RESEARCH ARTICLE   **Open Access**

Psychometric properties of the Brief Problem Monitor (BPM) in children with internalizing symptoms: examining baseline data from a national randomized controlled intervention study

Marit Løtveit Pedersen[1*] , Thomas Jozefiak[1] , Anne Mari Sund[1,2] , Solveig Holen[3] , Simon-Peter Neumer[3] , Kristin D. Martinsen[3] , Lene Mari P. Rasmussen[4] , Joshua Patras[4] and Stian Lydersen[1]

Abstract

Background: Prevention is essential to reduce the development of symptomology among children and adolescents into disorders, thereby improving public health and reducing costs. Therefore, easily administered screening and early assessment methods with good reliability and validity are necessary to effectively identify children's functioning and how these develop. The Brief Problem Monitor (BPM) is an instrument designed for this purpose. This study examined the psychometric properties of the Norwegian version of the BPM parent (BPM-P) and teacher (BPM-T) versions, including internal reliability and construct validity at assessing children with internalizing problems.

Methods: Baseline data were collected from a national randomized controlled intervention study. Children aged 8–12 years with self-reported symptoms of anxiety and/or depression with one standard deviation above a chosen population's mean were included in this study. Teachers (n = 750) and parents (n = 596) rated children using the BPM-T and BPM-P, respectively. Internal consistency was measured using Cronbach's alpha, and multi-informant agreement between the BPM-P and BPM-T was measured using Spearman's correlations. Construct validity was assessed via confirmatory factor analysis.

Results: Internal consistency was good throughout all domains for both the BPM-P and BPM-T, with a Cronbach's alpha ranging from .763 to .878. Multi-informant agreement between the parents and the teacher was moderate on the externalizing, attention, and total scales and low on the internalizing scale. The model fit for the three-factor structure of the BPM was excellent for the BPM-P and good for the BPM-T.

Conclusions: Internal consistency was good, and the original three-factor solution of the BPM-P and BPM-T was confirmed based on our sample of school children at-risk for emotional problems. These promising results indicate that the BPM may be a valid short assessment tool for measuring attentional, behavioral, and internalizing problems in children.

*Correspondence: marit.l.pedersen@ntnu.no
[1] Regional Center for Child and Youth Mental Health and Child Welfare,
Department of Mental Health, Norwegian University of Science
and Technology, Trondheim, Norway
Full list of author information is available at the end of the article

Background

Symptoms of anxiety and depression [internalizing problems] are among the most common psychological difficulties diagnosed in children and adolescents. Approximately seven percent of children from population-based samples in Norway present symptoms that are compatible with a mental disorder [1, 2]. Similarly, international studies have found a prevalence rate of mental health disorders from 7 to 23% [3–6]. Anxiety and depression often co-occur, and anxiety often precedes depression [7, 8]. Depressive problems affect youth negatively in different life domains (e.g., lower academic achievement, more peer and family problems) [4, 9, 10], and anxious youth are at greater risk for absenteeism, academic underachievement, low social acceptance, and impaired psychosocial functioning [11–13]. Previous research also indicates that children with symptoms of anxiety and/or depression clearly experience a reduction in their daily functioning, even though they do not qualify for a full diagnosis [14, 15]. Such negative outcomes can cause serious health consequences and costs, for the youth, his or her family, and the society at large [16]. Thus, it is important to prevent the development of anxiety and depression in children and adolescents. Having valid and reliable instruments to identify these children is, therefore, of utmost importance.

Both in research and clinical care, assessments of intervention progress and outcome that are quick, easily administered, valid, and reliable are needed [17] so that response to the intervention and possible adjustments to the intervention processes can be applied. Advantageously, the assessment method is suitable for children with comorbid conditions (e.g., anxiety, depression, conduct problems) and various informants. To be useful, a short survey should adequately assess progress and outcomes, accommodate comorbidity, take little time to administer, and show satisfactory psychometric properties. *The Brief Problem Monitor* (BPM) instrument is designed for this purpose [18]. This measure provides a uniform problem scale to assess attentional, behavioral, and internalizing problems in children and adolescents aged from 6 to 18 years. The BPM forms are developed from the longer corresponding versions of the ASEBA *Child Behavior Checklist for those aged* 6–18 *years* (CBCL/6–18), *Teachers' Report Form for children those aged* 6–18 (TRF), and *Youth Self-report for those aged* 11–18 *years* (YSR) [19]. The ASEBA long forms are widely used instruments for clinical and research psychopathology. They have for decades provided information from various informants and shown good psychometric properties in studies conducted in different countries [19, 20]. To date, the literature supporting the psychometric properties of the Brief Problem Monitor (BPM) measures is scarce. In the present study, we examined the psychometric properties of the Norwegian versions of the BPM for parents (BPM-P) and teachers (BPM-T).

Development of the BPM survey

The development of the BPM started with Chorpita and colleagues' Brief Problem Checklist interview (BPC), a brief 12-item index derived from YSR and CBCL, meant to be an easily administered, time-saving, and clinically relevant measure [17]. The index included internalizing (INT) and externalizing problems (EXT). Despite the reduced number of items in the subscales, internal consistency, test–retest reliability, and convergent and discriminant validity were considered good [17].

Achenbach et al. [18] expanded the index further and added an additional third scale in the BPM, assessing attention and hyperactive symptoms (ATT). Achenbach also developed an assessment for teachers based on the TRF. The final version of the BPM consists of one form for parents (BPM-P; 19 items), one form for adolescents (BPM-Y; 19 items), and one form for teachers (BPM-T; 18 items) [18]. The items are distributed on three subscales: attention/hyperactivity problems (6 items), externalizing problems (6 items in BPM-T and 7 items in BPM-P and BMP-Y), and internalizing problems (6 items).

Psychometric properties of BPM

Studies of the 2001 version of the ASEBA long forms were used to analyze the psychometric properties of the American versions of the BPM [18, 19]. In a US sample of youths, the CBCL- and TRF-forms showed excellent internal consistency [19]. Test–retest (8–16 days apart) yielded similar results. They also reported good criterion-related validity to differentiate between children with and without a diagnosis.

Apart from Achenbach and colleagues [19], there are, to our knowledge, four published studies evaluating the psychometric properties of the BPM forms, see Table 1 [21–24].

According to a systematic review of Scandinavian studies on the psychometric properties of the BPM [25], only

Table 1 Psychometric characteristics of BPM reported in four studies

Study (country)		BPM-P				BPM-T			
		ATT	INT	EXT	Total	ATT	INT	EXT	Total
Richter [21] (Norway)	Type of population	General population				Stratified random subsample			
	Age range (mean)	6–16 (10.6)				6–13 (9.4)			
	Reliability *(Internal consistency (Cronbach's alpha)*	.77	.70	.76	.93	.88	.71	.87	.96
Piper et al. [23] (USA/Canada)	Type of population	Convenience sample				BPM-T not conducted			
	Age range	6–18 (mean age 11.5)							
	Reliability *(Internal consistency (Cronbach's alpha)*	.87	.79	.86	.91				
Penelo et al. [22] (Spain)	Type of population	General population				BPM-T not conducted			
	Age range	6–8							
	Reliability *(Internal consistency (McDonald's Omega)*	.92	.83	.89	.93				
	Validity *Construct RMSEA*	.040[a]/.052[b]/.057[c]				"			
	CFI	.968[a]/.930[b]/.919[c]				"			
	TLI	.963[a]/.920[b]/.907[c]				"			
Rodenacker et al. [24] (Germany)	Type of population	Clinical/general population				Clinical/general population			
	Age range	Mean age 11.5/12.3				Mean age 11.5/12.3			
	Reliability *(Internal consistency (Cronbach's alpha)*	.83/.81	.72/.66	.81/.73	.83 /.83	.76/.85	.76/.74	.86/.87	.81/.87
	Validity *Construct RMSEA*	.077/.045				.119/.102			
	CFI	.920/.950				.890/.929			
	TLI	.906/.941				.872/.918			

BPM-P = Brief Problem Monitor–Parents; BPM-T = Brief Problem Monitor–Teacher; ATT = attention; EXT = externalizing problems; INT = internalizing problems; Total = total problems (ATT + EXT + INT). Validity: Robust weighted least squares estimator: RMSEA = root mean square error of approximation, CFI = comparative fit index, TLI = Tucker Lewis index. Age of the sample: [a] = 6-year-old, [b] = 7-year-old, [c] = 8-year-old

References: Richter [21], Piper et al. [23], Penelo et al. [22], Rodenacker et al. [24]

Pedersen *et al. BMC Psychology*   (2021) 9:185

Page 4 of 12

one study was found [21]. This study by Richter [21] included BPM-P, BPM-T, and BPM-Y in a Norwegian population sample of children ranging in age from 6 to 16 years. The study reported excellent internal consistency for the total scale for all versions according to the European Federation of Psychologists' Association's guidelines (EFPA) [26]. The internal consistency of the internalizing subscale was reported to be adequate for BPM-P and BPM-T. For the remaining two subscales, attention and externalizing problems, the internal consistency was adequate on BPM-P and good on BPM-T. The study also reported good construct and content (convergent) validity.

The second study was performed with an American/Canadian sample; caregivers of children/youths aged 6–18 years completed the CBCL/6–18 online [23]. The 19 items of the BPM-P were analyzed. The internal consistency for the BPM-P total scale was excellent, the attention and externalizing subscales were good, and the internalizing subscale was adequate. The correlation between the full-length CBCL/6–18 and the shorter BPM-P was considered high for the total score and the subscales. BPM-P was sensitive and could identify behavioral and emotional problems among children whose parents reported a psychiatric diagnosis (ADHD, depression, anxiety, autism spectrum disorders, etc.) when compared to the group that had not been diagnosed. However, the study was limited to caregivers, and the findings supported that additional information from other sources, for example, teachers, should be obtained.

The third study was conducted in a community sample of Spanish children aged from 6 to 8 years, where parents answered the CBCL/6–18. Nineteen items of the BPM-P were examined [22]. Internal consistency was good. The subscales showed higher values for attention problems and lower for internalizing problems. The concurrent validity was high with a significant correlation between BPM-P and CBCL/6–18. Construct validity, investigated by confirmatory factor analysis (CFA), showed that the 3-factor model (attention, externalizing, internalizing) was adequate.

The fourth study was conducted in a clinical and a general population-based sample from Germany. Children, parents, and teachers answered the BPM based on the long version of ASEBA from 1991 [24]. Two items on the attention scale were not present ("fails to finish things he/she starts" and "inattentive or easily distracted"), leaving 17 items. The internal consistency was considered adequate-to-good for most of the subscales and the total scale regarding BPM-P and BPM-T in both samples. The subscale of internalizing problems showed inadequate consistency in the general population sample of parents. BPM-P indicated an acceptable three-factor model fit

in the clinical sample and an excellent model fit in the general population sample. Regarding the BPM-T, the teacher-version did not have a satisfactory model fit in the clinical or the general population sample.

Multiple informant differences

A meta-analysis that evaluated the validity of multiple informants assessing child and adolescent mental health problems in 341 studies from 1989 to 2014 reported low-to-moderate cross-informant correspondence (mean internalizing: correlation 0.25; externalizing 0.30; mean overall 0.28) [27]. The meta-analysis indicated higher levels of correspondence when problems were easy to observe (externalizing behavior vs. internalizing problems), and informants came from the same setting (mother and father vs. parent and teacher). This is similar to studies evaluating BPC/BPM/ASEBA long forms, where the cross-informant agreement was low (0.22, 0.31, and 0.19 for the internalizing, externalizing, and total scales, respectively) between *child and parent* [17]. Achenbach [28] found a parent–child correlation of 0.25 and a higher cross-informant correlation for externalizing than internalizing problems across diverse types of cross-informant pairs.

Correlations between *parents and teachers* are somewhat higher in these studies—ranging from 0.38 to 0.44 for the attention scale, 0.32– 0.35 for the externalizing scale, 0.21 for the internalizing scale, and 0.33 for the total scale [19, 29]. All these findings are in accordance with Achenbach's multi-informant assessment approach, i.e., that high cross-informant agreement regarding psychological problems/mental health symptoms is not expected because mental health problems are perceived differently from different perspectives. Furthermore, the problems may only be present in certain settings.

To summarize, studies of the BPM are scarce and previous studies have been performed in clinical or population samples. Children with self-reported symptoms of anxiety and depression are an understudied at-risk population, and there is a call for valid, reliable, and easily administered instruments to assess these symptoms as early as possible. Moreover, the psychometric properties of translated versions of the BPM should be evaluated. The cultural norms and differences between countries are important and could influence the results and usefulness of the scale. The only study from Scandinavia [21] was population-based and did not include any factor analysis of the BPM. Furthermore, the only study including two types of informants and a confirmatory factor analysis did not contain all the BPM`s items [24]. This study aimed to evaluate the psychometric properties of the Norwegian version of the BPM in young children (aged

8–12 years) at-risk of developing anxiety and depression, based on reports from both parents and teachers.

Methods
Procedure
The current study used baseline data from a national randomized controlled intervention study [30] investigating the effectiveness of the intervention *EMOTION, Coping Kids Managing Anxiety and Depression* [31] for children aged 8–12 years. The EMOTION intervention aims to reduce symptoms of anxiety and depression and the likelihood of developing later disorders.

Primary schools from rural, suburban, and urban areas in Norway volunteered to participate in the study. The children in the eligible grades (8–12 years of age) and their parents were informed through age-appropriate information in class and at parental meetings about the study both in writing and orally. Children who considered themselves more anxious or sadder than their peers were invited to participate in the screening procedure. Participation required expressed interest from the child and written parental consent; teachers were also informed about the study. Data were collected electronically from 2014 until 2017 with new children entering every semester. For further information about the randomized controlled trial, see Patras and colleagues [30].

Participants
Children (n = 1692) from 36 primary schools were screened using web-based questionnaires for self-reported symptoms of anxiety and depression, using the Multidimensional Anxiety Scale for Children—MASC-C [32] and the Mood and Feelings Questionnaire-Short Version—SMFQ [33].

Children who reported symptom levels of at least one standard deviation above the population-based mean on measures of anxiety and/or depression were invited to participate in the study (n = 873). These cut-offs were based on national and international studies in the relevant age group [34–36]. Seven of the invited children, who were not expected to benefit from the intervention (having developmental delays, autism, severe behavioral disturbance), were excluded from the study. Due to a lack of resources (i.e., lack of group leaders to implement the intervention), 71 children were also randomly excluded.

Web-based questionnaires were sent to the participating children's parents and teachers approximately two weeks after their screening. Although both parents were encouraged to participate, participation was voluntary; however, the children were invited regardless of whether their parents had answered the questionnaires. The child's primary teacher answered questionnaires about the child. The parents and teachers rated the children on matters concerning attention/hyperactivity, internalizing problems, and externalizing problems now or within the last two weeks, using BPM-P and BPM-T, respectively.

Of the included children, 750 (n = 435, 58% girls) students had a teacher response and were included in the current study. Grade level was used as a proxy for age: Third to sixth grade represented 8–12 years of age. Approximately 4% of the participating children were in the third grade, 36.1% in fourth, 46.1% in fifth, and 13.7% in sixth. Only one parental response per child was analyzed in the present study; 596 children had a response from one of the parents (482 mothers, 80.6%).

Instruments
Brief Problem Monitor (BPM). The BPM-P (19 items) and BPM-T (18 items) has an age range from 6 to 18 years. The instruments include three subscales with six items each: ATT, INT, and EXT. The extra question on the parent version is about disobedient behavior at home. The ATT subscale contains questions like whether the child "can't concentrate or pay attention for long;" or "can't sit still, restless or hyperactive." Within the EXT subscale, there are questions on whether the child "argues a lot" or "has temper tantrums or a hot temper." In the INT subscale, the questions ask if the child is "feeling too fearful or anxious" or "unhappy, sad, or depressed." The items are rated over user-selected rating periods (e.g., 5,7,14,30, and 45 days), and are supposed to describe the child, on a scale ranging from 0 to 2 (0 = *not true*, 1 = *somewhat true*, or 2 = *very true*).

The Norwegian versions of the CBCL, TRF, and YSR were translated and published in 1986/1988, 1993, and 2002, respectively [37]. The Norwegian version of the BPM was based on these translations.

The Multidimensional Anxiety Scale for Children–Child (MASC-C). In this study, children were included using the MASC-C [32] to assess anxiety symptoms, a 39-item self-report questionnaire that assesses anxiety symptoms in children and adolescents aged 8–19 years based on the past two weeks.

The MASC-C has shown favorable psychometric properties both internationally and in Norway [32, 38–40]. The internal consistency of self-reported symptoms of anxiety at baseline in the current study was good, with a Cronbach's alpha of 0.84.

The Mood and Feelings Questionnaire-short version (SMFQ). To assess depressive symptoms, SMFQ [33], a 13-items questionnaire targeting children from 8 to 18 years, was used. The SMFQ assesses cognitive, affective, and behavioral-related symptoms of depression during the last two weeks.

Previous studies indicated good psychometric properties for the Norwegian version of the SMFQ [41, 42].

Pedersen *et al. BMC Psychology*   (2021) 9:185

Page 6 of 12

In this study, the internal consistency of the scale was good, with a Cronbach's alpha = 0.80 at baseline.

Statistics analysis

For the children who had teacher reports (n = 750), 154 parental answers were missing. For the participating parents (n = 596) and the teachers, no missing items were reported.

Reliability was measured by internal consistency (Cronbach's alpha) of the subscales and the total problem scale. According to EFPA [26], the internal consistency is considered excellent if Cronbach's alpha > 0.90, good between 0.80 and 0.90, adequate between 0.70 and 0.79, and inadequate when < 0.70.

The items on BPM-P and BPM-T have only three ordinal response categories (*not true, somewhat true, very true*). The responses were not normally distributed. To assess multi-informant agreement between the subscales on BPM-P and BPM-T, we used Spearman's correlation coefficient. According to Cohen [43], correlation coefficients were considered low between 0.10 and 0.29, moderate between 0.30 and 0.49, and high for 0.50 and above.

To assess construct validity, we conducted a CFA for ordinal categorical variables to confirm the conceptual three-factor model with the three subscales of the BPM-P and BPM-T. Weighted least square estimator with robust standard errors and mean- and variance-adjusted chi-square test statistics (WLSMV) were used as an estimator, owing to the ordinal structure of the data. The robust weighted least squares (WLSMV) estimator with Delta parameterization is recommended for the analysis of skewed categorical data [44]. In this method, error points are accounted for, and the items are interpreted as observable indicators of the non-observable (latent) factor to which they belong [45]. The following goodness-of-fit indices were used: RMSEA, CFI, and TLI. The indices were recommended by Brown [46]: RMSEA < 0.08 and CFI and TLI > 0.90 were considered acceptable fit, and RMSEA < 0.06 and CFI and TLI > 0.95 indicate excellent fit. The factor loadings for each item on the associated subscale were evaluated according to Tabachnick and Fidell's [47] suggestions, where a rating of 0.71 is considered excellent, 0.63 very good, 0.55 good, 0.45 fair, and 0.32 and lower poor.

Two-sided *p*-values < 0.05 were regarded as significant, and we reported 95% confidence intervals (CI) where relevant. CFA analyses were conducted using Mplus 8 [48]. Other analyses were conducted using SPSS 26.0 (IBM, Armonk, NY, USA).

Table 2 Mean scores and standard deviation for the BPM-P and BPM-T

Domains	BPM-P (n = 596)		BPM-T (n = 750)	
	Number of items	Mean (SD)	Number of items	Mean (SD)
ATT	6	2.92 (2.88)	6	3.09 (3.27)
EXT	7	2.26 (2.50)	6	1.30 (2.07)
INT	6	2.58 (2.49)	6	2.57 (2.61)
Total	19	7.76 (6.15)	18	6.96 (6.12)

BPM-P = Brief Problem Monitor–Parents; BPM-T = Brief Problem Monitor–Teacher; ATT = attention; EXT = externalizing problems; INT = internalizing problems; Total = total problems (ATT + EXT + INT); SD = standard deviation

Table 3 Spearman's correlations between BPM-P and BPM-T subscales and total score

BPM-T	BPM-P (n = 596)			
	ATT	EXT	INT	Total
ATT	.451**	.491**	.047	.325**
EXT	.255**	.391**	.089*	.270**
INT	.211**	.186**	.290**	.293**
Total	.409**	.315**	.181**	.384**

BPM-P = Brief Problem Monitor-parents; BPM-T = Brief Problem Monitor-teacher; ATT = attention; EXT = externalizing problems; INT = internalizing problems; Total = total problems (ATT + EXT + INT). *p < .05 (two-tailed), **p < .01 (two-tailed)

Results

Descriptive statistics for the BPM-P and BPM-T subscales and total scores

Table 2 shows the total mean scores and standard deviation for the different subscales and the total score on BPM-P and BPM-T.

Reliability

The BPM-P showed good internal consistency on attention problems (α = 0.834), on externalizing problems (α = 0.805), and the total problems scale (α = 0.871); and adequate internal consistency on internalizing problems (α = 0.763). The BPM-T showed good internal consistency on all the subscales and total problems scale: attention (0.878), externalizing problems (0.805), internalizing problems (0.818), and total problems (0.877).

Multi-informant agreement

The correlations between BPM-P and BPM-T were moderate on the total problem scale (0.384) and the subscales of ATT (0.451) and EXT (0.391). The correlation was low on INT (0.290). See Table 3.

Validity

Construct validity

The CFA analyses produced an excellent model fit for the BPM-P and a good model fit for the BPM-T for the three-factor model. See Table 4. The chi-square statistics were significant, as expected for a large sample size. The χ^2/df-ratio was 2.6845 for the BPM-P and 6.6287 for the BPM-T; below 3 is commonly regarded as acceptable [46, 49, 50].

Further, all items on each subscale contributed significantly ($p < 0.001$) to the three latent constructs—EXT, ATT, and INT—with satisfactory factor loadings. See Figs. 1 and 2.

For the BPM-P, the factors that correlated highest were attention and externalizing problems (ATT–EXT = 0.658). See Fig. 1. The lowest factor correlation was between attention and internalizing problems (ATT–INT = 0.403). The items with the highest loading on the subscales were "Can't concentrate; can't pay attention for long" (ATT = 0.900); "Argues a lot" (EXT = 0.870); and "Unhappy, sad, or depressed" (INT = 0.785).

For the BPM-T, the factors that correlated highest were attention and externalizing problems (0.744). The lowest factor correlation was between attention and internalizing problems (0.390). The items with the highest loading on the subscales were "Can't concentrate; can't pay attention for long" (ATT = 0.925), "Threatens people" (EXT = 0.907), and "Worries" (INT = 0.865). See Fig. 2.

Discussion

This is the first study to present the psychometric properties concerning both the reliability and construct validity of the Norwegian version of the BPM-P and BPM-T, used to assess schoolchildren with internalizing problems. The internal consistency for both versions was good. The multi-informant agreement was moderate to low. The model fit for the three-factor

Table 4 Confirmatory factor analyses of the BPM-P and BPM-T: fit statistics for the model

Fit statistic	BPM-P	BPM-T
χ^2/df	400.928/149 = 2.6845	875.361/132 = 6.6287
RMSEA	0.053 (CI 0.047–0.060)	0.087 (CI 0.081–0.092)
CFI	0.966	0.947
TLI	0.961	0.938

BPM-P = Brief Problem Monitor-parents, BPM-T = Brief Problem Monitor-teacher, Robust weighted least squares estimator, χ^2/df = chi-square relative to its degree of freedom, RMSEA = root mean square error of approximation, CFI = comparative fit index, TLI = Tucker Lewis index

structure of the BPM was confirmed: excellent for the BPM-P and good for the BPM-T.

The internal consistency, assessed by Cronbach's alpha, for the total scale and both the BPM-P and BPM-T were good, and the subscales generally showed good estimates. Hence, internal consistency was higher in the current study than in a representative sample from the general population in Norway [21]. In line with previous studies [21–24], this study also found higher values of attention problems than internalizing problems on both the BPM-P and BPM-T. Attention problems might be expressed through behaviors that are more visible to others, such as failing to finish tasks the child has started, inability to sit still, acting without thinking, etc. [51]. In contrast, internalizing problems exist more within the individual (e.g., feeling worthless, having worries, etc.), which are not so observable by others [52]. Children's self-reports might be better suited to identify these problems, and triangulation of responses will create a better overall picture of children's problem areas [27, 53].
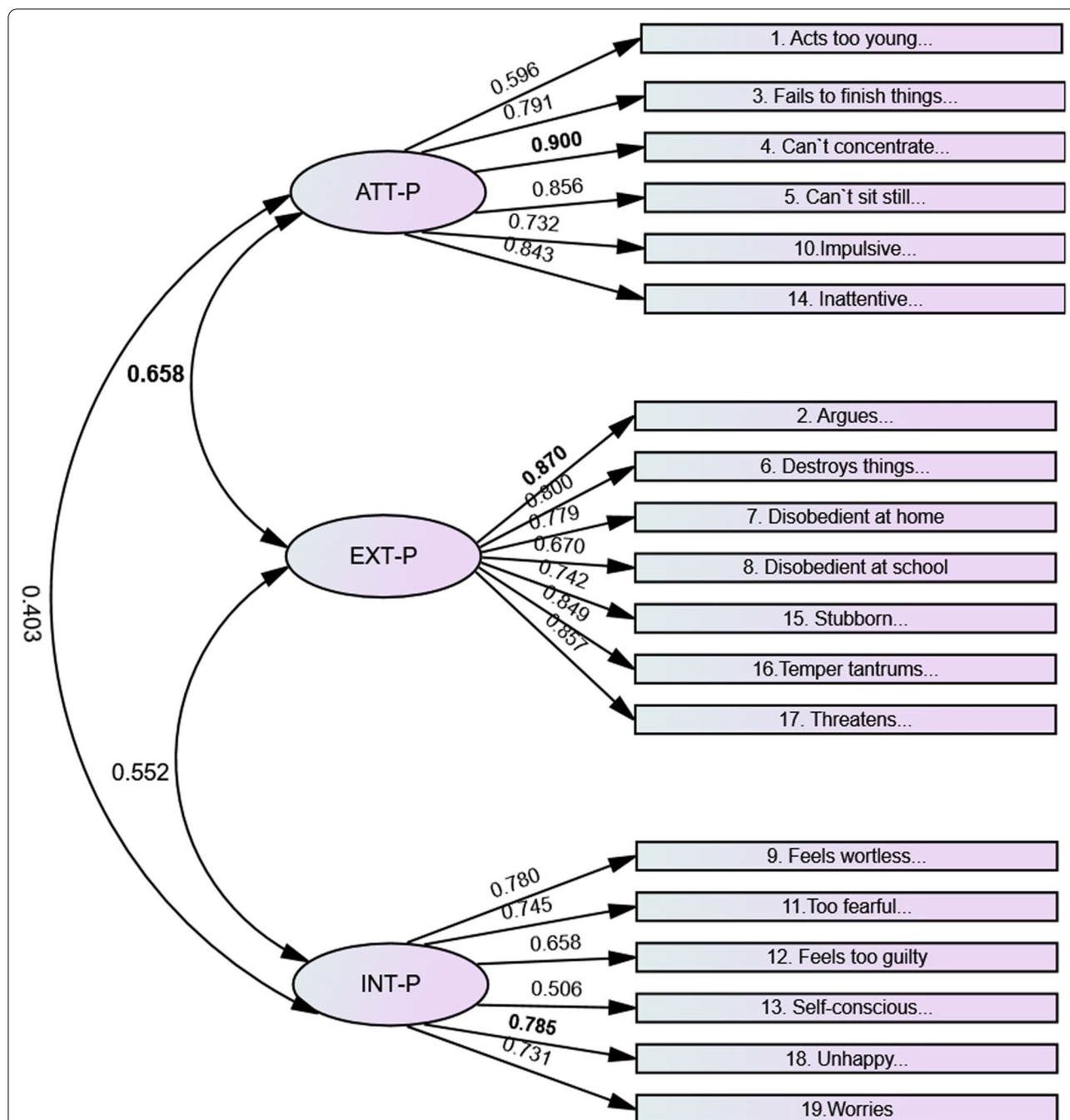
Considering multi-informant perspectives from parents and teachers, we found moderate associations between corresponding subscales of attention, externalizing problems, and the total problem scale, and a weak association between the scorings of internalizing problems. Our findings corroborate previous research indicating that it is more common to agree on externalizing behavioral and attention problems than internalizing problems [19, 27–29]. It is important to underline that in child psychology, the associations between responses from different informants are expected to be low-to-moderate because mental health problems, per se, are perceived from different perspectives and in different environments. Thus, our results are in line with previous studies [19, 27–29].

Concerning the children in our high-scoring at-risk sample, our results showed a better model fit regarding both BPM-P and BPM-T compared to clinical and population-based samples [22, 24]. In the present study, all the items of the BPM were included in the analyses of construct validity, while in a German study, only 17 items were used [24]. The German study also showed that the teacher version did not have a satisfactory model fit in the clinical or the population sample. Our study, however, yielded an acceptable model fit for the teacher version. This indicates that in the school-based population studied, the complete version of the BPM can assess different problems in young school children from teachers' perspectives as well. The results provide an opportunity for future use of the teacher version when examining children's mental health problems to include significant school perspectives. However, more research is needed to confirm our findings. It is crucial to identify children

Fig. 1 BPM-P. Standardized factor loadings and factor correlations. *Note*: BPM-P = Brief Problem Monitor-Parents. ATT = Attention; EXT = Externalizing problems; INT = Internalizing problems; P = parents. All the factors are significant at *p* < 0.01. In bold; highest correlation between factors, and items with the highest factor loading

with mental health issues early to prevent possible negative trajectories leading to anxiety and depression and reduced functioning in different life domains [4, 9, 12, 13]. Therefore, having available instruments that are easily administered, reliable and valid both for clinical and municipal health services are essential. It is also

important to include different perspectives from different areas of the children's lives, such as at home and school. Children can, for example, behave differently at home, where they feel safe. However, in school, which may have more unpredictable surroundings, the same child may struggle. Teachers can observe children's behavior and

Pedersen *et al. BMC Psychology* (2021) 9:185

Page 9 of 12
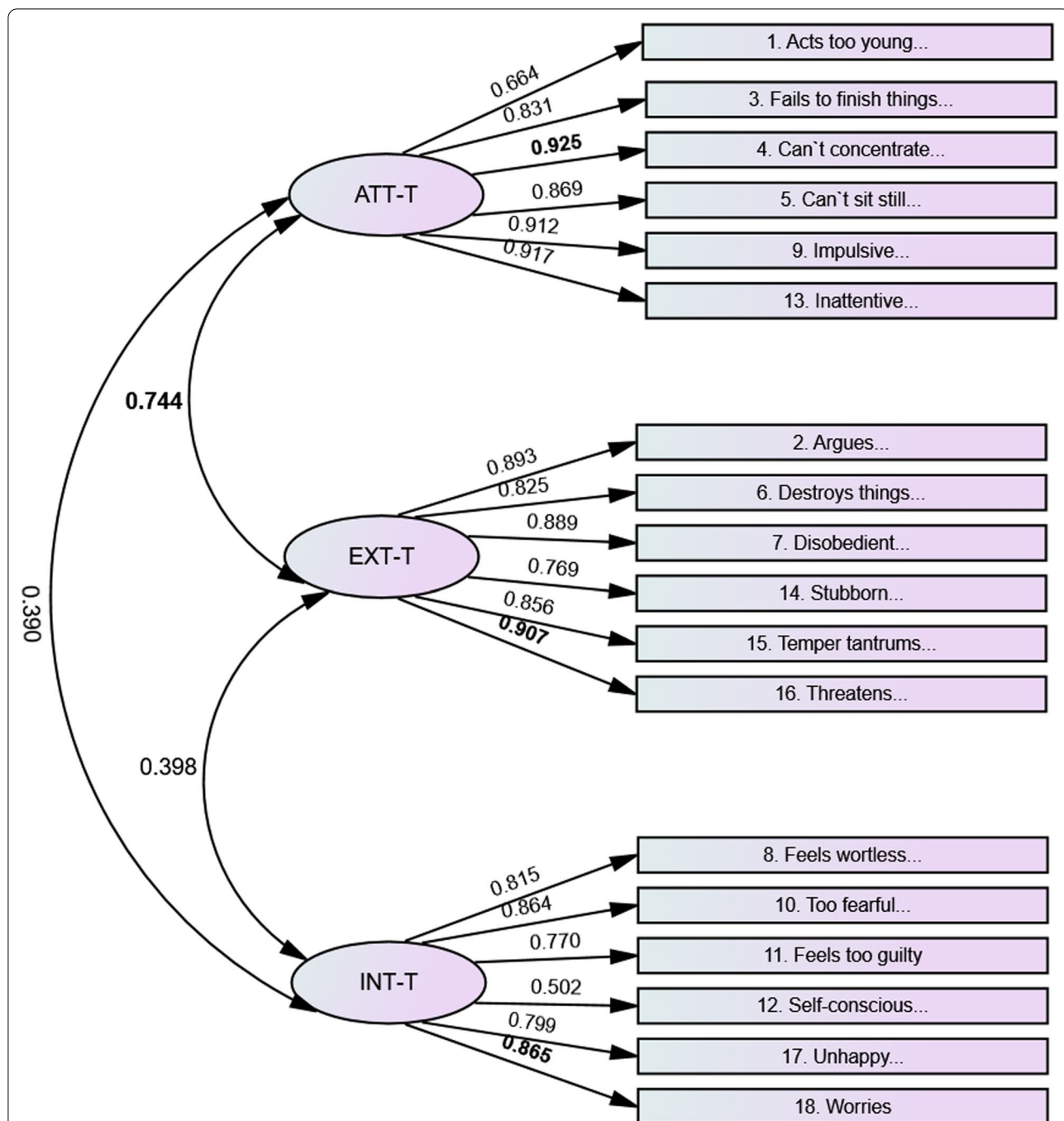
Fig. 2 BPM-T. Standardized factor loadings and factor correlations. *Note*: BPM-T = Brief Problem Monitor-Teacher. ATT = Attention; EXT = Externalizing problems; INT = Internalizing problems; P = parents. All the factors are significant at *p* < 0.01. In bold; highest correlation between factors, and items with the highest factor loading

well-being at school. The positive findings from our study are a much-needed supplement to the research on BPM in an understudied population of at-risk children.

Regarding the factor correlations on both BPM-P and BPM-T, attention and externalizing problems had the highest correlations, whereas attention and internalizing problems showed the lowest factor correlations. This result was not surprising given that externalizing issues may co-occur with attentional problems [51]. Children with internalizing problems may, however,

Pedersen *et al. BMC Psychology*   (2021) 9:185

Page 10 of 12

often show avoidant behavior to their surroundings to cope with their fears [54], which may be more difficult to observe. Furthermore, the problems may be present at school and not at home or vice versa, which may make it even more difficult to identify. When exploring the factor loadings on the subscale of ATT, five of the six items showed excellent values. The exception was the item, "Acts too young for his/her age"; however, it was still considered either good (BPM-P) or very good (BMP-T). This item was included in the BPM based on earlier research indicating high factor loadings in both population and clinical samples [55]. Although this item was within the acceptable range, one might consider that the participating children were young, i.e., aged between 8 and 12 years, which is the lower age range recommended for the use of the BPM [18]. The individual differences may be large in this age group, which might have had an impact on the scoring of this item. The two items, which were not present in the German study [24], contributed significantly to the factor of attention in both the BPM-P and BMP-T in the current study. For the EXT subscale, the factors loaded well on all items for both versions. The subscale INT showed larger differences between the factor loadings in the BPM-P than in the BMP-T. However, four items were considered excellent, and the item "feels too guilty" was considered very good, whereas "self-conscious or easily embarrassed" also reflected good factor loadings. The latter item had between a good and fair factor loading on the BPM-T, while the five other items showed excellent values. Self-consciousness and embarrassment might be more difficult to observe if the child does not express these feelings overtly. Children at this age, and especially children with internalizing problems, may have difficulties expressing their emotions. An alternative explanation could be that increased self-consciousness and embarrassment typically become a challenge at an older age, i.e., around puberty [56]. Nevertheless, when examining all the items in our analysis within each of the subscales, all factor loadings were high and significant. The suggested three-factor model of the BPM was confirmed, indicating that the instrument in our sample assessed what was intended, thus displaying strong construct validity.

The BPM has formerly been psychometrically evaluated in samples from the clinical and general population [21–24]. We add to existing knowledge by evaluating it in an "at-risk sample for internalizing problems". This fits well to its use by school- or community services for children and adolescents to identify internalizing problems among children where one suspects such problems. Moreover, such a short early systematic assessment for children at risk would provide a reliable and valid base to prevent further development of anxiety or depression. It also could be a starting point for further referral to treatment in specialized mental health services.

Strengths and limitations

One strength of this study was that it included a large heterogeneous sample of children from both rural and urban areas. The at-risk population of children presenting elevated symptom levels of anxiety and/or depression represents a sample less often studied than general population samples or clinical samples, which is also a strength of this study. A multi-informant approach with a high response rate, particularly from the teachers, together with sophisticated analyses, strengthened the results.

However, there are some limitations. The study was not initially designed to investigate the psychometric properties of BPM; therefore, data for tests of other instances of reliability (e.g., test–retest, sensitivity for change) and validity (e.g., convergent and divergent validity) were not available.

Conclusion

The reliability of the Norwegian version of the BPM-P and BPM-T showed overall satisfactory internal consistency on all subscales and the total scale. Multi-informant agreement between the parents and the teacher reports were as expected—moderate on the externalizing, attention, and total scales, and low on the internalizing scale. Regarding validity, the original three-factor solution of the parents' and teachers' versions of the BPM was confirmed based on our sample of school children. Owing to the psychometric findings in this study, the BPM may be a valid, quick assessment tool for measuring attentional, behavioral, and internalizing problems in children. Further research in other Norwegian populations is needed to recommend the BPM for use in community health services. In addition, further evaluations that compare different short screening instruments developed during the last decade are desirable.

Availability of data and materials
The datasets generated and/or analyzed are not publicly available owing to privacy policy; however, they are available from the author on reasonable request.

Declarations

Ethics approval and consent to participate
The Regional Committee of the Medical and Health Research Ethics of Norway, South East, approved this study (No. 2013/1909; Project title: "Coping Kids: a randomized controlled study of a new indicated preventive intervention for children with symptoms of anxiety and depression"). The authors were provided with written informed consent from parents regarding their children's participation.

Consent for publication
Not applicable.

Competing interests
Author KM receives royalties from sales of the EMOTION manual. The other authors declare that they have no competing interest with publishing this article.

Author details
[1]Regional Center for Child and Youth Mental Health and Child Welfare, Department of Mental Health, Norwegian University of Science and Technology, Trondheim, Norway. [2]Department of Child and Adolescent Psychiatry, St. Olavs University Hospital, Trondheim, Norway. [3]Centre for Child and Adolescent Mental Health, RBUP East and South, Oslo, Norway. [4]RKBU – North, Health Sciences Faculty, UiT The Arctic University of Norway, Tromsø, Norway.

References
1. Heiervang E, Stormark KM, Lundervold AJ, Heimann M, Goodman R, Posserud MB, et al. Psychiatric disorders in Norwegian 8- to 10-year-olds: an epidemiological survey of prevalence, risk factors, and service use. J Am Acad Child Adolesc Psychiatry. 2007;46:438–47.
2. Wichstrøm L, Berg-Nielsen TS, Angold A, Egger HL, Solheim E, Sveen TH. Prevalence of psychiatric disorders in preschoolers. J Child Psychol Psychiatry. 2012;53(6):695–705.
3. Kessler RC, Avenevoli S, Costello EJ, Georgiades K, Green JG, Gruber MJ, et al. Prevalence, persistence, and sociodemographic correlates of DSM-IV disorders in the National Comorbidity Survey Replication Adolescent Supplement. Arch Gen Psychiatry. 2012;69(4):372–80.
4. Costello EJ, Mustillo S, Erkanli A, Keeler G, Angold A. Prevalence and development of psychiatric disorders in childhood and adolescence. Arch Gen Psychiatry. 2003;60(8):837–44.
5. Merikangas KR, He JP, Brody D, Fisher PW, Bourdon K, Koretz DS. Prevalence and treatment of mental disorders among US children in the 2001–2004 NHANES. Pediatrics. 2010;125(1):75–81.
6. Ford T, Goodman R, Meltzer H. The British child and adolescent mental health survey 1999: the prevalence of DSM-IV disorders. J Am Acad Child Adolesc Psychiatry. 2003;42(10):1203–11.
7. Angold A, Costello EJ, Erkanli A. Comorbidity. J Child Psychol Psychiatry. 1999;40(1):57–87.
8. Costello EJ, Egger HL, Angold A. The developmental epidemiology of anxiety disorders: Phenomenology, prevalence, and comorbidity. Child Adolesc Psychiatr Clin N Am. 2005;14(4):631–48.
9. Birmaher B, Ryan ND, Williamson DE, Brent DA, Kaufman J. Childhood and adolescent depression: a review of the past 10 years, Part II. J Am Acad Child Adolesc Psychiatry. 1996;35(12):1575–83.
10. Fergusson DM, Woodward LJ. Mental health, educational, and social role outcomes of adolescents with depression. Arch Gen Psychiatry. 2002;59(3):225–31.
11. Kearney CA, Albano AM. The functional profiles of school refusal behavior diagnostic aspects. Behav Modif. 2004;28(1):147–61.
12. Mychailyszyn MP, Méndez JL, Kendall PC. School functioning in youth with and without anxiety disorders: Comparisons by diagnosis and comorbidity. Sch Psych Rev. 2010;39(1):106–21.
13. Swan AJ, Kendall PC. Fear and missing out: Youth anxiety and functional outcomes. Clin Psychol Sci Pract. 2016;23(4):417–35.
14. Egger HL, Angold A. Common emotional and behavioral disorders in preschool children: Presentation, nosology, and epidemiology. J Child Psychol Psychiatry. 2006;47(3–4):313–37.
15. Mathiesen KS. Psykiske lidelser blant barn og unge i Norge. Psykiske lidelser i Norge: et folkehelseperspektiv [Mental disorders among children and young people in Norway. Mental disorders in Norway: a public health perspective]. Oslo: Nasjonalt folkehelseinstitutt; 2009. pp. S41–S56.
16. Reneflot A, Aarø LE, Aase H, Reichborn-Kjennerud T, Tambs K, Øverland S. Psykisk helse i Norge [Mental health in Norway]; 2018.
17. Chorpita BF, Reise S, Weisz JR, Grubbs K, Becker KD, Krull JL, et al. Evaluation of the Brief Problem Checklist: Child and caregiver interviews to measure clinical progress. J Consult Clin Psychol. 2010;78(4):526–36.
18. Achenbach TM, McConaughy SH, Ivanova MY, Rescorla LA. Manual for the ASEBA Brief Problem Monitor (BPM). Burlington, VT: University of Vermont, Research Center for Children, Youth, & Families.: ASEBA; 2011.
19. Achenbach TM, Rescorla LA. Manual for the ASEBA school-age forms & profiles: Child Behavior Checklist for ages 6–18. Teacher's Report form, Youth Self-report: An integrated system of multi-informant assessment. Burlington, VT: University of Vermont, Research Center for Children, Youth, & Families: ASEBA; 2001.
20. Rescorla LA, Achenbach TM, Ginzburg S, Ivanova M, Dumenci L, Almqvist F, et al. Consistency of teacher-reported problems for students in 21 countries. Sch Psych Rev. 2007;36(1):91–110.
21. Richter J. Preliminary evidence for good psychometric properties of the Norwegian version of the Brief Problems Monitor (BPM). Nord J Psychiatry. 2015;69(3):174–8.
22. Penelo E, De la Osa N, Navarro JB, Domènech JM, Ezpeleta L. The Brief Problem Monitor-parent form (BPM-P), a short version of the Child Behavior Checklist: psychometric properties in Spanish 6- to 8-year-old children. Psychol Assess. 2017;29(11):1309–20.
23. Piper BJ, Gray HM, Raber J, Birkett MA. Reliability and validity of Brief Problem Monitor, an abbreviated form of the Child Behavior Checklist. Psychiatry Clin Neurosci. 2014;68(10):759–67.
24. Rodenacker K, Plück J, Döpfner M. Fragebogen zum Problem-Monitoring für Eltern, Lehrer und Jugendliche–eine Deutsche Fassung des Brief Problem Monitor (BPM) [Questionnaire for problem monitoring for parents, teachers and young people - A German version of the Brief Problem Monitor (BPM)]. Z Klin Psychol Psychother. 2015;44:197–206.
25. Backer-Grøndahl A, Martinussen M. Måleegenskaper ved den norske versjonen av Brief Problem Monitor (BPM). [Psychometric properties of the Norwegian version of Brief Problem Monitor (BPM)]. *PsykTestBarn*. 2018;1(2).

Pedersen *et al. BMC Psychology*   (2021) 9:185

Page 12 of 12

26. EFPA. EFPA review model for the description and evaluation of psychological tests: test review form and notes for reviewers, v 4.2. 6. European Federation of Psychologists' Association; 2013.
27. De Los RA, Augenstein TM, Wang M, Thomas SA, Drabick DAG, Burgers DE, et al. The validity of the multi-informant approach to assessing child and adolescent mental health. Psychol Bull. 2015;141(4):858–900.
28. Achenbach TM, McConaughy SH, Howell CT. Child/adolescent behavioral and emotional problems: implications of cross-informant correlations for situational specificity. Psychol Bull. 1987;101(2):213–32.
29. Achenbach TM, Rescorla LA. Multicultural supplement to the manual for the ASEBA school-age forms and profiles. Burlington VT: University of Vermont Research Center for Children, Youth, & Families; 2007.
30. Patras J, Martinsen KD, Holen S, Sund AM, Adolfsen F, Rasmussen L-MP, et al. Study protocol of an RCT of EMOTION: An indicated intervention for children with symptoms of anxiety and depression. BMC Psychol. 2016;4(1):48.
31. Kendall PC, Stark KD, Martinsen K, O'Neil KA, Arora P. EMOTION: "Coping kids" managing anxiety and depression; groupleaders manual. Ardmore, PA: Workbook Publishing; 2013.
32. March JS, Parker JD, Sullivan K, Stallings P, Conners CK. The multidimensional anxiety scale for children (MASC): factor structure, reliability, and validity. J Am Acad Child Adolesc Psychiatry. 1997;36(4):554–65.
33. Angold A, Costello EJ, Messer SC, Pickles A. Development of a short questionnaire for use in epidemiological studies of depression in children and adolescents. Int J Methods Psychiatr Res. 1995;5(4):237–49.
34. Angold A, Erkanli A, Silberg J, Eaves L, Costello EJ. Depression scale scores in 8–17 year olds: effects of age and gender. J Child Psychol Psychiatry. 2002;43(8):1052–63.
35. Ólason DT, Sighvatsson MB, Smári J. Psychometric properties of the multidimensional anxiety scale for children (MASC) among Icelandic schoolchildren. Scand J Psychol. 2004;45(5):429–36.
36. Rhew IC, Simpson K, Tracy M, Lymp J, McCauley E, Tsuang D, et al. Criterion validity of the short mood and feelings questionnaire and one- and two-item depression screens in young adolescents. Child Adolesc Psychiatry Ment Health. 2010;4(1):8.
37. Kornør H, Jozefiak T. Måleegenskaper ved den norske versjonen av Child Behavior Checklist–versjon 2–3, 4–18, 1½-5 og 6–18 (CBCL). [Psychometric properties of the Norwegian version of Child Behavior Checklist Version ages 2–3, 4–18, 1½-5 and 6–18 (CBCL)]. *PsykTestBarn*. 2012;2(1).
38. March JS, Sullivan K, Parker J. Test-retest reliability of the multidimensional anxiety scale for children. J Anxiety Disord. 1999;13(4):349–58.
39. Villabø MA, Gere M, Torgersen S, March JS, Kendall PC. Diagnostic efficiency of the child and parent versions of the Multidimensional Anxiety Scale for Children. J Clin Child Adolesc Psychol. 2012;41(1):75–85.
40. Martinsen KD, Holen S, Neumer S-P, Waaktaar T, Rasmussen LMP, Kendall PC, et al. The factor structure of MASC youth report in Norwegian school children. J Child Fam Stud. 2017;26(7):1808–16.
41. Larsson B, Ingul JM, Jozefiak T, Leikanger E, Sund AM. Prevalence, stability, 1-year incidence and predictors of depressive symptoms among Norwegian adolescents in the general population as measured by the short mood and feelings questionnaire. Nord J Psychiatry. 2016;70(4):290–6.
42. Richter J, Sund AM. Måleegenskaper ved den norske versjonen av the Mood and Feelings Questionnaire-Humøret Ditt (MFQ) [Psychometric properties of the Norwegian version of mood and feelings questionnaire (MFQ)]. *PsykTestBarn*. 2013;1(1).
43. Cohen JW. Statistical power analysis for the behavioral sciences. 2nd ed. Hillsdale: Erbaum Press; 1988.
44. Muthén B. A general structural equation model with dichotomous, ordered categorical, and continuous latent variable indicators. Psychometrika. 1984;49(1):115–32.
45. Li CH. Confirmatory factor analysis with ordinal data: comparing robust maximum likelihood and diagonally weighted least squares. Behav Res Methods. 2016;48(3):936–49.
46. Brown TA. Confirmatory factor analysis for applied research. Guilford Publications; 2015.
47. Tabachnick BG, Fidell LS. Using Multivariate Statistics. Pearson; 2019.
48. Muthén LK, Muthén BO. Mplus User's Guide. Eighth Edition ed. Los Angeles, CA: Muthén & Muthén. 1998–2017.
49. Hooper D, Coughlan J, Mullen M, editors. Evaluating model fit: A synthesis of the structural equation modelling literature. In: 7th European conference on research methodology for business and management studies; 2008.
50. West SG, Taylor AB, Wu W. Model fit and model selection in structural equation modeling. Handb Struct Eq Model. 2012;1:209–31.
51. APA. Diagnostic and statistical manual of mental disorders (DSM-5®). American Psychiatric Publishing; 2013.
52. Ogden T, Hagen KA. Adolescent mental health: Prevention and intervention. Second Edition. Routledge; 2018.
53. De Los RA, Cook CR, Gresham FM, Makol BA, Wang M. Informant discrepancies in assessments of psychosocial functioning in school-based services and research: Review and directions for future research. J Sch Psychol. 2019;74:74–89.
54. Kendall PC. Child and adolescent therapy: Cognitive-behavioral procedures. Kendall PC, editor. Guilford Publications; 2012.
55. Dumenci L, McConaughy SH, Achenbach TM. A hierarchical three-factor model of inattention-hyperactivity-impulsivity derived from the attention problems syndrome of the teacher's report form. Sch Psych Rev. 2004;33(2):287–301.
56. Harter S. The construction of the self 2nd. ed. New York: Guilford Press; 2012.

Publisher's Note