

Psychometric properties of the Norwegian version of the Strength and Difficulties Questionnaire in a preschool sample

Nina Jakhelln Laugen^a , Hege Midtli^a, Ulrika Löfkvist^b and Kenneth Stensen^c

^aDepartment of Psychology, Norwegian University of Science and Technology, Trondheim, Norway; ^bDepartment of Public Health and Caring Sciences, Uppsala University, Uppsala, Sweden; ^cRegional Centre for Child and Youth Mental Health and Child Welfare (RKBU Central Norway), Department of Mental Health, Norwegian University of Science and Technology, Trondheim, Norway

ABSTRACT

Purpose: The Strength and Difficulties Questionnaire (SDQ) is widely used internationally, however less so in preschool populations and validation studies are thus needed. This study examined the psychometric properties of the Norwegian version parent report of the SDQ – preschool version (SDQ 2–4).

Materials and methods: Parents of 289 Norwegian children in the age span 1–6 years old filled out the SDQ 2–4, the Child Behavior Checklist (CBCL), and background information. Internal consistency, factor structure, and convergent validity were assessed.

Results: The results showed satisfying internal consistency for the total difficulties score, but worse for some of the subscales. The five-factor structure showed a good fit. Good convergent and divergent validity was found in terms of correlations with CBCL. Sex differences were found on all scales, boys scoring higher on all problem scales.

Conclusions: The SDQ 2–4 can be a promising instrument to screen for emotional and behavioral difficulties among Norwegian preschoolers, particularly in high-risk populations.

ARTICLE HISTORY

Received 3 October 2023

Revised 11 April 2024

Accepted 30 April 2024

KEYWORDS

Psychometry; screening; socio-emotional functioning; preschool

Psychosocial difficulties in preschool age are common, with an estimated prevalence of 10–20% [1–3]. For some children, such difficulties develop into psychiatric disorders [1, 4–9]. Early intervention can ameliorate internalizing and externalizing difficulties in children [10–12], however, it is challenging to identify the children in need of support or treatment [13–15]. Standardized screening instruments, such as parent-reported symptom checklists, are frequently used to identify these children [16, 17], however, documentation of their psychometric properties is often lacking [18]. Because symptoms may manifest themselves differently in preschool age compared to older children [19] and cultural and linguistic factors can influence the reporting and interpretation of checklists, they need to be sensitive to the age and culture in which it is used.

The Strength and Difficulties Questionnaire (SDQ) [20, 21] is widely used to screen for psychosocial difficulties in clinical practice and in research and holds good psychometric properties [22]. The preschool version, the SDQ 2–4, has the potential to contribute to earlier and better-targeted support to children at risk. Per date, the Norwegian version has not yet been subject to a psychometric evaluation. This study aims to assess reliability and validity of the Norwegian SDQ 2–4.

The Strength and Difficulties Questionnaire in preschool children

The SDQ [20, 23] is an internationally well-known measure of psychosocial functioning in children and adolescents. It is

used in research on psychosocial functioning in the general population [24] as well as in clinical population, such as children with hearing loss [25] and children who are exposed to trauma [26]. It is also used in clinical practice, and can be used as a screening tool to detect children at risk for developing psychiatric disorders [3]. Over the past decade, psychometric properties in preschool populations have been reported in a range of studies, covering several linguistic and cultural contexts. Some of these studies have tested the properties of the school-age version of SDQ in preschool population, whereas others have used the adapted preschool version (SDQ 2–4) in which three items have been altered for better fit with the age group (Table 1). The Norwegian version of SDQ 2–4 has not yet been subject to such evaluation.

Internal consistency

Previous studies have reported acceptable to satisfactory internal consistency for SDQ in preschool samples, with Cronbach's alpha values ranging from 0.71 to 0.88 for the total difficulties score [3, 29, 31, 33, 34, 36, 38–40]. As for the subscales, studies report heterogeneous results; for example, in a study of 1–2-year-olds [38], report $\alpha = 0.38$ for the subscale 'conduct problems' whereas [30] found $\alpha = 0.75$ in their study of 3–7-year-olds. Across studies, the Peer problems scale yields relatively low scores, in fact as low as 0.33 [36], and to our knowledge, only one study has reported Cronbach's α higher than 0.70 [31].

Cronbach's alpha is the most frequently used measure of internal consistency. However, it relies on assumptions of normal distribution and equal variance, which may be hard to meet in many studies. Thus, in many cases, Cronbach's alpha estimations may be inaccurate or biased. McDonald's Omega has been suggested as an alternative coefficient, being less dependent on strict assumptions and therefore less exposed to bias [41]. A few studies have reported high omega coefficients for SDQ 2–4 in preschool samples [30, 35].

Factor structure

SDQ 2–4 has five subscales: emotional problems, problems with friends, conduct problems, hyperactivity/inattention, and prosocial behavior. For the school version of SDQ, a three-factor structure has been proposed, including internalizing problems (emotional problems+problems with friends), externalizing problems (hyperactivity/inattention and conduct problems), and prosocial behavior [42]. In preschool samples, the original five-factor structure is mostly supported and rated as having an acceptable or a good fit [30, 31, 35, 39]. The three-factor structure has received some support [35], and although the five-factor structure seems to have a better fit, the difference between the two structures is relatively low [30]. One study proposes a modified five-factor model, adding a 'positivity factor' including the prosocial scale and other items with positive wording [34].

Convergent and divergent validity

Regarding convergent and divergent validity estimates, comparisons with the Child Behavior Checklist (CBCL) [43, 44], which is a well-esteemed measure of emotional and

behavioral difficulties, yield positive correlations between the SDQ 2–4 subscales and corresponding subscales in CBCL in Dutch and British samples [38, 39]. Also, negative correlations between the SDQ 2–4 prosocial subscale and the CBCL total difficulties score were reported [38]. Many of the studies report sex differences, boys being reported to having higher problems scores and girls having higher prosocial scores [32, 34–36, 40]. These findings align with the sex differences in mental health reported in epidemiological studies [45, 46].

Overall, previous studies suggest promising reliability and validity estimates; however, psychometric properties of the Norwegian translation of the SDQ 2–4 have not yet been investigated. The aim of this study is to examine internal consistency, factor structure, convergent and divergent validity in a sample of Norwegian preschool children.

Methods

Recruitment and participants

Parents of children aged 1–6 years were recruited primarily via preschools in both rural and urban areas of Norway. Information was also provided on social media. Because more than 90% of Norwegian children are enrolled in a preschool [47], a representative sample of the population was likely to be reached. It should be noted, however, that 51% of the participating parents had a master's degree or more, which is a higher proportion than the national average for adults aged 25–50 years which is 18% [48]. A total of 295 parents entered the study, six participants were excluded due to their child being outside of the age range for the study or due to missing data. For the included 289 participants, 54 of the respondents were male. Mean age of the child was 40.56 months (SD 15.00, range 11–74). See Table 2 for demographic details.

Data were collected as an online survey that provided anonymity for the participants. The need for ethical approval was waived by the Regional Committee for Medical and Health Research Ethics (No.140084). The full dataset is available at DataverseNO [49].

Instruments

SDQ 2–4

The SDQ questionnaires (www.sdqinfo.org) consist of 25 items, each describing a psychological attribute (e.g. 'many

Table 1. Psychometric evaluation studies of SDQ 2–4 in preschool populations.

Study	Population	N (age)	Examined properties	Evaluation
[27]	Denmark	1344 (2–6)	Norms	Norms
[28]*	Finland	1774 (4–9)	Test-retest and interrater rel.	Satisfactory
[29]*	Finland	2666 (4–9)	Criterion val., norms	Satisfactory
[30]**	UK	14 444 (3–7)	Construct and predictive val., internal rel.	Satisfactory
[31]	Sweden	7113 (3–5)	Construct val., internal rel.	Satisfactory
[32]	Sweden	11 196 (3–5)	Norms	Norms
[33]	Japan	383 (4–6)	Internal and test-retest rel.	Acceptable
[34]	New Zealand	5481 (2)	Construct val., internal rel.	Satisfactory
[35]	Spain	1341 (3–4)	Construct and criterion val., internal rel.	Acceptable
[32]*	Germany	1738 (3–4)	Construct val., internal rel.	Satisfactory
[36]	Netherlands	376 (2–5)	Internal rel., norms	Satisfactory (total score)
[37]**	Denmark	1176 (3–8)	Predictive val.	Promising
[38]	UK	93 (1–2)	Construct val., internal rel.	Acceptable
[3]*	Norway	845 (4)	Criterion val.	Moderate support
[39]	Netherlands	839 (3–4)	Construct and criterion val., internal rel.	Satisfactory

*SDQ 4–17 was used. **Longitudinal study.

Rel.: reliability; Val.: validity

Table 2. Demographic information about participants.

	Girls	Boys	Total
Sex, no (%)	142 (49)	147 (51)	289 (100)
Age, mean (SD) months	41.8 (14.8)	40.1 (15.3)	40.6 (15.0)
Gestation week <37, no (%)	13 (9.2)	12 (8.2)	25 (8.7)
Reduced hearing, no (%)	2 (1.4)	4 (2.7)	6 (2.1)
Reduced vision not aided with glasses, no (%)	1 (.7)	0 (0)	1 (.3)
Other chronic conditions, no (%)	7 (4.9)	6 (4.1)	13 (4.5)
Attending preschool, no (%)	140 (98.6)	146 (99.0)	286 (99.0)
Informant's sex male, no (%)	28 (19.7)	26 (17.7)	54 (18.7)
Informant's education, no (%)	65 (45.8)	83 (56.5)	148 (51.2)
Master's degree or higher			

fears, easily scared') to be rated by a parent, teacher, or the child itself as not true, somewhat true, or certainly true for the child in question. In our study, the form for parents was administered. The SDQ 2–4 is a modified version of the original instrument, in which 22 items are identical and three items are changed (www.sdqinfo.org). Just as for the other versions, the SDQ 2–4 yields five subscales; emotional problems, conduct problems, hyperactivity/inattention, peer relationship problems, and prosocial behavior. The scales can be combined to form three composite scores: Internal difficulties, external difficulties, and prosocial behavior.

The Norwegian version of SDQ was introduced in 1999 after a translation- and back-translation process [50], and has been used in many larger studies in Norway [51]. The SDQ 2–4, with the three modified items, is available from www.sdqinfo.org. To our knowledge, the SDQ 2–4 is rarely used in Norway and has not been subject to a psychometric evaluation.

CBCL

The CBCL1 ½–5 [43] is part of the Achenbach System of Empirical Based Assessment (ASEBA). It has been referred to as the 'gold standard' of parent-report assessments [43, 52] and is widely used in Norway. The CBCL consists of 100 items describing children's behaviors or problems, and the parent responds on a 3-point Likert scale. Acceptable psychometric properties have been reported in a range of linguistic and cultural contexts [43, 53, 54] including Norwegian for school-age [44]. In addition to the total problems score, there are seven subscales: Emotionally reactive, anxious/depressed, somatic complaints, withdrawn, attention problems, aggressive behavior, and sleep problems.

Analyses

Statistical analyses were performed using SPSS version 27 (SPSS Inc., Chicago, IL) [55] unless otherwise noted. McDonald's omega total (ω) was chosen as the main indicator for scales internal consistency, as it is less sensitive to deviations from the assumptions associated with Cronbach's alpha, such as unidimensionality, tau-equivalence, and normally distributed scores [41, 56–59]. However, to allow

comparison with previous studies, Cronbach's α is also reported. For both measures, scores of 0.70 or higher were considered acceptable, in accordance with guidelines from the European Federation of Psychologists' Associations [60]. The analyses for internal consistency were performed with RStudio using the 'psych' package [61].

The factor structure was examined by performing a confirmatory factor analysis (CFA) using Mplus version 8.4 (Muthén & Muthén, Los Angeles, CA) [62]. Both the original five-factor structure and the proposed three-factor structure were tested. The weighted least square mean variance (WLSMV) estimator was used to account for the non-normal distribution of data. The WLSMV estimator produces accurate parameter estimates in ordered categorical data [63]. To assess the model fit, the root mean square error of approximation (RMSEA), comparative fit index (CFI), and Tucker–Lewis index (TLI) were examined. RMEA values of ≤ 0.05 were considered a good fit and 0.05–0.10 as acceptable [64]. For CFI and TLI, values > 0.90 were considered a good fit and 0.80–0.90 as acceptable [65].

Convergent and divergent validity was assessed with correlations between SDQ 2–4 and CBCL, including subscales, composite scores, and total scores, using Spearman's rho (ρ). Correlations of $\rho = 0.55$ –0.64 were considered satisfactory, $\rho = 0.65$ –0.74 good, and $\rho \geq 0.75$ as excellent [60]. In addition, as boys tend to exhibit higher degrees of psychosocial difficulties in preschool age [46], Welch t-tests for sex differences on all subscales were conducted to examine whether this effect was evident for SDQ 2–4. Bootstrapping with 1000 samples was applied.

Results

Internal consistency

Internal consistency estimates are presented in Table 3. The internal consistency for the total score was satisfactory ($\omega = 0.82$), however the estimates of the subscales were varying. The conduct problems subscale had the lowest of $\omega = 0.60$ and the hyperactivity/inattention subscale reached the highest, $\omega = 0.82$.

Factor structure

The original five-factor structure produced a good to acceptable fit, RMSEA = 0.051, CFI = 0.898, and TLI = 0.884.

Table 3. SDQ internal consistency scores and correlations with CBCL.

Variable	SDQ emotion	SDQ conduct	SDQ hyper/ inattention	SDQ peer problems	SDQ prosocial	SDQ internalizing	SDQ externalizing	SDQ total
SDQ Cronbach's α	0.64	0.57	0.76	0.73	0.73	0.64	0.76	0.85
SDQ McDonald's ω	0.73	0.60	0.82	0.54	0.67	0.72	0.81	0.82
CBCL Emotionally reactive	0.52	0.44	0.31	0.16	–0.27	0.45	0.43	0.52
CBCL Anxious/depressed	0.56	0.18	0.22	0.19	–0.22	0.50	0.23	0.40
CBCL Somatic complaints	0.34	0.23	0.15	0.15	–0.10 ^{ns}	0.50	0.23	0.40
CBCL Withdrawn	0.30	0.31	0.31	0.26	–0.29	0.37	0.35	0.44
CBCL Sleep problems	0.31	0.25	0.20	0.17	–0.14	0.32	0.26	0.35
CBCL Attention problems	0.32	0.34	0.64	0.29	–0.23	0.38	0.61	0.50
CBCL Aggressive behavior	0.40	0.65	0.49	0.19	–0.38	0.39	0.64	0.64
CBCL Internalizing	0.59	0.38	0.32	0.24	–0.29	0.56	0.40	0.56
CBCL Externalizing	0.42	0.62	0.57	0.23	–0.37	0.42	0.69	0.69
CBCL Total problems	0.55	0.52	0.49	0.28	–0.36	0.55	0.58	0.68

$N = 289$. $r > 0.55$ are marked in bold.

ns: non-significant

The three-factor structure including internalizing problems, externalizing problems, and prosocial behavior also showed acceptable fit: RMSEA = 0.062, CFI = 0.843, TLI = 0.827.

Convergent and divergent validity

Spearman's correlations show high convergence between SDQ 2–4 and CBCL scores on several subscales (Table 3). The subscales measuring externalizing problems have somewhat higher correlations than those of internalizing problems, the highest being the correlation between the Externalizing composite scores of both instruments ($\rho = 0.69$). The prosocial subscale of SDQ 2–4 yielded negative correlations with all CBCL scales (range -0.10 to -0.38).

Sex differences were evident on all the scales of the SDQ 2–4 except for the Conduct problems subscale (total score: $t = 3.80$, $p < .001$. Externalizing score: $t = 3.40$, $p < .001$. Internalizing score: $t = 2.98$, $p = .003$), boys showing higher problem scores. However, effect sizes were low to moderate with Cohen's d ranging from 0.23 to 0.45 [66].

Discussion

This study examined psychometric properties of the SDQ 2–4 in a sample of Norwegian preschool children aged 1–5 years. Satisfactory internal consistency for the total score was established, though scores for the subscales varied. Both the five-factor and three-factor structures were found to be acceptable, the five-factor structure having a slightly better fit. Convergent and divergent validity scores, using CBCL scores as comparison, were also satisfactory, sex differences were found on some subscales, boys being reported to have higher problem scores.

Internal consistency

The relatively high internal consistency coefficient for the total score, and the more diverse scores for the subscales, align well with previous studies using SDQ 2–4 in other preschool populations. Many studies have reported $\alpha > 0.70$ for the total score and with a wide variability among the subscales [33, 36, 38–40]. However, some also report satisfactory internal consistency for the subscales; for example, D'Souza et al. [34] reported $\alpha > 0.70$ for all subscales except the Peer problems subscale, in a large sample of 2-year-olds. It has been suggested that although a few items have been altered from the SDQ 4–17 version, some items might still be less appropriate for the preschool age than for older children [36]. Combined with only a few items per subscale, problematic items may have a substantial impact on internal consistency scores. That said, it should also be noted that both Cronbach's α and McDonald's ω rely on the assumption that the scales are unidimensional; thus, the subscales, more likely to be unidimensional than composite scores such as the total score, could be regarded as a more reliable estimate of internal consistency. McDonald's ω is also considered more robust against violations of the assumption of unidimensionality [41].

In sum, the total problems score holds good internal consistency as do most of the subscales, however, the conduct problems subscale should be used with caution.

Factor structure

Our study found good to acceptable model fit for the five-factor structure, in line with previous studies [30, 35]. The three-factor structure yielded a slightly lower fit, yet still acceptable. In school-age children, the internalizing and externalizing scales are recommended for screening in low-risk samples, whereas the five subscales might be more useful in high-risk populations [42]. As our data yielded acceptable fit for both models, and more complex models tend to be more specific, it is worth exploring whether the three-factor model could be recommended in screening in low-risk samples and the five-factor model is more useful in high-risk samples. Future studies should test this notion.

Convergent and divergent validity

Positive correlations were found between several SDQ 2–4 subscales and CBCL subscales intended to measure the same construct, and negative correlations between constructs that are considered to be opposite of each other. These results are consistent with previous studies [35, 38, 39]. Higher correlations were found for scales that address externalizing difficulties, compared to scales that tap into internalizing difficulties. It should be noted that the wording in the externalizing items in the two instruments are quite similar, which could result in more overlapping scores. It is also possible that externalizing difficulties are more apparent in children's behavior compared to internalizing difficulties, thus making the scoring of such difficulties easier to operationalize.

Across all subscales, boys were rated higher than girls on all the problem scales and lower than girls on the prosocial subscale. Notably, effect sizes were low to moderate, and it is uncertain whether these differences are clinically significant. The difference in scores could reflect actual differences between boys and girls, however it is also possible that parents interpret behavior differently in boys and girls, thus having different response patterns when reporting the same behavior [67]. Measurement invariance has not been addressed in this study and should be addressed in future studies. Still, the findings corroborate the construct validity of the SDQ 2–4 as similar sex differences are reported in preschool children not only in studies that use SDQ 2–4 [32, 34–36, 40] but also with studies using diagnostic assessments and observational methods [45, 46]. Future studies should further investigate sex differences at different ages, not least regarding norms and cut-off scores. To some extent, one might argue that sex-specific norms are more accurate; on the other hand, scoring might be more cumbersome and thus undermine the intentions of effective screening procedures. Moreover, a higher cut-off score for boys would imply that they would have to exhibit more difficulties than girls would, to qualify for intervention. This would in many cases violate the aim of equitable access to healthcare [68].

It should be noted that studies who use the original version intended for school-age children, the SDQ 4–17, have also reached satisfactory psychometric properties in pre-school populations [3, 40]. Because the SDQ 2–4 and the SDQ 4–17 only differs by three items, in which two are changed and one is slightly altered, it is possible that the 4–17 version could be used for the whole age range from 2 to 17, for easier administration. We are not aware of any studies that compare the two versions in the same sample. However, even with similar psychometric properties, one should also acknowledge how the questionnaire is accepted by the parent. For example, the item ‘often lies or cheats’ can be difficult to report for a 2-year-old, whereas ‘often argumentative with adults’ might be easier to observe, even in children with limited language abilities.

To summarize, there are several indications of good construct validity for SDQ 2–4. Constructs are meaningfully correlated with corresponding constructs in CBCL, subscales that are intended to measure separate constructs are negatively or non-significantly correlated. Our findings align well with evaluations of SDQ 2–4 in other populations.

Strengths and limitations

This study is the first to assess psychometric properties of SDQ 2–4 in a Norwegian sample. Sample size was close to $N=300$, in line with recommendations for psychometric evaluations [69]. Yet, the study also has some limitations that need to be addressed.

Efforts were made to ensure recruit a representative sample of the Norwegian population, by including both rural and urban municipalities and mainly recruiting through kindergartens in which over 90% of all children between 1 and 6 years are enrolled [47]. Still, parents in our sample had higher education than the Norwegian population, thus suggesting a ‘supernormal sample’ as has been observed also in other studies [22]. If we assume that education level is associated with socioeconomic status (SES), we expect that low-SES families are not well represented in our study. The majority of respondents were mothers, and although we did not find any differences in mothers’ and fathers’ ratings which align with the findings of Dahlberg et al. [31], future studies should aim to include a larger proportion of fathers. Our recruitment procedure did not enable us to obtain information about non-respondents. Thus, we do not have any data on response rate or possible differences between respondents and non-respondents.

The CBCL was used as a comparison measure to assess the validity of the SDQ 2–4. The CBCL is commonly used in clinical work as well as in research. Notably, it is often used as the reference criterion for screening tools [52, 70]. Nonetheless, it is not a diagnostic tool and importantly, the preschool version of CBCL has not undergone proper psychometric evaluation in Norway. Future studies using other reference criteria, for example diagnostic interviews, would be needed to investigate criterion validity and to determine cut-off scores for the Norwegian population; however, this was not feasible in this study.

Conclusion

Our study presents psychometric properties of the SDQ 2–4 in a Norwegian sample of preschool children aged 1–6 years. Even though methodological limitations apply, the findings still suggest that the Norwegian translation of the SDQ 2–4 is promising and can be recommended as a tool for screening for children at risk for socioemotional risk. Internal consistency scores suggest that the total problem score should be used, and that the subscales must be interpreted with caution.

Acknowledgments

The authors wish to thank the parents who participated in the study.

Disclosure statement

The authors report there are no competing interests to declare.

Notes on contributors

Nina Jakhelln Laugen is a clinical psychologist and associate professor. Research interests are emotional development, hearing loss and competence development in psychotherapists.

Hege Midtli is a clinical psychologist with background from department of psychology, NTNU. Her current clinical practice is with children and adolescents.

Ulrika Löfkvist is Associate professor and senior lecturer. Her current research includes etiology and developmental outcomes in children and adolescents who are deaf or hard-of-hearing.

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

Funding

No funding was received.

ORCID

Nina Jakhelln Laugen  <http://orcid.org/0000-0001-7920-2585>

Data availability statement

The dataset is available at DataverseNO: <https://doi.org/10.18710/YAC3ZP>
The analytic code necessary to reproduce the analyses is available from the first author.

References

- [1] 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.
- [2] Drugli MB, Stensen K. Psykiske vansker hos norske barnehagebarn. *Første Steg*. 2019;1:46–49.
- [3] Sveen TH, Berg-Nielsen TS, Lydersen S, et al. Detecting psychiatric disorders in preschoolers: screening with the Strengths and

- Difficulties Questionnaire. *J Am Acad Child Adolesc Psychiatry*. 2013;52(7):728–736. doi: [10.1016/j.jaac.2013.04.010](https://doi.org/10.1016/j.jaac.2013.04.010).
- [4] Caspi A, Moffitt TE, Newman DL, et al. Behavioral observations at age 3 years predict adult psychiatric disorders: longitudinal evidence from a birth cohort. *Arch Gen Psychiatry*. 1996;53(11):1033–1039. doi: [10.1001/archpsyc.1996.01830110071009](https://doi.org/10.1001/archpsyc.1996.01830110071009).
- [5] D'Souza S, Underwood L, Peterson ER, et al. Persistence and change in behavioural problems during early childhood. *BMC Pediatr*. 2019;19(1):259. doi: [10.1186/s12887-019-1631-3](https://doi.org/10.1186/s12887-019-1631-3).
- [6] Klein AM, Schlesier-Michel A, Otto Y, et al. Latent trajectories of internalizing symptoms from preschool to school age: a multi-informant study in a high-risk sample. *Dev Psychopathol*. 2019;31(2):657–681. doi: [10.1017/S0954579418000214](https://doi.org/10.1017/S0954579418000214).
- [7] Lavigne JV, Arend R, Rosenbaum D, et al. Psychiatric disorders with onset in the preschool years: i. Stability of diagnoses. *J Am Acad Child Adolesc Psychiatry*. 1998;37(12):1246–1254. doi: [10.1097/00004583-199812000-00007](https://doi.org/10.1097/00004583-199812000-00007).
- [8] Morken IS, Viddal KR, Ranum B, et al. Depression from preschool to adolescence – five faces of stability. *J Child Psychol Psychiatry*. 2021;62(8):1000–1009. doi: [10.1111/jcpp.13362](https://doi.org/10.1111/jcpp.13362).
- [9] National Scientific Council on the Developing Child. Establishing a level foundation for life: mental health begins in early childhood: working paper 6. 2008–2012. Available from: <https://developingchild.harvard.edu/resources/establishing-a-level-foundation-for-lifemental-health-begins-in-early-childhood/>
- [10] Halperin JM, Bédard A-CV, Curchack-Lichtin JT. Preventive interventions for ADHD: a neurodevelopmental perspective. *Neurotherapeutics*. 2012;9(3):531–541. doi: [10.1007/s13311-012-0123-z](https://doi.org/10.1007/s13311-012-0123-z).
- [11] Luby JL, Lenze S, Tillman R. A novel early intervention for preschool depression: findings from a pilot randomized controlled trial. *J Child Psychol Psychiatry*. 2012;53(3):313–322. doi: [10.1111/j.1469-7610.2011.02483.x](https://doi.org/10.1111/j.1469-7610.2011.02483.x).
- [12] Luby JL, Barch DM, Whalen D, et al. A randomized controlled trial of parent-child psychotherapy targeting emotion development for early childhood depression. *Am J Psychiatry*. 2018;175(11):1102–1110. doi: [10.1176/appi.ajp.2018.18030321](https://doi.org/10.1176/appi.ajp.2018.18030321).
- [13] Fält E, Sarkadi A, Fabian H. Exploring nurses', preschool teachers' and parents' perspectives on information sharing using SDQ in a Swedish setting – A qualitative study using grounded theory. *PLoS One*. 2017;12(1):e0168388. doi: [10.1371/journal.pone.0168388](https://doi.org/10.1371/journal.pone.0168388).
- [14] Sheldrick RC, Merchant S, Perrin EC. Identification of Developmental-Behavioral problems in primary care: a systematic review. *Pediatrics*. 2011;128(2):356–363. doi: [10.1542/peds.2010-3261](https://doi.org/10.1542/peds.2010-3261).
- [15] Stokholm JR, Lykke K. The Strengths and Difficulties Questionnaire is a usable way to address mental health at well-child visits in general practice - a qualitative study of feasibility. *BMC Fam Pract*. 2020;21(1):126. doi: [10.1186/s12875-020-01156-3](https://doi.org/10.1186/s12875-020-01156-3).
- [16] Vallance AK, Kramer T, Churchill D, et al. Managing child and adolescent mental health problems in primary care: taking the leap from knowledge to practice. *Prim Health Care Res Dev*. 2011;12(4):301–309. doi: [10.1017/S1463423611000338](https://doi.org/10.1017/S1463423611000338).
- [17] Weitzman C, Wegner L, Blum NJ. Promoting optimal development: screening for behavioral and emotional problems. *Pediatrics*. 2015;135(2):384–395. doi: [10.1542/peds.2014-3716](https://doi.org/10.1542/peds.2014-3716).
- [18] Peltonen K, Kurki M, Reedt C, et al. Psychological tests for expectant parents and young children in the Nordic countries: a review of the evidence. *Eur J Dev Psychol*. 2023;20(3):401–435. doi: [10.1080/17405629.2022.2067141](https://doi.org/10.1080/17405629.2022.2067141).
- [19] Luby JL. Preschool depression: the importance of identification of depression early in development. *Curr Dir Psychol Sci*. 2010;19(2):91–95. doi: [10.1177/0963721410364493](https://doi.org/10.1177/0963721410364493).
- [20] Goodman R, Ford T, Simmons H, et al. Using the Strengths and Difficulties Questionnaire (SDQ) to screen for child psychiatric disorders in a community sample. *Br J Psychiatry*. 2000;177(6):534–539. doi: [10.1192/bjp.177.6.534](https://doi.org/10.1192/bjp.177.6.534).
- [21] SDQinfo. Scoring the strengths & difficulties questionnaire for 2-4 year olds. Youthmind. 2015. Available from: <https://www.sdqinfo.org/py/sdqinfo/c0.py>.
- [22] Goodman R. Psychometric properties of the Strengths and Difficulties Questionnaire. *J Am Acad Child Adolesc Psychiatry*. 2001;40(11):1337–1345. doi: [10.1097/00004583-200111000-00015](https://doi.org/10.1097/00004583-200111000-00015).
- [23] Goodman R. The Strengths and Difficulties Questionnaire: a research note. *J Child Psychol Psychiatry*. 1997;38(5):581–586. doi: [10.1111/j.1469-7610.1997.tb01545.x](https://doi.org/10.1111/j.1469-7610.1997.tb01545.x).
- [24] Steinsbekk S, Wichstrøm L. Cohort profile: the Trondheim early secure study (TESS)—a study of mental health, psychosocial development and health behaviour from preschool to adolescence. *Int J Epidemiol*. 2018;47(5):1401–1401i. doi: [10.1093/ije/dyy190](https://doi.org/10.1093/ije/dyy190).
- [25] Stevenson J, Kreppner J, Pimperton H, et al. Emotional and behavioural difficulties in children and adolescents with hearing impairment: a systematic review and meta-analysis. *Eur Child Adolesc Psychiatry*. 2015;24(5):477–496. doi: [10.1007/s00787-015-0697-1](https://doi.org/10.1007/s00787-015-0697-1).
- [26] Burkhart K, Agarwal N, Kim S, et al. A scoping review of Trauma-Informed pediatric interventions in response to natural and biologic disasters. *Children (Basel)*. 2023;10(6):1017. doi: [10.3390/children10061017](https://doi.org/10.3390/children10061017).
- [27] Arnfred J, Svendsen K, Rask C, et al. Danish norms for the Strengths and Difficulties Questionnaire. *Dan Med J*. 2019;66(6):A5546.
- [28] Borg AM, Kaukonen P, Joukamaa M, et al. Finnish norms for young children on the Strengths and Difficulties Questionnaire. *Nord J Psychiatry*. 2014;68(7):433–442. doi: [10.3109/08039488.2013.853833](https://doi.org/10.3109/08039488.2013.853833).
- [29] Borg AM, Kaukonen P, Salmelin R, et al. Reliability of the Strengths and Difficulties Questionnaire among Finnish 4–9-year-old children. *Nord J Psychiatry*. 2012;66(6):403–413. doi: [10.3109/08039488.2012.660706](https://doi.org/10.3109/08039488.2012.660706).
- [30] Croft S, Stride C, Maughan B, et al. Validity of the Strengths and Difficulties Questionnaire in preschool-aged children. *Pediatrics*. 2015;135(5):E1210–E1219. doi: [10.1542/peds.2014-2920](https://doi.org/10.1542/peds.2014-2920).
- [31] Dahlberg A, Ghaderi A, Sarkadi A, et al. SDQ in the hands of fathers and preschool teachers—psychometric properties in a non-clinical sample of 3–5-year-olds [journal article]. *Child Psychiatry Hum Dev*. 2019;50(1):132–141. doi: [10.1007/s10578-018-0826-4](https://doi.org/10.1007/s10578-018-0826-4).
- [32] Dahlberg A, Fält E, Ghaderi A, et al. Swedish norms for the Strengths and Difficulties Questionnaire for children 3–5 years rated by parents and preschool teachers. *Scand J Psychol*. 2020;61(2):253–261. doi: [10.1111/sjop.12606](https://doi.org/10.1111/sjop.12606).
- [33] Doi Y, Ishihara K, Uchiyama M. Reliability of the Strengths and Difficulties Questionnaire in Japanese preschool children aged 4–6 years. *J Epidemiol*. 2014;24(6):514–518. doi: [10.2188/jea.JE20140050](https://doi.org/10.2188/jea.JE20140050).
- [34] D'Souza S, Waldie KE, Peterson ER, et al. Psychometric properties and normative data for the preschool Strengths and Difficulties Questionnaire in two-year-old children. *J Abnorm Child Psychol*. 2017;45(2):345–357. doi: [10.1007/s10802-016-0176-2](https://doi.org/10.1007/s10802-016-0176-2).
- [35] Ezpeleta L, Granero R, de la Osa N, et al. Psychometric properties of the Strengths and Difficulties Questionnaire(3-4) in 3-year-old preschoolers. *Compr Psychiatry*. 2013;54(3):282–291. doi: [10.1016/j.comppsy.2012.07.009](https://doi.org/10.1016/j.comppsy.2012.07.009).
- [36] Maurice-Stam H, Haverman L, Splinter A, et al. Dutch norms for the Strengths and Difficulties Questionnaire (SDQ) – parent form for children aged 2–18 years. *Health Qual Life Outcomes*. 2018;16(1):123. doi: [10.1186/s12955-018-0948-1](https://doi.org/10.1186/s12955-018-0948-1).
- [37] Nielsen LG, Rimvall MK, Clemmensen L, et al. The predictive validity of the Strengths and Difficulties Questionnaire in preschool age to identify mental disorders in preadolescence. *PLoS One*. 2019;14(6):e0217707. doi: [10.1371/journal.pone.0217707](https://doi.org/10.1371/journal.pone.0217707).
- [38] Patel J, Smith R, O'Farrelly C, et al. Assessing behavior in children aged 12–24 months using the Strengths and Difficulties Questionnaire. *Infancy*. 2021;26(5):724–734. doi: [10.1111/inf.12425](https://doi.org/10.1111/inf.12425).
- [39] Theunissen MHC, Vogels AGC, de Wolff MS, et al. Characteristics of the Strengths and Difficulties Questionnaire in preschool children. *Pediatrics*. 2013;131(2):e446–e454. doi: [10.1542/peds.2012-0089](https://doi.org/10.1542/peds.2012-0089).

- [40] Klein AM, Otto Y, Fuchs S, et al. Psychometric properties of the parent-rated SDQ in preschoolers. *Eur J Psychol Assess.* 2013;29(2):96–104. doi: [10.1027/1015-5759/a000129](https://doi.org/10.1027/1015-5759/a000129).
- [41] McNeish D. Thanks coefficient alpha, we'll take it from here. *Psychol Methods.* 2018;23(3):412–433. doi: [10.1037/met0000144](https://doi.org/10.1037/met0000144).
- [42] Goodman A, Lamping DL, Ploubidis GB. When to use broader internalising and externalising subscales instead of the hypothesised five subscales on the Strengths and Difficulties Questionnaire (SDQ): data from British parents, teachers and children. *J Abnorm Child Psychol.* 2010;38(8):1179–1191. doi: [10.1007/s10802-010-9434-x](https://doi.org/10.1007/s10802-010-9434-x).
- [43] Achenbach TM, Rescorla LA. 2000). Manual for the ASEBA preschool forms and profiles. Vol. 30. Burlington (VT): University of Vermont, Research Center for Children, Youth and Families.
- [44] Kornør H, Jozefiak T. Måleegenskaper ved den norske versjonen av Child Behavior Checklist (CBCL). *PTB.* 2012;2(1):1–8. doi: [10.21337/0014](https://doi.org/10.21337/0014).
- [45] Wichstrøm L, Berg-Nielsen TS, Angold A, et al. Prevalence of psychiatric disorders in preschoolers. *J Child Psychol Psychiatry.* 2012;53(6):695–705. doi: [10.1111/j.1469-7610.2011.02514.x](https://doi.org/10.1111/j.1469-7610.2011.02514.x).
- [46] Zahn-Waxler C, Shirtcliff EA, Marceau K. Disorders of childhood and adolescence: gender and psychopathology. *Annu Rev Clin Psychol.* 2008;4(1):275–303. doi: [10.1146/annurev.clinpsy.3.022806.091358](https://doi.org/10.1146/annurev.clinpsy.3.022806.091358).
- [47] Statistics Norway. Kindergartens. Statistics Norway. 2023. Available 05.07.2023 from <https://www.ssb.no/en/utdanning/barnehager/statistikk/barnehager>.
- [48] Statistics Norway. Higher education. Statistics Norway. 2023. Available 30.08.2023 from <https://www.ssb.no/utdanning/utdanningsniva/statistikk/befolkningens-utdanningsniva>.
- [49] Laugen NJ. Replication Data for: "Validation of the Norwegian version of the Parents' Evaluation of Aural/Oral Performance of Children (PEACH+) for children with typical hearing aged 12-72 months" Version V1 [Survey data]. 2023. DataverseNO. Available from: doi: [10.18710/YAC3ZP](https://doi.org/10.18710/YAC3ZP).
- [50] Heyerdahl S. SDQ - Strength and difficulties questionnaire: en orientering om et nytt spørreskjema for kartlegging av mental helse hos barn og unge, brukt i UNGHUBRO, OPPHED og TROFINN [SDQ – strength and difficulties questionnaire: information about a new questionnaire for assessment of mental health in children and adolescents, used in UNGHUBRO, OPPHED and TROFINN]. *Nor J Epidemiol.* 2011;13(1):127–135. doi: [10.5324/nje.v13i1.320](https://doi.org/10.5324/nje.v13i1.320).
- [51] Kornør H, Heyerdahl S. Måleegenskaper ved den norske versjonen av strengths and difficulties questionnaire, foreldrerapport (SDQ-P). *PTB.* 2017;7(1):1–31. [10.21337/0048](https://doi.org/10.21337/0048)
- [52] Stone LL, Otten R, Engels RC, et al. Psychometric properties of the parent and teacher versions of the Strengths and Difficulties Questionnaire for 4- to 12-year-olds: a review. *Clin Child Fam Psychol Rev.* 2010;13(3):254–274. doi: [10.1007/s10567-010-0071-2](https://doi.org/10.1007/s10567-010-0071-2).
- [53] Gross D, Fogg L, Young M, et al. The equivalence of the Child Behavior Checklist/1 1/2-5 across parent race/ethnicity, income level, and language. *Psychol Assess.* 2006;18(3):313–323. doi: [10.1037/1040-3590.18.3.313](https://doi.org/10.1037/1040-3590.18.3.313).
- [54] Ivanova MY, Achenbach TM, Rescorla LA. 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(12):1215–1224. doi: [10.1016/j.jaac.2010.08.019](https://doi.org/10.1016/j.jaac.2010.08.019).
- [55] Corp IBM. IBM SPSS statistics for MacOS x, version 27.0. Armonk (NY): IBM Corp; 2020.
- [56] 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. doi: [10.1111/bjop.12046](https://doi.org/10.1111/bjop.12046).
- [57] Peters GJ. The alpha and the omega of scale reliability and validity: why and how to abandon Cronbach's alpha and the route towards more comprehensive assessment of scale quality. *Eur Health Psychol.* 2014;16:56–69.
- [58] Sijtsma K. On the use, misuse, and the very limited usefulness of Cronbach's alpha. *Psychometrika.* 2009;74(1):107–120. doi: [10.1007/s11336-008-9101-0](https://doi.org/10.1007/s11336-008-9101-0).
- [59] Yang, Y. & Green, S. B. (2011). Coefficient alpha: a reliability coefficient for the 21st century? *J Psychoeduc Assess.* 29(4):377–392 doi: [10.1177/0734282911406668](https://doi.org/10.1177/0734282911406668).
- [60] European Federation of Psychologists' Associations (EFPA). EFPA review model for the description and evaluation of psychological and educational tests, version 4.2.6. 2013. Available from: <https://www.efpa.eu/ongoing-revision-european-test-review-model>
- [61] Revelle W. Psych: procedures for psychological, psychometric, and personality research. Northwestern University, Evanston, Illinois. R package version 2.3.6. 2023. Available from: <https://CRAN.R-project.org/package=psych>
- [62] Muthén LK, Muthén BO. Mplus user's guide. 8th ed. Los Angeles (CA): Muthén & Muthén; 1998–2017.
- [63] DiStefano C, Morgan GB. A comparison of diagonal weighted least squares robust estimation techniques for ordinal data. *Struct Eq Model A Multidisciplin J.* 2014;21(3):425–438., doi: [10.1080/10705511.2014.915373](https://doi.org/10.1080/10705511.2014.915373).
- [64] MacCallum RC, Browne MW, Sugawara HM, et al. Power analysis and determination of sample size for covariance structure modeling. *Psychol Method.* 1996;1(2):130–149. doi: [10.1037/1082-989X.1.2.130](https://doi.org/10.1037/1082-989X.1.2.130).
- [65] Browne C. 1993. Alternative ways of assessing model fit in. In: K. A. Bollen, J. S. Long, editors. *Testing structural equation models.* Newbury Park (CA): Sage; p. 136–162.
- [66] Cohen J. A power primer. *Psychol Bull.* 1992;112(1):155–159. doi: [10.1037/0033-2909.112.1.155](https://doi.org/10.1037/0033-2909.112.1.155).
- [67] Stensen K, Lydersen S, Ranøyen I, et al. Investigating the measurement invariance of the Caregiver-Teacher report form (C-TRF) factors in a Norwegian early childhood education and care context. *J Psychopathol Behav Assess.* 2023. doi: [10.1007/s10862-023-10074-4](https://doi.org/10.1007/s10862-023-10074-4).
- [68] World Health. Organization. Breaking barriers: Towards more gender-responsive and equitable health systems. Geneva, Switzerland: World Health. Organization; 2019. Available from: <https://www.who.int/publications/m/item/breaking-barriers-towards-more-gender-responsive-and-equitable-health-systems>.
- [69] Frost MH, Reeve BB, Liepa AM, et al. What is sufficient evidence for the reliability and validity of Patient-Reported outcome measures? *Value Health.* 2007;10: s 94–S105. doi: [10.1111/j.1524-4733.2007.00272.x](https://doi.org/10.1111/j.1524-4733.2007.00272.x).
- [70] Heian F. Instrumenter til bruk i sped- og småbarnsalder. Instruments used in infancy and toddlerhood]. *Norsk barne- og ungdomspsykoatrisk forening.* 2019. Available from: <https://www.legeforeningen.no/foreningsledd/fagmed/norsk-barne-og-ungdomspsykiatrisk-forening/veiledere/veileder-i-bup/del-1-diagnostikk-ogutredning/instrumenter-som-bidrag-ved-diagnostikk/instrumenter-til-bruk-i-sped-og-smabarnsalder/>.