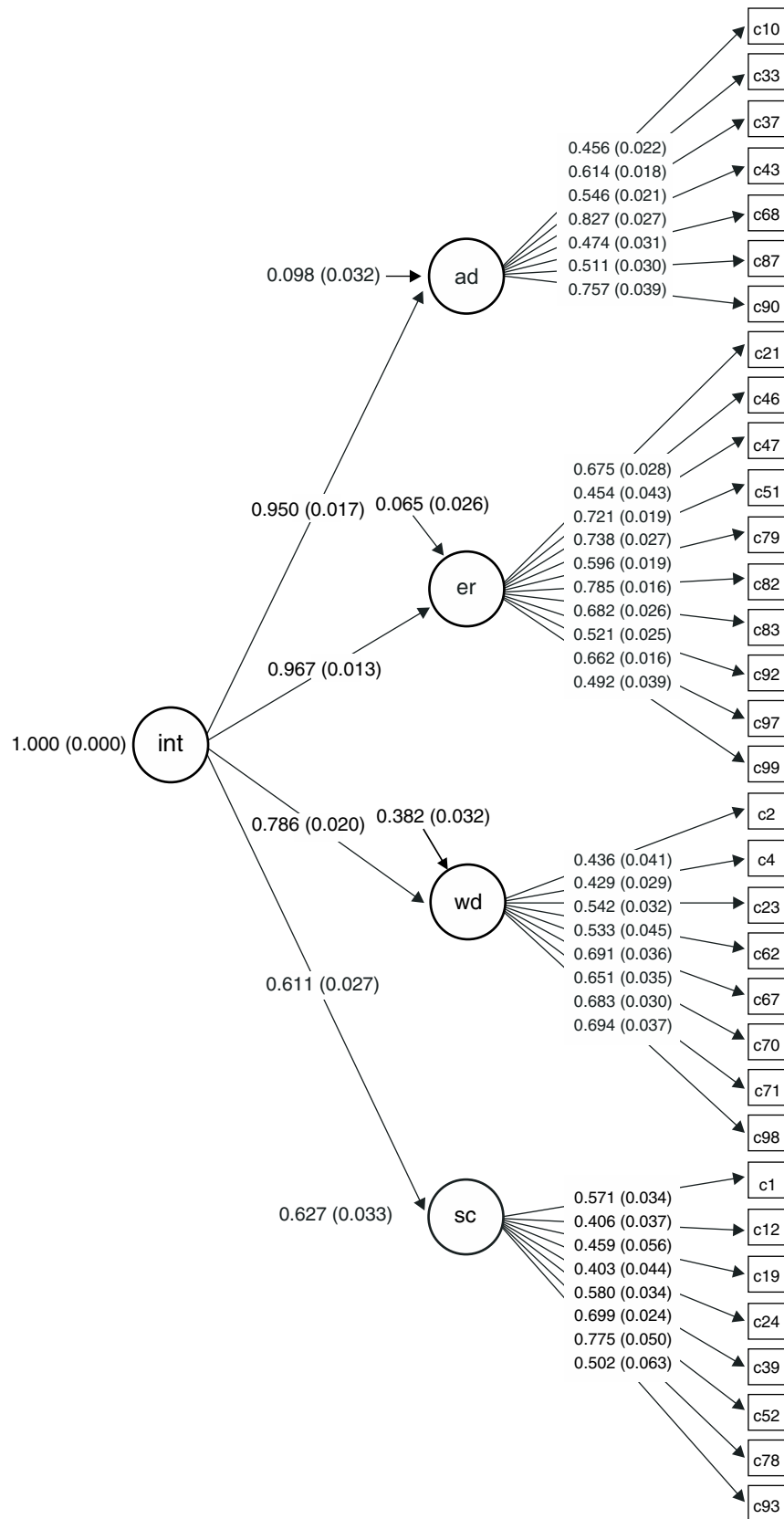


Figure 2 Factor loadings for the selected model combining boys and girls. ad = anxiety/depression; er = emotionally reactive; int = internalizing; sc = somatic complaints; wd = withdrawn.

Table 3 Mean internalization scores by sex in preschoolers

Syndrome	Girls	Boys	t (df)	p-value
Anxiety/depression	26.03	23.89	6.09 (29)	< 0.001
Emotionally reactive	22.89	23.34	-1.31 (29)	0.200
Withdrawn	10.76	13.19	-10.94 (29)	< 0.001
Somatic complaints	13.22	11.20	8.14 (29)	< 0.001
Total internalizing problems	19.42	19.18	0.97 (29)	0.339

df = degrees of freedom.

Discussion

The selected model of the structure of internalizing problems in our sample, a hierarchical model with four first-order factors or syndromes (anxious/depressed, somatic complaints, withdrawn and emotionally reactive) and one second-order factor (internalizing problems), is in line with the proposed structure of the CBCL 1.5-5. It differs from p-factor proponents, who recommended that all behavioral problems be considered only one factor of psychopathology (38-4). However, externalizing symptoms should always be measured when studying internalizing symptoms, since they are distinct problems that are highly correlated.

Analyzing each syndrome rather than using a single score for internalizing problems provides a better understanding of the construct of internalizing problems in these children. In this study, two highly correlated syndromes, anxious/depressed and emotionally reactive, explained most of the variance of the internalizing problems in preschoolers. Hence, the two syndromes are the core of internalizing problems in these children, and their prevention and treatment should be prioritized over the syndromes of withdrawn and somatic complaints, since these syndromes could be reduced when anxiety/depressive problems are treated. Moreover, if, for example, a child tends to be withdrawn, it is essential first to identify if they could also have anxiety/depressive problems instead of trying to find activities that could be more appealing to the child. Common recommendations such as validating the child’s emotions and strengthening the bond with caregivers and teachers³² could also result in the child becoming more engaged in activities and more affectionate toward others. However, since not all problems arise from the same cause, it is suggested that future studies examine the potential predictors of internalizing problems in these children to plan for more effective prevention and treatment efforts.

The CFA was a useful method to identify items that did not explain much of the construct of internalizing problems, namely “Can’t stand things out of place” and “Too concerned with cleanliness and neatness,” which explained less than 7% of the internalizing factor and led to worse fit estimates in all tested models. These two symptoms refer to obsessive thoughts or compulsions rather than somatic complaints, as described in the CBCL 1.5-5 guidelines.¹⁰ Studies on the CBCL 4-18³³ include these questions in screenings for obsessive-compulsive disorder (OCD). It is recommended that, for more accurate estimates of internalizing problems, these two questions be excluded from the analysis, since they could

be part of a separate construct. In addition, there could be low endorsement because the onset of problematic OCD-related symptoms usually occurs after age 16 years.³⁴

Configural invariance was not rejected because estimates showed that the hierarchical model is a good fit for internalizing problems for boys and girls. However, upon exploration, some correlations were problematic in boys and deleting or moving an item was necessary to ensure an appropriate fit of the model for both groups. For researchers interested in this topic, we recommend examining the correlation of the items in the anxious/depressed and emotionally reactive syndromes; the two were only marginally distinguishable in the analyses, especially in boys.

Metric invariance was also not rejected because the model had appropriate fit indices in boys and girls, implying that the internalizing syndromes have the same structure in boys and girls. Since the focus of this study was to increase the understanding of the structure of internalizing problems by sex, we accounted even for minimal differences, calculating the scores from factor loadings separately for boys and girls. However, since the differences in the estimates of boys and girls were negligible, we recommend that scores continue to be calculated as recommended in the CBCL 1.5-5 guidelines,¹⁰ not stratifying by sex. The calculated scores were standardized to make them comparable and were only marginally different in girls versus boys (19.42 vs. 19.18). There is no evidence in this sample to suggest that the disparate rates of total internalizing problems between females and males seen in adolescents and adults are observed from an early age.

There is insufficient evidence to indicate that the withdrawn factor has a more significant association with total internalizing problems in boys than in girls, or that the somatic complaints factor has a more significant association with total internalizing problems in girls than in boys. However, the mean scores in boys and girls in the two syndromes were significantly different by more than two points. These findings suggest that there are more boys than girls with a clinically significant withdrawn syndrome and more girls than boys with a clinically significant somatic complaints syndrome. Interestingly, these findings, along with the differences in anxious/depressed (significantly higher in girls than in boys), indicate that boys and girls can have disparate estimates of specific internalizing syndromes even in preschool years, when there have been fewer exposures to gender norms than in adolescence and adulthood³⁵ and they have not yet been through puberty. Thus, the differences in the prevalence of syndromes in earlier years may lead to

different psychopathology risks and related conditions in later years through different pathways.

Since it was estimated that preschool girls score higher than boys in anxious/depressed and somatic complaints, it is possible that a lack of focus on these problematic behaviors has led to their normalization in girls (it is usually expected that females be “more emotional” than males,³⁶ even in childhood and adolescence). Females could consequently learn inappropriate coping skills when facing stressful situations, which summed to the expectation that they juggle several responsibilities at home and at work³⁷ may lead to higher rates than males of anxiety/depression and their related conditions (such as fibromyalgia³⁸ and severe gastrointestinal symptoms³⁹) in later years.

Since it was estimated that preschool boys score higher than girls in the withdrawn syndrome, boys may start learning to hide their feelings and appear stronger from an early age.³⁶ In adolescence and adulthood, males are less likely than females to analyze their emotions or seek treatment.²² They are expected by society at large, particularly less gender-egalitarian societies, to be more competitive and successful than their female counterparts. Because of this, they could either be more reactive when they underachieve³⁶ or disregard problems until they become unmanageable, potentially increasing their suicide risk compared to females.⁴⁰

The different prevalence of the internalizing syndromes in preschool years suggests there risk patterns could also differ between boys and girls. Exploring the effect of high-risk exposures on internalizing syndromes by sex in preschoolers and assessing confounding in these associations is needed. Analysis of confounding variables (such as child externalizing symptomatology, food insecurity, parental substance use, neighborhood disadvantage, parental social support, and neighborhood violence) is key to understanding what could have led to the difference in estimates between boys and girls.

The main limitations of this study, like those of most similar research, arise from the use of caregiver-reported data. Since the CBCL 1.5-5 scores are based on a third-person report, potential recall problems could have led to overestimation or underestimation of symptom severity. Caregivers may have overestimated the severity of symptoms because they overprotected the child or because they were depressed themselves, or underestimated the severity of symptoms and underreported violent behaviors or neglect to avoid raising questions regarding their parenting skills. Moreover, regarding sex differences in the syndromes, parents or caregivers could have gender expectations of their own that bias their perceptions of gender-coded behaviors, leading to overestimation of these differences.

Prospective studies where parents can record their children's behaviors and these children then be observed by an independent evaluator, small clinical studies with comprehensive assessments, or studies with reports by more than one caregiver are all potential solutions to this limitation.

Another contribution of this study, based on the non-rejection of configural and metric invariance between boys and girls and the acceptable fit of the four-

syndromes model of internalizing problems before adjustments (SRMR = 0.077), is the validation of the internalizing structure of the CBCL 1.5-5. Since all the models tested had good fit indices, the CBCL 1.5-5 can be considered valid for assessment of internalizing problems, at least in São Paulo. The CBCL 1.5-5, in its original form, can be administered in surveys mainly to maintain comparability with studies in other populations and regions. However, for examination of causal pathways, in clinical work, and in other settings where accuracy and understanding of the nuances of internalizing problems are key, we suggest implementation of the modifications made herein, including the deletion and reorganization of some items. Additionally, our suggested considerations regarding gender differences among preschool children should be taken into account.

Regarding the strengths of this study, one major strength was the use of the PreK Survey. Since it was a representative sample of children aged 4 to 5 years attending public preschool in an urban region of Brazil, the findings suggest a similar factor structure of internalizing problems would be relevant to other cities in São Paulo and elsewhere in the country due to a common culture and language. The study had a large sample appropriate for CFA, and there were no missing data. The current study innovated in its approach by placing the focus on internalizing problems of preschoolers, examining in detail the construct of internalization and the related syndromes and symptoms, and by exploring sex differences to aid in the understanding of these problems and inform treatment and prevention efforts in early childhood.

Acknowledgements

The authors thank the participants of the PreK Survey. This study is adapted from AMM's doctoral dissertation, and all other pertinent acknowledgments are duly incorporated therein.

Disclosure

The authors report no conflicts of interest.

Author contributions

AMM: Conceptualization, Data curation, Formal analysis, Funding acquisition, Investigation, Methodology, Project administration, Resources, Software, Supervision, Validation, Visualization, Writing – original draft, Writing – review & editing.

SC: Conceptualization, Supervision, Writing – original draft, Writing – review & editing.

MR: Validation, Visualization, Writing – original draft, Writing – review & editing.

AR: Validation, Visualization, Writing – original draft, Writing – review & editing.

PO: Validation, Visualization, Writing – original draft, Writing – review & editing.

TF: Validation, Visualization, Writing – original draft, Writing – review & editing.

PS: Validation, Visualization, Writing – original draft, Writing – review & editing.

LS: Validation, Visualization, Writing – original draft, Writing – review & editing.

SM: Conceptualization, Data curation, Funding acquisition, Investigation, Methodology, Project administration, Resources, Software, Supervision, Validation, Visualization, Writing – original draft, Writing – review & editing. All authors have read and approved of the final version to be published.

Handling Editor: Giovanni Salum

References

- World Health Organization (WHO). Depression and other common mental disorders: global health estimates. Geneva: WHO; 2017.
- Carpenter AL, Puliafico AC, Kurtz SM, Pincus DB, Comer JS. Extending parent-child interaction therapy for early childhood internalizing problems: new advances for an overlooked population. *Clin Child Fam Psychol Rev*. 2014;17:340-56.
- Berg-Nielsen TS, Solheim E, Belsky J, Wichstrom L. Preschoolers' psychosocial problems: In the eyes of the beholder? Adding teacher characteristics as determinants of discrepant parent-teacher reports. *Child Psychiatry Hum Dev*. 2012;43:393-413.
- Matijasevich A, Murray E, Stein A, Anselmi L, Menezes AM, Santos IS, et al. Increase in child behavior problems among urban Brazilian 4-year olds: 1993 and 2004 Pelotas birth cohorts. *J Child Psychol Psychiatry*. 2014;55:1125-34.
- Ashford J, Smit F, van Lier PA, Cuijpers P, Koot HM. Early risk indicators of internalizing problems in late childhood: a 9-year longitudinal study. *J Child Psychol Psychiatry*. 2008;49:774-80.
- Slemming K, Sørensen MJ, Thomsen PH, Obel C, Henriksen TB, Linnet KM. The association between preschool behavioural problems and internalizing difficulties at age 10–12 years. *Eur Child Adolesc Psychiatry*. 2010;19:787-95.
- Rafii-Ferreira R, Ferreira de Mattos Silveiras E, Pires MLN, Assumpção FB Jr, de Moura CB. Sleep and behavior in children assisted in a psychology service. *Psicol Teor Prat*. 2016;18:159-72.
- Lins T, Alvarenga P. Controle psicológico materno e problemas internalizantes em pré-escolares. *Psic Teor Pesq*. 2015;31:311-9.
- Pearson RM, Culpin I, Loret de Mola C, Matijasevich A, Santos IS, Horta BL, et al. Grandmothers' mental health is associated with grandchildren's emotional and behavioral development: a three-generation prospective study in Brazil. *BMC Psychiatry*. 2019;19:184.
- Achenbach TM, Rescorla LA. Manual for the ASEBA preschool forms and profiles. Burlington, VT: University of Vermont; 2000.
- Dos Santos LM, Queirós FC, Barreto ML, Dos Santos DN. Prevalence of behavior problems and associated factors in preschool children from the city of Salvador, state of Bahia, Brazil. *Braz J Psychiatry*. 2016;38:46-52.
- Ivanova MY, Achenbach TM, Rescorla LA, Harder VS, Ang RP, Bilenberg N, et al. Preschool psychopathology reported by parents in 23 societies: testing the seven-syndrome model of the child behavior checklist for ages 1.5-5. *J Am Acad Child Adolesc Psychiatry*. 2010;49:1215-24.
- Lecannelier F, Ewert JCP, Groissman S, Gallardo D, Bardet AM, Bascuñan A, et al. Validación del Inventario de Conductas Infantiles para niños de entre 1½-5 años (CBCL 1½-5) en la Ciudad de Santiago. *Univ Psychol*. 2014;13:491-500.
- Albores-Gallo L, Hernández-Guzmán L, Hasfura-Buenaga C, Navarro-Luna E. Consistencia interna y validez de criterio de la versión mexicana del child behavior checklist 1.5-5 (CBCL/1.5-5). *Rev Chil Pediatr*. 2016;87:455-62.
- Mahler C. The influence of model components and misspecification type on the performance of the Comparative Fit Index (CFI) and the Root Mean Square Error of Approximation (RMSEA) in structural equation modeling. *Calgary: Graduate Studies*; 2016.
- Martel MM, Pan PM, Hoffmann MS, Gadelha A, do Rosário MC, Mari JJ, et al. A general psychopathology factor (P factor) in children: structural model analysis and external validation through familial risk and child global executive function. *J Abnorm Psychol*. 2017;126:137-48.
- Caspi A, Houts RM, Belsky DW, Goldman-Mellor SJ, Harrington H, Israel S, et al. The p factor: one general psychopathology factor in the structure of psychiatric disorders? *Clin Psychol Sci*. 2014;2:119-37.
- Starr LR, Conway CC, Hammen C, Brennan PA. Transdiagnostic and disorder-specific models of intergenerational transmission of internalizing pathology. *Psychol Med*. 2014;44:161-72.
- Bordin IA, Rocha MM, Paula CS, Teixeira MCTV, Achenbach TM, Rescorla LA, et al. Child Behavior Checklist (CBCL), Youth Self-Report (YSR) and Teacher's Report Form (TRF): an overview of the development of the original and Brazilian versions. *Cad Saude Publica*. 2013;29:13-28.
- Rocha M, Rescorla L, Emerich DR, Silveiras EFM, Borsa JC, Araújo LGS, et al. Behavioural/emotional problems in Brazilian children: findings from parents' reports on the Child Behavior Checklist. *Epidemiol Psychiatr Sci*. 2013;22:329-38.
- Kuehner C. Why is depression more common among women than among men? *Lancet Psychiatry*. 2017;4:146-58.
- Clement S, Schauman O, Graham T, Maggioni F, Evans-Lacko S, Bezborodovs N, et al. What is the impact of mental health-related stigma on help-seeking? A systematic review of quantitative and qualitative studies. *Psychol Med*. 2015;45:11-27.
- Canavan C, West J, Card T. The epidemiology of irritable bowel syndrome. *Clin Epidemiol*. 2014;6:71-80.
- Rosenfield S, Mouzon D. Gender and mental health. In: Aneshensel CS, Phelan JC, Bierman A. Handbook of the sociology of mental health. 2nd ed. Berlin: Springer; 2013. p. 277-96.
- Aina OE, Cameron PA. Why does gender matter? Counteracting stereotypes with young children. *Dimens Early Child*. 2011;39:11-9.
- Caetano SC, Martins SS, Askari M, Surkan P, Sanchez ZM, Rosario MC, et al. The Embu Das Artes, Brazil, preschool mental health study: an epidemiological study of mental health and traumatic events: aims and methods. *JAACAP Open*. 2016;55:S116.
- Caetano SC, Ribeiro MVV, Askari MS, Sanchez ZM, do Rosário MC, Perissinoto J, et al. An epidemiological study of childhood development in an urban setting in Brazil. *Braz J Psychiatry*. 2021;43:43-54.
- Kline RB. Principles and practice of structural equation modeling: Guilford publications; 2015.
- Putnick DL, Bornstein MH. Measurement invariance conventions and reporting: the state of the art and future directions for psychological research. *Dev Rev*. 2016;41:71-90.
- Muthén L, Muthén B. Statistical analysis with latent variables using Mplus. Los Angeles, CA: Muthén & Muthén; 2007.
- SAS Help Center. SAS/ACCESS® 9.4 Interface to ADABAS: Reference [Internet]. 2013 [cited 2024 Sep 03]. documentation.sas.com/doc/en/pgmsascdc/9.4_3.5/acadbas/titlepage.htm
- Winston R, Chicot R. The importance of early bonding on the long-term mental health and resilience of children. *London J Prim Care (Abingdon)*. 2016;8:12-4.
- Andersen PAS, Bilenberg N. Comparison of child behavior checklist subscales in screening for obsessive-compulsive disorder. *Dan Med J*. 2012;59:A4523.
- Millet B, Kochman F, Gallarda T, Krebs MO, Demonfaucon F, Barrot I, et al. Phenomenological and comorbid features associated in obsessive-compulsive disorder: influence of age of onset. *J Affect Disord*. 2004;79:241-6.
- Rohde P, Lewinsohn PM, Klein DN, Seeley JR, Gau JM. Key characteristics of major depressive disorder occurring in childhood, adolescence, emerging adulthood, and adulthood. *Clin Psychol Sci*. 2013;1:41-53.
- Gavranidou M, Rosner R. The weaker sex? Gender and post-traumatic stress disorder. *Depress Anxiety*. 2003;17:130-9.
- Ramados K, Rajadhyaksha U. Gender differences in commitment to roles, work-family conflict and social support. *J Soc Sci*. 2012;33:227-33.
- Furness PJ, Vogt K, Ashe S, Taylor S, Haywood-Small S, Lawson K. What causes fibromyalgia? An online survey of patient perspectives. *Health Psychol Open*. 2018;5:2055102918802683.
- Windgassen S, Moss-Morris R, Goldsmith K, Chalder T. Key mechanisms of cognitive behavioural therapy in irritable bowel syndrome: the importance of gastrointestinal related cognitions, behaviours and general anxiety. *J Psychosom Res*. 2019;118:73-82.
- Martini M, da Fonseca RC, de Sousa MH, de Azambuja Farias C, de Azevedo Cardoso T, Kunz M, et al. Age and sex trends for suicide in Brazil between 2000 and 2016. *Soc Psychiatry Psychiatr Epidemiol*. 2019;54:857-60.