# Psychometric properties of the caregiver-teacher report form in a sample of Norwegian preschool children

Kenneth Stensen, Thomas Jozefiak & Stian Lydersen

Published online: 24 Jan 2022.

Submit your article to this journal ⬚

View related articles ⬚

View Crossmark data ⬚

Taylor & Francis
Taylor & Francis Group

RESEARCH ARTICLE

🔓 OPEN ACCESS | Check for updates

# Psychometric properties of the caregiver-teacher report form in a sample of Norwegian preschool children

Kenneth Stensen, Thomas Jozefiak and Stian Lydersen [ID]

Regional Centre for Child and Youth Mental Health and Child Welfare, Faculty of Medicine and Health Sciences, Department of Mental Health, Norwegian University of Science and Technology (NTNU), Trondheim, Norway

**ABSTRACT**

**Purpose:** The caregiver-teacher report form (C-TRF) is included in the Achenbach system of empirically based assessment (ASEBA) and widely used to measure child psychopathology. In Norway, the C-TRF is frequently used by the Educational and Psychological Counselling Service or for referrals to special health services, however, its psychometric properties in the Norwegian context have not yet been explored. The aim of this study was to investigate the internal consistency of the C-TRF and its factorial validity in a Norwegian preschool context.

**Method:** This study is based on baseline data from the project *Children in Central Norway*, where a total of 169 preschool teachers reported on the C-TRF for 1430 children aged 1–6 years.

**Results:** The findings indicate promising psychometric properties for the C-TRF in terms of internal consistency and factorial validity, however, the *somatic complaints* scale seems problematic because of its poor psychometric properties.

**Conclusion:** Users of the C-TRF can be confident in the instrument's applicability in a Norwegian context, however, careful considerations when applying the *somatic complaints* scale in clinical decision making is warranted.

## Introduction

Globally, approximately 20% of children suffer from mental health problems, while 15–20% of Norwegian children display mental health problems [1,2]. In addition, prevalence estimates indicate that 13–20% of children worldwide meet diagnostic criteria for a psychiatric disorder [3–5]. Moreover, a third of 1- to 7-year-old children who meet criteria for a psychiatric disorder also fulfills the criteria for at least one additional psychiatric disorder [5]. Even though preschool children show similar prevalence estimates as older children, preschool children's mental health problems tend to be overlooked, leading to under-referral and under-treatment [6–8]. For instance, only one-tenth of Norwegian 4-year-old children who meet diagnostic criteria for an emotional or behavioral disorder have received professional help for their problems [9]. Consequently, chances to intervene early in emerging psychopathology and increase the odds for healthy development, both in the short and long term, may be lost. Thus, it is important to intervene at an early stage to prevent mental health problems developing further into stable patterns and disorders. One prerequisite to the success rate of early intervention is the identification of children who would benefit from an intervention. It is essential to have psychometrically sound instruments available to map children's mental health status, however, instruments need to be validated for its intended population to ascertain its accuracy.

In Norway, 92% of children aged one to five years attend childcare centers [10]. Parents and preschool teachers constitute the most viable source of information regarding young children's mental health. Even though parents of preschool children are regarded as 'gatekeepers' to mental health services, preschool teachers also play a major role in initiating contact [9]. Gathering information from other caregivers regarding children's behavior, such as preschool teachers, may capture context-specific behaviors that may not be present in the home environment. It also reduces the over-reliance on parent reports, which may be influenced by other factors (e.g. parental psychopathology) regardless of the child's actual mental health status [11,12]. Identification attempts by preschool teachers without the use of sound psychometric tools could result in misclassifications [13] that consequently may overlook children in need of support or follow-up assessment through non-identification (false negatives), cause stress for the parents and children involved, and burden the support systems unnecessary for those wrongly identified (false positives). As it may be more difficult to distinguish normal from abnormal behavior in younger children

CONTACT Kenneth Stensen ✉ kenneth.stensen@ntnu.no 🏛 Regional Centre for Child and Youth Mental Health and Child Welfare, P.O. Box 8905, MTFS, Trondheim, N-7491, Norway

compared to older children, the preschool period requires developmentally sensitive assessment tools that can capture the full range of behavior relevant to psychopathology, as well as the intensity, frequency, duration, and context in which the behavior occurs. Usually, the discrimination between normal and abnormal behavior focuses on behavioral constructs, clusters, or sets of symptoms relevant to psychopathology, such as anxiety, fear, sadness, and irritability [14,15].

Reliable and valid assessment instruments are important for the identification and treatment of psychopathology in young children [16]. One assessment instrument that has been used since the turn of the millennium is the caregiver-teacher report form (C-TRF) [17], which is a part of Achenbach system of empirically based assessment (ASEBA). The ASEBA is widely used and has been translated into more than 110 languages and thousands of studies have been published that have applied the ASEBA as a measure of child psychopathology [18]. Additionally, the C-TRF and its parent-reported equal, the child behavior checklist (CBCL), are often used as 'gold standard' comparators or benchmark tests for other instruments [19]. To be regarded as a 'gold standard' an instrument's psychometric properties need rigorous and thorough documentation (e.g. reliability and validity indices). One important form of validity is construct validity, more specifically, how constructs relate to other constructs as specified by the theory [20]. In Achenbach and Rescorla's original C-TRF validation study on children in the United States (2000), six syndrome scales and two broadband scales were developed from factor analyses. This two-level factor structure for the C-TRF has generally been supported in 10 out of 14 countries outside the United States [16]. Identification of commonalities in factor structure across societies may stimulate international collaboration between clinicians, mental health practitioners, and researchers [16].

Denmark and Iceland, two Nordic countries that share many cultural similarities with Norway, participated in [16] study of the C-TRF's factor structure across societies. The norm scores of the C-TRF for Denmark, Iceland, and Norway also rank them as low scoring societies in teacher-rated mental health problems among preschool children [21,22]. Although norm scores exist for the C-TRF in a Norwegian preschool context, little is known about the psychometric properties (e.g. reliability and validity) of the instrument in the same context. Even so, the C-TRF is frequently used in research and for multi-informant follow-up assessments or first-assessments in the educational and psychological counselling service, or for referrals to the special health service in Norway. It is important to validate instruments for its intended population to ensure the appropriateness of the instrument, as prevalence and informant perception of mental health problems may vary across societies. Additionally, for instruments considered as a 'gold standard' or benchmark test for other instruments, the documentation of its psychometric properties is of high importance The reliability and validity of a test may strongly influence clinical decision-making and research outcomes.

The aim of the current study was to examine the psychometric properties (internal consistency and factorial validity) of the C-TRF in a Norwegian preschool context.

## Methods

Data used in the current study are based on baseline data from the Children in Central Norway project, which were collected over the period 2012 − 2014. The project aimed to enhance preschool teachers' competence in addressing preschool children's mental health and to improve the quality of the relationship between preschool teachers and children. The study was approved by the Regional Committee for Medical and Health Research Ethics.

### Procedure and participants

Recruitment letters with information regarding the project and an informed consent form were sent to parents with children in childcare centers in three municipalities (Steinkjer, Volda, and former Klaebu) in Central Norway. The children in the childcare centers were from age 1 to 6 years old. Information was also provided in parent meetings before the project started. The recruitment letter provided an option for parents to consent either digitally or by returning the consent form to the childcare center. Participation was voluntary and parental consent could be withdrawn without reprisal at any time until the participation registry was deleted. Parental consent gave the preschool teacher who was most familiar with their child permission to complete a survey regarding the child. Children are usually enrolled in childcare centers in the autumn, and the data were collected in January the following year. Thus, it was assumed that most preschool teachers would have known the child for at least a few months. Preschool teachers themselves provided consent digitally with their own invitation codes. Of the invited parents, 1631 (77%) consented to enroll their child in the study, and 169 teachers (7% men) reported on 1430 children (88% of eligible). The gender distribution of the children was 51% boys and 49% girls, with a mean age of 45 months. The preschool teachers responded for all the children in the same sitting.

### Measurements

#### The caregiver-teacher report form (C-TRF)
The preschool teachers completed the C-TRF [17], which contains 100 items describing symptoms of mental health problems for children aged from 1.5 to 5 years old. Each item has three response options: 0= 'not true (as far as you know)', 1= 'somewhat or sometimes true' and 2= 'very often or often true'. The C-TRF contains the following syndrome scales: *emotionally reactive* (7 items), *anxious/depressed* (8 items), *withdrawn* (10 items), *somatic complaints* (7 items), *attention problems* (9 items), and *aggressive behavior* (25 items. In addition, two broadband scales can be calculated by adding the corresponding syndrome scales into *internalizing problems* (*emotionally reactive, anxious/depressed,*

Table 1. Cronbach's α and McDonald's ω with 95% CI for the caregiver-teacher report form (C-TRF) syndrome scales compared with the α for the original normative United States sample.

| Scale (n items) | Cronbach's α [95% CI] | McDonald's ω [95% CI] | α [17] |
|---|---|---|---|
| Emotionally reactive (7) | .683 [.656, .701] | .703 [.679, .727] | .71 |
| Anxious/depressed (8) | .690 [.668, .711] | .689 [.664, .713] | .76 |
| Somatic complaints (7) | .441 [.400, .484] | .376 [.323, .433] | .52 |
| Withdrawn (10) | .752 [.734, .769] | .757 [.738, .776] | .83 |
| Attention problems (9) | .874 [.863, .884] | .888 [.879, .897] | .89 |
| Aggressive behavior (25) | .921 [.915, .927] | .926 [.921, .932] | .96 |

withdrawn, and somatic complaints) and externalizing problems (attention problems and aggressive behavior). A total problem score can be calculated by summing all 100 items, thus yielding a score between 0 and 200. The two broadband scales constitute 66 of the 100 items in the C-TRF, while the remaining 34 items are labeled as other problems.

## Data analyses

First, Cronbach's alpha (α) and McDonald's omega (ω) with 95% confidence interval (CI) were calculated for the syndrome scales of the C-TRF. Cronbach's α and McDonald's ω are estimates of the internal consistency of multi-item scales, where the former is more frequently reported than the latter. McDonald's ω was chosen as the primary indices for scale internal consistency because it has better and more realistic data assumptions [23–26]. For instance, Cronbach's α assumes constant item variance for true scores, while allowing the true score means and the error variance of items to vary [27]. In contrast, McDonald's ω is less restrictive, allowing the means and variance of true scores, as well as the error variance to vary [23]. Cronbach's α and McDonald's ω were calculated using the MBESS (v. 4.8.0) package in RStudio using the ci.reliability command. The maximum likelihood estimator and confidence intervals (CI) (mll command) based on logistic transformation were used in the calculation of reliability coefficients and CI. One benefit of using logistic transformation to calculate CI is that logistic transformation does not assume a symmetric sampling distribution of reliability [28]. A commonly used threshold for acceptable α regarding scale internal consistency is ≥ .70 [29]. In the current study we applied the same threshold for the ω coefficient.

The intraclass correlation coefficient ranged between .11 and .14 at the preschool teacher level for the individual syndrome scales, while the residual intraclass correlation coefficient was .035 for all syndrome scales combined. Due to the low values of these correlation coefficients, we chose to not apply multilevel analyses.

Next, we performed a hierarchical confirmatory factor analysis (CFA) of the 66 C-TRF items that loaded significantly on the six syndrome scales and the two broadband scales based on the derived factor structure from the original United States sample [17]. Items to syndrome scales constituted level one, and syndrome scales to broadband scales constituted level two. Each item in the CFA could only load to its originally intended factor. CFA was performed in the Mplus

(v. 8) software using the weighted least square mean variance (WLSMV) estimator. The root mean square error of approximation (RMSEA) was chosen as the main indicator of model fit. RMSEA's ≤.05 indicates a good fit and values between .05 and .10 indicates an acceptable fit [30]. The comparative fit index (CFI) and Tucker-Lewis index (TLI) are also reported to follow the convention of multiple fit indices. CFI and TLI values ≥.95 are commonly used to indicate a good model fit [31], however, the ≥.95 threshold for CFI and TLI has been criticized for being too conservative for complex models [32]. Thus, following the recommendation of [33], a CFI and TLI threshold of >.90 were used in the current study to indicate a good model fit and .80–.90 indicated an acceptable model fit. This said, the CFI and TLI estimates are considered secondary to the RMSEA as the appropriateness for their use with categorical data is uncertain [31].

There were no missing data. Except where otherwise noted, the analyses were carried out in Stata 17.

## Results

The internal consistency estimates for the C-TRF syndrome scales are presented in Table 1. The Cronbach's α estimates for the syndrome scales withdrawn, attention problems, and aggressive behavior all surpassed the acceptable threshold of ≥ .70, while emotionally reactive, anxious/depressed, and somatic complaints did not. However, the 95% CI of the emotionally reactive and anxious/depressed do include this threshold. The highest α was found for the syndrome scale aggressive behavior (.921), while the lowest was found for the syndrome scale somatic complaints (.441). For the McDonald's ω coefficient, the syndrome scales emotionally reactive, withdrawn, attention problems, and aggressive behavior all surpassed the acceptable threshold of ≥ .70, while the anxious/depressed and somatic complaints syndrome scales did not. As in the case of the α, for the anxious/depressed, the upper 95% CI for the ω also included the threshold for acceptable internal consistency. Similar to α, the highest ω coefficient was found for the aggressive behavior syndrome scale (.926) and the lowest for somatic complaints syndrome scale (.376). As seen in Table 1, the difference between the α and ω coefficients in the current study is small, while the biggest difference found is in the somatic complaints syndrome scale (α .441 and ω .376). The α and ω coefficients are also lower in general than the α from the original United States sample [17].

The standardized factor loadings and correlation can be seen in Figure 1 and Table 2 shows the model fit indices from the CFA analyses of the first- and second- order scales. The RMSEA indicates an acceptable to good fit for the anxious/depressed, withdrawn, attention problems, and aggressive behavior syndrome scales. Emotionally reactive slightly exceeded the RMSEA criterion for an acceptable model fit, however, for the emotionally reactive syndrome scale the CFI and TLI estimates indicate an acceptable model fit. The CFI and TLI for anxious/depressed, withdrawn, attention problems, and aggressive behavior indicates a good model fit. The only syndrome scale that meets all three model fit criteria for a good model fit is the attention problems syndrome scale
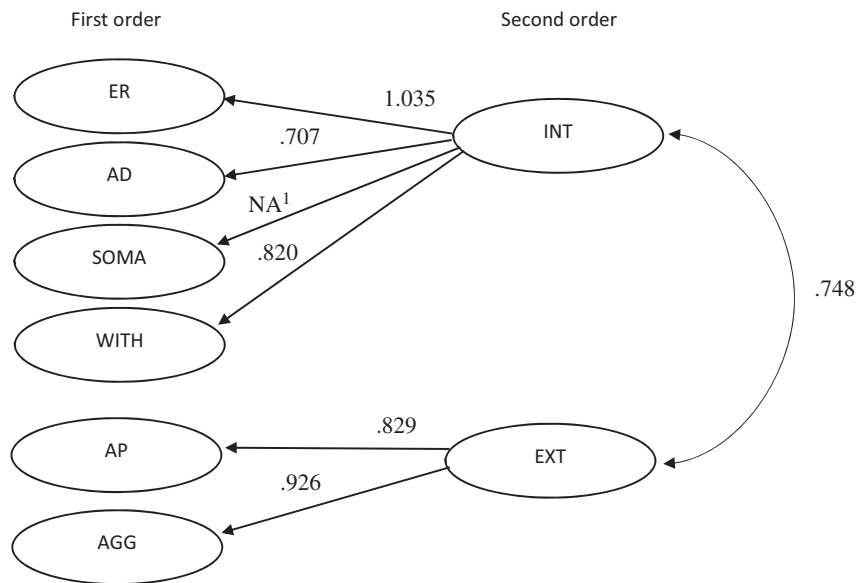
First order                                    Second order



**Figure 1.** Caregiver-teacher report form (C-TRF) factor model with standardized factor loadings and correlation. *Note*: ER: emotionally reactive; AD: anxious/depressed; SOMA: somatic complaints; WITH: withdrawn; AP: attention problems; AGG: aggressive behavior; INT: internalizing problems; EXT: externalizing problems. NA[1] = Somatic complaints omitted due to non-convergence.

**Table 2.** Model fit indices for first order syndrome scales and second order broadband scales for the caregiver-teacher report form (C-TRF).

|  | RMSEA | CFI | TLI |
|---|---|---|---|
| First order |  |  |  |
| Emotionally reactive | .11 | .89 | .84 |
| Anxious/depressed | .06 | .95 | .94 |
| Somatic complaints | NA[1] | NA[1] | NA[1] |
| Withdrawn | .06 | .96 | .94 |
| Attention problems | .05 | .99 | .99 |
| Aggressive behavior | .06 | .93 | .92 |
| Second order |  |  |  |
| Internalizing and externalizing problems | .05 | .87 | .87 |

*Note.*[1]Somatic complaints omitted due to non-convergence; RMSEA: root mean square error of approximation; CFI: comparative fit index; TLI: Tucker-Lewis index.

(RMSEA .05; CFI .99; TLI .99). The *somatic complaints* syndrome scale was omitted because of non-convergence. Consequently, only five syndrome scales were used in the second-order CFA. The second- order CFA with the five remaining syndrome scales indicated a good model fit for the second-order broadband scales *internalizing* and *externalizing problems* (RMSEA .05), as well as an acceptable CFI (.87) and TLI (.87).

## Discussion

The aim of the current study was to investigate the psychometric properties of the C-TRF in a Norwegian context, which has not been explored previously. Despite this, the C-TRF is frequently used in research and clinical work. Findings from the current study add to the knowledge of the applicability of the C-TRF. Overall, the reported findings agree with the original two-level model structure of the C-TRF [17], which has also been supported in several other countries outside the United States [16]. However, the syndrome scale *somatic complaints* show poor internal consistency and does not converge in the model framework. Thus, the applicability of this

scale is questionable and should perhaps be omitted all together, as the second-order broadband scales *internalizing* and *externalizing problems* show a good model fit without the *somatic complaints* syndrome scale.

### Internal consistency

Regarding internal consistency, the C-TRF shows promising estimates for the externalizing scales *attention problems* and *aggressive behavior*, while for the internalizing syndrome scales, only *withdrawn* surpasses an acceptable level for both $\alpha$ and $\omega$. This said, both the *emotionally reactive* and *anxious/depressed* syndrome scales are at or approaches the threshold of acceptable internal consistency, and the upper limit includes this threshold. However, as arbitrary cutoff values may make researchers discard important findings, more attention should be given to the CIs rather than the point estimates [34]. Thus, stating that a scale has an unacceptable internal consistency because of a coefficient slightly below threshold might hinder the development of instruments and models.

Similar to the original study by [17], the two externalizing scales of *attention problems* and *aggressive behavior* exhibit the most internal consistency. The high internal consistency estimate found for the *aggressive behavior* syndrome scale may also be explained by the share number of items (25), which is almost three times the number of items compared to the other syndrome scales. If all internal consistency parameters are kept constant across scales and only the number of items is allowed to vary, the number of items will influence the internal consistency coefficients [35]. The syndrome scale *somatic complaints* were found to be the least consistent in the original study of the C-TRF [17], which corresponds with the findings of the current study as well. This may be due to issues with the scale itself, but it can also reflect preschool teachers' difficulty to distinguish somatic

symptoms from other types of internalizing and externalizing symptoms. In addition, younger children have a more limited language repertoire and ability to communicate how they feel. For instance, a child's inability to sit still due to stomachache may be interpreted by preschool teachers as an attention symptom rather than a somatic symptom. Additionally, the C-TRF specifies that the somatic complaints should be without a medical cause, which may also be difficult to distinguish for preschool teachers. Another interesting point to mention regarding internal consistency in this study is the case of the *emotionally reactive* syndrome scale, which is below the threshold with the $\alpha$ coefficient but above the threshold with $\omega$. Thus, researchers may reach different inferences based on which coefficient is used. This underlines the importance of reporting confidence intervals (CIs).

### Factorial validity

Consistent with earlier research on the factor structure of the C-TRF [16], the current study also indicates promising factorial validity indices for the C-TRF in a Norwegian context. In some countries, support has been found for the original factor structure of the C-TRF with six first-order syndrome scales and the two second-order broadband scales, while in other countries, support has been found for a slightly modified factor structure with five first order syndrome scales, mainly due to the high correlation between e*motionally reactive* and *anxious/depressed* [16]. The current study supports a five-syndrome scale first order model, but with the exclusion of the *somatic complaints* syndrome scale due to its non-convergence and low internal consistency. Based on the RMSEA of the *emotionally reactive* syndrome scale, some adjustment might be needed to increase the model fit. However, the CFI and TLI indicate that the *emotionally reactive* syndrome scales have an acceptable model fit. The applicability of the rest of the first-order syndrome scales (*anxious/depressed*, *withdrawn*, *attention problems*, *aggressive behavior*) and the two second-order broadband scales of *internalizing* (without the *somatic complaints* syndrome scale) and *externalizing problems* seems promising. Consequently, it is advisable to remove *somatic complaints* from the *internalizing* broadband scale and instead place *somatic complains* together with *other problems* on the C-TRF, if the items should be used at all. Findings from the current study indicate that researchers, clinicians, and other practitioners may use the five syndrome factors and two broadband factor structures of the C-TRF with greater confidence in mapping a broad range of symptoms related to childhood psychopathology.

### Strengths and limitations

One of the major strengths of the current study is the large sample size and the inclusion of the full age range of preschool children. In addition, the current study reports an alternative coefficient for internal consistency to the popular and frequently reported Cronbach's $\alpha$. An increasing number of researchers have advocated moving away from Cronbach's $\alpha$ to McDonald's $\omega$ or other coefficients of internal consistency, as the assumptions for $\alpha$ are rarely met [26,36,37]. This said, the $\omega$ also shares some of the assumptions as the $\alpha$ (e.g. unidimensionality), however, owing to the less restrictive nature of $\omega$, it is less biased when the assumptions are violated (e.g. less likely to under- or overestimate). Thus, $\omega$ is a more robust estimate of the internal consistency than $\alpha$. Another strength of the current study is the inclusion of the CIs associated with the internal consistency coefficients. The use of point estimates without CI and 'rule of thumb' cutoff values (e.g. .70) is problematic as it does not reflect the uncertainty for values that are around this cutoff [38]. A final strength of this study is the model fit evaluation for both the first- and second-order factor structures rather than just the overall model, as it is important to know how the first-order syndrome scales relates to the second-order broadband scales. One limitation of the current study is that each preschool teacher reported on multiple children. Thus, reports are not independent of each other, and the estimated CIs may be slightly too narrow. However, low values of the intraclass indicate that this is not an issue in the current study. Future studies should investigate how the characteristics of the preschool teachers influence their reporting on the C-TRF and how this may influence the model fit for the C-TRF. For instance, it has been shown that preschool teachers who perceive their relationship to children as conflictual report children with more symptoms than they actually have [39]. Additionally, there is a need for future research to investigate the psychometric properties and applicability of the C-TRF's parent-reported counterpart, the CBCL, in a Norwegian preschool population.

## Conclusion

Findings from the current study support the applicability of the C-TRF in a Norwegian context. The C-TRF exhibited promising internal consistency and factorial validity, however, the syndrome scale *somatic complaints* seem problematic, and its usefulness seems limited. These findings suggest a five-factor syndrome scale model with two broadband scales, rather than the original six syndrome scales with two broadband scales. Researchers and professionals using the C-TRF to gather information regarding childhood psychopathology from preschool teachers and other caregivers might use the five-by-two model structure with greater confidence.

## Ethical approval

The study was approved by the Regional Committee for Medical and Health Research Ethics (REK) (reference number: 2011/2252; 2012/1030; 2013/1363)

## Disclosure statement

None of the authors declare any competing or potential conflicts of interest.

## Notes on contributors

*Kenneth Stensen*, PhD, is an associate professor in medicine and health sciences. His research interests cover topics such as developmental psychopathology and psychometrics.

*Thomas Jozefiak*, MD, PhD, is a child- and adolescent psychiatrist and professor in clinical medicine. His research interests are Quality of life and well-being, epidemiology, mental health/psychopathology and neuropsychiatry in children and adolescents, child maltreatment/childhood trauma, adolescents in child protection services and residential youth care.

*Stian Lydersen*, PhD, is professor of medical statistics. He is one of the authors of the books 'Medical Statistics in Clinical and Epidemiological Research' (2012) and 'Statistical Analysis of Contingency Tables' (2017). He is a contributor to the column 'Medicine and numbers' in The Journal of the Norwegian Medical Association.

## ORCID

Stian Lydersen http://orcid.org/0000-0001-6613-8596

## References

[1] Belfer ML. Child and adolescent mental disorders: the magnitude of the problem across the globe. J Child Psychol Psychiatry. 2008;49(3):226–236.

[2] Skogen JC, Kjeldsen A, Knudsen AK, et al. 2014. Psykisk helse hos barn og unge (Folkehelserapporten 2014:4). Oslo: Nasjonalt Folkehelseinstitutt.

[3] Charach A, Mohammadzadeh F, Belanger SA, et al. Identification of preschool children with mental health problems in primary care: systematic review and meta-analysis. J Can Acad Child Adolesc Psychiatry. 2020;29(2):76–105., Szatmari P,

[4] Polanczyk GV, Salum GA, Sugaya LS, et al. Annual research review: a meta-analysis of the worldwide prevalence of mental disorders in children and adolescents. J Child Psychol Psychiatry. 2015;56(3):345–365. ).

[5] Vasileva M, Graf RK, Reinelt T, et al. Research review: a meta-analysis of the international prevalence and comorbidity of mental disorders in children between 1 and 7 years. J Child Psychol Psychiatr. 2021;62(4):372–381.

[6] Egger HL, Angold A. Common emotional and behavioral disorders in preschool children: presentation, nosology, and epidemiology. J Child Psychol Psychiatry. 2006;47(3–4):313–337.

[7] Horwitz SM, Gary LC, Briggs-Gowan MJ, et al. Do needs drive services use in young children? Pediatrics. 2003;112(6 Pt 1):1373–1378.

[8] Horwitz SM, Kelleher KJ, Stein REK, et al. Barriers to the identification and management of psychosocial issues in children and maternal depression. Pediatrics. 2007;119(1):e208–218.

[9] Wichstrøm L, Belsky J, Jozefiak T, et al. Predicting service use for mental health problems among young children. Pediatrics. 2014;133(6):1054–1060.

[10] Statistics Norway. 2020. Kindergartens. https://www.ssb.no/barnehager.

[11] Harvey EA, Fischer C, Weieneth JL, et al. Predictors of discrepancies between informants' ratings of preschool-aged children's behavior: an examination of ethnicity, child characteristics, and family functioning. Early Child Res Q. 2013;28(4):668–682.

[12] Salomonsson B, Sleed M. The ages & stages questionnaire: social-emotional: a validation study of a mother-report questionnaire on a clinical mother-infant sample. Infant Ment Health J. 2010;31(4):412–431.

[13] Stensen K, Lydersen S, Stenseng F, et al. Childcare providers'nominations of preschool children at risk for mental health problems: does it discriminate well compared to the caregiver-teacher report form? Scand J Educ Res. 2021;1–15. DOI: 10.1080/00313831.2020.1869089

[14] Carter AS, Briggs-Gowan MJ, Davis NO. Assessment of young children's social-emotional development and psychopathology: recent advances and recommendations for practice. J Child Psychol Psychiatry. 2004;45(1):109–134.

[15] Dougherty LR, Leppert KA, Merwin SM, et al. Advances and directions in preschool mental health research. Child Dev Perspect. 2015;9(1):14–19.

[16] Ivanova MY, Achenbach TM, Rescorla LA, et al. Syndromes of preschool psychopathology reported by teachers and caregivers in 14 societies using the caregiver-teacher report form (C-TRF). J Early Child Infant Psychol. 2011;7:87–103.

[17] Achenbach TM, Rescorla LA. 2000. Manual for the ASEBA preschool forms & profiles. Burlington (VT): University of Vermont, Research Center for Children, Youth, & Families.

[18] ASEBA.org. 2020. Achenbachs system of empirically based assessment. https://aseba.org/.

[19] Lavigne JV, Meyers KM, Feldman M. Systematic review: classification accuracy of behavioral screening measures for use in integrated primary care settings. J Pediatr Psychol. 2016;5:1–19.

[20] Strauss ME, Smith GT. Construct validity: advances in theory and methodology. Annu Rev Clin Psychol. 2009;5:1–25.

[21] Drugli MB, Stensen K. Psykiske vansker hos norske barnehagebarn. Første Steg. 2019;1:46–49.

[22] Ivanova MY, Achenbach TM, Rescorla LA, et al. Testing the teacher's report form syndromes in 20 societies. School Psychol Rev. 2007;36(3):468–483.

[23] Dunn TJ, Baguley T, Brunsden V. From alpha to omega: a practical solution to the pervasive problem of internal consistency estimation. Br J Psychol. 2014;105(3):399–412.

[24] Hayes AF, Coutts JJ. Use omega rather than Cronbach's alpha for estimating reliability. But … . Commun Methods Meas. 2020;14(1):1–24.

[25] Peters G-JY. The alpha and omega of scale reliability and validity: why and how to abandon the cronbach's alpha and the route to a more comprehensive assessment of scale quality. Eur Health Psychol. 2014;16:56–69.

[26] Ravinder B, Saraswathi AB. Litterature review of cronbachalpha-coefficient (A) and McDonald's omega coefficient (ω). Eur J Mol Clin Med. 2020;7:2943–2949.

[27] Raykov T. Scale reliability, Cronbach's coefficient alpha, and violations of essential tau-equivalence with fixed congeneric components. Multivariate Behav Res. 1997;32(4):329–353.

[28] Browne MW. Covariance structures. In: Hawkins DM, editor. Topics in applied multivariate analysis. Cambridge (UK): Cambridge University Press; 1982. p. 72–141.

[29] Taber KS. The use of cronbach's alpha when developing and reporting research instruments in science education. Res Sci Educ. 2018;48(6):1273–1296.

[30] MacCallum RC, Browne MW, Sugawara HM. Power analysis and determination of sample size for covariance structure modeling. Psychol Methods. 1996;1(2):130–149.

[31] Hu L, Bentler PM. Cutoff criteria for fit indexes in covariance structure analysis: conventional criteria versus new alternatives. Struct Equat Model. 1999;6(1):1–55.

[32] Marsh HW, Hau K-T, Wen Z. In search of golden rules: comment on hypothesis-testing approaches to setting cutoff values for fit indexes and dangers in overgeneralizing Hu and bentler's (1999) findings. Struct Equation Model. 2004;11(3):320–341.

[33] Browne MW, Cudeck R. Alternative ways of assessing model fit. In Bollen KA, Long JS, editors. Testing structural equation models. Newbury Park (CA): Sage. 1993. p. 136–162.

[34] Agbo AA. Cronbach's alpha: Review of limitations and associated recommandations. J Psychol Africa. 2010;20(2):233–239.

[35]    Cortina JM. What is coefficient alpha? An examination of theory and applications. J Appl Psychol. 1993;78(1): 98–104.

[36]    McNeish D. Thanks coefficient alpha. We'll take it from here. Psychol Methods. 2018;23(3):412–433.

[37]    Sijtsma K. On the use, misuse, and the very limited usefulness of Cronbach's alpha. Psychometrika. 2009;74(1):107–120.

[38]    Pfadt JM, van den Bergh D, Sijtsma K, et al. Bayesian estimation of single-test reliability coefficients. Multivariate Behav Res. 2021; 1–30. DOI: 10.1080/00273171.2021.1891855

[39]    Berg-Nielsen TS, Solheim E, Belsky J, et al. 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(3):393–413.