



Teacher Nominations of Preschool Children at Risk for Mental Health Problems: how False Is a False Positive Nomination and What Make Teachers Concerned?

Kenneth Stensen¹ · Stian Lydersen¹ · Frode Stenseng² · Jan Lance Wallander^{1,3} · May Britt Drugli^{1,4}

Accepted: 3 February 2021 / Published online: 5 March 2021
© The Author(s) 2021

Abstract

Identification attempts in populations with a low prevalence of problems usually result in a considerable number of false positives. Thus, the aim of the current study was to investigate the false positive rate following nomination of developmental concerns by preschool teachers and the reasons for which teachers raise developmental concerns about children who display non-clinical levels of mental health problems. A total of 1430 children aged 1 to 6 years in Norwegian childcare centers were classified as true positive, false positive, true negative, or false negative by comparing preschool teachers' nomination with their ratings on the Caregiver-Teacher Report Form, resulting in 127 (9%) false positives and 1142 (80%) true negatives. Compared to the true negative group, the false positive group received significantly higher scores on internalizing problems, externalizing problems than true negatives, conflict and significantly lower scores on closeness. Children's internalizing and externalizing problems and age were the main factors that increased the likelihood of teachers raising concerns, while increased closeness in the teacher-child relationship reduced the likelihood of being nominated. Children's gender and conflict level were not significant when adjusting for other factors. These findings suggest that preschool teachers' concerns about children's development should not be discarded as the false positive group did show elevated levels of problem behavior and poorer teacher-child relationship compared to the true negative group. Scrutinizing concerns in collaboration with parents and other mental health professionals may be beneficial to ensure healthy development for children with elevated problem levels.

Keywords Preschool children · Teacher nominations · Mental health problems · False positives

Introduction

Globally, approximately 20% of all children experience mental health difficulties (Belfer 2008) and 13% meet the criteria for a psychiatric disorder (Polanczyk et al. 2015). In Norway,

the prevalence rates are slightly lower with 15% to 20% of the preschool children exhibiting some mental health problems (Skogen et al. 2014) and 7% showing a symptom load that would qualify for a psychiatric disorder (Wichstrøm et al. 2012). However, very few preschoolers who meet diagnostic criteria are referred for mental health evaluation or receive treatment (Egger and Angold 2006; Horwitz et al. 2003; Horwitz et al. 2007) with approximately half of all children with behavioral disabilities not being identified before school entry (Glascoe and Marks 2011). Additionally, in Norway, only 10% of four-year-old children with emotional and behavioral difficulties have received help (Wichstrøm et al. 2014).

At the community level, parents and other caregivers, such as preschool teachers, are the only viable source of information regarding young children's development (Sveen et al. 2013). Parents are the primary initiators of contact with health services when there are concerns about a child's development (Ellingson et al. 2004). Usually, the task of identifying children with emotional and behavioral difficulties has been carried out by pediatric practitioners in collaboration with

✉ Kenneth Stensen
kenneth.stensen@ntnu.no

¹ Regional Centre for Child and Youth Mental Health and Child Welfare, Faculty of Medicine and Health Science, Department of Mental Health, Norwegian University of Science and Technology (NTNU), P.O. Box 8905, MTFS, N-7491 Trondheim, Norway

² Department of Education and Lifelong Learning, Faculty of Social and Educational Sciences, Norwegian University of Science and Technology (NTNU) and Queen Maud University College, Trondheim, Norway

³ University of California, Merced, CA, USA

⁴ Center for Studies of Educational Practice, Inland Norway University of Applied Science, Hamar, Norway

parents. However, many cases may go undetected if responsibility is placed solely on parents and pediatric practitioners (Lavigne et al. 2016a). Alternatively, preschool may provide a valuable context to screen for early developmental concerns. Thus, more attention should be directed towards preschool teachers' perceptions of children's difficulties (Poulou 2015) as their observational accuracy may be an important factor in connecting children in need of help with relevant mental health services (Berkhout et al. 2012; Eklund et al. 2009). However, surprisingly little research has been carried out on the ability of preschool teachers to identify children with emotional and behavioral difficulties.

There is an increasing awareness of the great importance of early identification of emotional and behavioral difficulties (Council on Children with Disabilities et al., 2006; Essex et al. 2009; Glascoe and Marks 2011; Njoroge and Bernhart 2011; Radecki et al. 2011). The rapid development occurring in preschool-aged children may place them at risk of developing emotional/internalized problems and behavioral/externalized problems. While some children overcome these difficulties, others struggle to return to a normal developmental trajectory (Essex et al. 2009; Lavigne et al. 1998; Poulou 2015). Early emotional and behavioral difficulties are predictors of later maladjustment, underlining the importance of identifying those children with high levels of internalized and externalized problems or those with continuity in their problem behavior over time (Basten et al. 2016; Briggs-Gowan and Carter 2008; Essex et al. 2009; Fanti and Henrich 2010; Gilliom and Shaw 2004). Well-timed and targeted interventions may disrupt these negative trajectories and enhance the probability of better adjustment (Masten and Cicchetti 2010). This presumes that children in need of intervention are identified at an early stage; however, such identification may be challenging in a period during which children's development proceeds rapidly (Keenan et al. 1998). Given the potentially serious consequences of early difficulties in children's development and lifelong health (Center on the Developing Child Harvard University, 2010), developing procedures to identify those in need of help should be a public health priority (Essex et al. 2009; Sawyer et al. 2013).

In contrast to diagnostic assessment, the main purpose of identifying children at risk through screening is to detect which children are in need of further assessment to decide whether treatment is necessary or not. In other words, screening may be viewed as a first step that indicates the need for a more thorough diagnostic assessment and identifies children who do not meet diagnostic criteria but may still be at risk developmentally. Screening procedures are usually shorter and less costly to administer than diagnostic assessments, and the briefest and most low-cost screening procedure is the nomination method. Simply, this approach involves informants nominating the children who they perceive to meet a given criteria. This can be in the form of either a certain

number of risk factors being present or the informant's perception of developmental concerns. The nomination method may be regarded as pre-screening or as a subjective judgment call as it can direct attention towards subsets of children in need of further screening.

Preschool teachers are a logical source for the developmental screening of children during early development. However, accuracy in identifying children with difficulties has predominantly been investigated for teachers of school-aged children rather than teachers of preschool children. Several studies of children in elementary school have shown that teachers are more likely to elicit concern and exhibit higher precision regarding externalizing symptoms compared with internalizing symptoms (Dwyer et al. 2006; Loades and Mastroyannopoulou 2010; Soles et al. 2008). Furthermore, children identified as needing mental health services show significantly more adjustment problems than their peers (Layne et al. 2006; Roeser and Midgley 1997) and children nominated by a teacher as at-risk for problems differ significantly from non-nominees with respect to academic grades, sociometric status, and social behavior five years after nomination (Ollendick et al. 1990). However, previous studies have reported low to moderate accuracy when teachers are asked to identify children with anxiety and depression difficulties (Dadds et al. 1997; Moor et al. 2000). Similar to elementary school teachers, preschool teachers also tend to under-report internalizing symptoms (e.g., anxiety, depression, and withdrawal) compared with parent reports. In addition, if preschool teachers perceive their relationship with a child as conflictual, they tend to over-report both internalizing and externalizing symptoms (Berg-Nielsen et al. 2012). That said, it has been reported that preschool teachers can identify a considerable portion of children rated with a clinical symptom load (true positives), especially among the oldest preschool children, and leave very few false negative cases (children rated with a clinical symptom load for which teachers have no concerns). However, preschool teachers' nominations can also produce a high rate of false positive cases with about every other nomination identifying a child rated with a non-clinical symptom load (Stensen et al. 2021). Being identified as a false negative, that is having a negative screening test with positive follow-up or presence of problems that are not identified in a timely manner, may deprive children of appropriate help through non-referral for a more thorough assessment. Misclassifications in the form of false positives, that is a positive screening test with negative follow-up results, indicating an absence of problems, may have a stigmatizing effect for the child and create unnecessary anxiety for the child and parents. This will also result in a waste of clinical resources, which ideally should be allocated to those clearly needing services.

The high rates of false positive cases associated with screening in a population with a low base rate of difficulties, such as a normal population of preschool children, remain

problematic (Lavigne et al. 2016a). The ability of screening procedures to classify cases correctly tends to decline as the base rate of the target problem declines, thereby leading to a higher misclassification rate (Lavigne et al. 2016a; Young and Takala 2018). However, some findings suggest that the validity of a false positive screening result can be debated. By drawing from various screening instrument validation studies, Glascoe (2001) recruited 512 parents and their children, age 7 months to 8 years, to undergo screening followed by diagnostic assessment for all. The results showed that the children classified as false positives performed significantly worse on diagnostic measures than the children classified as true negatives (i.e., those who showed negative screening results and an absence of problems). In another study, Jensen and Watanabe (1999) used DSM criteria and symptom checklists, as well as other survey measures, to compare true positive, false positive, true negative, and false negative cases. They found that the false positive cases exhibited higher levels of a range of risk factors than the true negative cases.

Even though reports of estimated accuracy are somewhat mixed, the nomination method can be used to identify a considerable proportion of children with emotional and behavioral problems. Children classified as false positive cases may still present problems, although not necessarily at a clinical level. Indeed, these children have been reported to receive poorer outcome scores on various measures than children classified as true negative (Glascoe 2001; Jensen and Watanabe 1999). In addition, it has been reported that children displaying symptoms of psychopathology experience considerable impairment even when they do not meet DSM criteria for disorders (Angold et al. 1999). Subthreshold conditions may be effective targets for preventive interventions as these can be precursors for disorders later in life (Shankman et al. 2009). Because preschool teachers can potentially play an important role in identifying and helping children with emotional and behavioral problems, the validity of false positive teacher nominations needs to be investigated by comparing the characteristics of those classified as false positive with those classified as true negative. The current study thus sought to test the following hypothesis in a sample of preschool children:

- (H1) Children with false positive teacher nominations will receive higher scores for problem behaviors and lower scores for teacher-child relationship.
- (H2) Children with an elevated but non-clinical, symptom load will have higher odds of a false positive classification.
- (H3) Negative teacher-child relationships, represented by either high levels of conflict or low levels of closeness, will increase the odds of a false positive classification.

Methods

Data were collected from 2012 to 2014 as part of the Children in Central Norway study, which aimed to enhance teacher competence in addressing preschool children's mental health and to improve the quality of relationships between teachers and children. The study was approved by the Regional Committee for Medical and Health Research Ethics.

Procedure and Participants

Parents with children in childcare centers, serving children from age one to six years old, in three municipalities in Central Norway received recruitment letters with information regarding the project as well as an informed consent form. Information was also provided in parent meetings before the project started. The recruitment letter provided the option for parents to consent either by logging in with a personal invitation code or by returning the consent form to the childcare center. Participation was voluntary and parental consent could be withdrawn at any time until the participation registry was deleted without reprisal. Parental consent gave the teacher in the childcare center who was most familiar with their child permission to complete a survey regarding that parent's child. Children are usually enrolled in childcare centers in the autumn, and the data were collected in January the following year. Thus, most teachers would be expected to have known the child for at least a few months. The teachers provided consent electronically with their own invitation codes. Of the invited parents, 1631 (77%) consented to enroll their child in the study, and 169 teachers (7% male) reported on 1431 children (88% of eligible). The gender distribution of the children was 51% boys and 49% girls, and the mean age was 45 months. In the survey, the preschool teachers were first asked to decide whether or not to nominate the child as being subject of developmental concern. Next, the teachers were asked to respond to the Student-Teacher Relationship Scale (STRS) and the Caregiver-Teacher Report Form (C-TRF). The teachers responded for all children in the same sitting.

Measures

Teacher Nomination

Teachers were asked to make a global subjective judgment concerning each child's risk status by answering "yes" or "no" to indicate whether they perceived that the child had any developmental concerns. This question was located at the start of the survey before the standardized questionnaires were presented. If "yes" was answered, teachers could specify their nomination by checking one or more reasons for the nomination, including aggression, attention, emotional, social, motoric, language, and home situation. However, only those

nominated with specifications of aggression, attention, emotional, or social concerns were considered in the analyses to have been nominated at being at risk to match the types of problems addressed in the criterion (see below).

Student-Teacher Relationship Scale (S-TRS)

The S-TRS (Pianta 2001) is a teacher-report form developed to measure teachers' perception of their relationship to a child or student. It contains the subscales *closeness*, *conflict*, and *dependency* with item responses ranging from 1 ("definitely does not apply") to 5 ("definitely applies"). In the current study, only the *closeness* (11 items) and *conflict* (12 items) subscales were used. A total score for each subscale is obtained by summing the individual items, where higher scores in *closeness* indicate a higher degree of warmth in the relationship and higher scores in *conflict* indicate a higher degree of problem interactions in the relationship. The *closeness* and *conflict* subscales have been demonstrated to have high internal consistency ($\alpha = .86$ for *closeness* and $.92$ for *conflict*) and test-retest reliability (4-week $r = .88$ for *closeness* and $.92$ for *conflict*) (Pianta 2001). In addition, these two subscales have been shown to have good concurrent and discriminant validity in Norway and have been shown to have good factor validity in a slightly modified version (Drugli and Hjemdal 2013; Solheim et al. 2012).

Caregiver-Teacher Report Form (C-TRF)

Teachers also completed the C-TRF (Achenbach and Rescorla 2000), which contains 100 items describing problem behaviors for children who are between 1.5 and 5 years old. Each item has three response options: "not true (as far as you know)", "somewhat or sometimes true", and "very often or often true". These answers correspond to scores from zero to two. The C-TRF contains the following subscales: *emotionally reactive* (7 items), *anxious/depressed* (8 items), *withdrawn* (10 items), *somatic complaints* (7 items), *attention problems* (9 items), and *aggressive behavior* (25 items). A total problem score (ranging from zero to 200) can be calculated by adding the scores across all items. In addition, two broadband scales can be calculated by adding certain subscales for *internalizing problems* (*emotionally reactive*, *anxious/depressed*, *withdrawn*, and *somatic complaints*) and *externalizing problems* (*attention problems* and *aggression problems*). The validity, reliability, and factor structure of the C-TRF have demonstrated to be excellent across cultures (de Groot et al. 1994; Ivanova et al. 2007; Ivanova et al. 2010; Ivanova et al. 2011; Koot et al. 1997; Rescorla et al. 2012; Rescorla et al. 2014; Verhulst and Koot 1992). A score at or above the 90th percentile defines the clinical range on the C-TRF total problem score and has been shown to discriminate well between referred and non-referred children (Achenbach and Rescorla 2000). In addition, the parent-reported counterpart to the C-TRF, the Child Behavior Checklist (CBCL), has been

reported to show good correspondence and predictive validity to DSM diagnoses both for preschool children and older children (Bellina et al. 2013; Ebesutani et al. 2010; Krol et al. 2006; de la Osa et al. 2016).

We defined children with a score at or above the 90th percentile on the C-TRF's *Total Problem*, *Internalizing*, or *Externalizing* scale as having a clinical level of mental health problems. In addition, children in the top 2% on at least one C-TRF subscale (except *somatic complaints*, as this scale does not match the concern specification options in the nomination process and, thus, would have created mismatched data) but who were not rated in the clinical range (90th percentile) on the C-TRF broader scales were also considered to be at risk. Consistent with recommendations by Achenbach and Rescorla (2000), this was done because the subscales comprise a smaller and more homogeneous set of problems, which are believed to require more stringent cutoff values to indicate a need for professional help. Thus, we ensured that children scoring very high on a specific set of problems were classified in the clinical range and in need of mental health services, even though they might have scored below the cutoff value on a broader scale. Because teachers tend to score boys higher than girls on the C-TRF (Achenbach and Rescorla 2000; Kristensen et al. 2010), separate norms for girls and boys were used to establish gender-specific cutoffs. The cutoff values used are shown in Table 1.

Statistical Analyses

Before establishing cutoff values defining the clinical range on the C-TRF, one child was excluded due to missing age information. Based on teacher nominations of children with developmental concerns (246/1430 = 17%) and the C-TRF cutoff values described above (161/1430 = 11% in the clinical range), children were placed in one of the following categories: true positive (119/1430 = 8%), false positive (127/1430 = 9%), true negative (1142/1430 = 80%), and false negative (42/1430 = 3%) (Table 2).

Table 1 Sample-specific cutoff values used on the C-TRF to define the clinical range ($N = 1430$)

Scale (Percentile cutoff)	Girls	Boys
Total Problems (90%)	27	33
Internalizing (90%)	9	10
Externalizing (90%)	14	18
Emotionally Reactive (98%)	5	6
Anxious/Depressed (98%)	6	6
Withdrawn (98%)	7	8
Attention (98%)	10	12
Aggression (98%)	20	26

Independent sample t-tests (equal variance not assumed) were carried out to test for group differences between the true negative cases and the false positives cases, followed by two-level (children nested within preschool teachers) binominal logistic regression analyses to investigate the covariates of group membership for the false positive (target group, $n = 127$) compared to the true negative (reference group, $n = 1142$). None of the 1269 total children in these two groups had missing data, and all were thus included in the current study. Due to the relatively low number of false positive cases, the number of covariates were limited to children's age and gender, *S-TRS conflict* score, *S-TRS closeness* score, *C-TRF internalizing problems* score, and *C-TRF externalizing problems* score. The covariates were entered in the analytic model in three blocks to yield unadjusted (single covariate entry), adjusted (single covariate adjusted for children's age and gender), and fully adjusted (full model with all covariates) odds ratio (OR) estimates. The analyses were performed using SPSS25 and STATA16.

Results

The true negative group contained 50% (569/1142) boys, while the false positive group contained 60% (76/127) boys. The mean age for the true negatives was 3.70 years and the mean age for the false positives was 4.25 years. Both gender ($p = .03$) and age ($p = <.001$) yielded significant group differences. In general, as seen in Table 3, the false positive group had a significantly higher mean score on all covariates except the closeness scale, which was significantly lower, all indicating more negative evaluations for this group, as well as larger variation. In addition, when comparing the false positive group's mean scores for internalizing ($M = 5.76$) and externalizing ($M = 8.54$) problems (Table 3) with the clinical cutoff values for the same scales (Table 1), the estimates indicate that, on average, the false positive cases would have to display approximately twice the symptom load to approach a clinical level on the C-TRF.

Table 4 shows that all covariates in the unadjusted and age- and gender adjusted analyses were significantly associated

with group membership, while male gender and conflict were non-significant covariates in the fully adjusted analyses. With the exception of closeness, all covariates (age, male gender, conflict, internalizing problems, and externalizing problems) showed ORs greater than one, indicating as these covariates increased, the chance of being classified as a false positive increased. For closeness, the OR was significant and less than one, indicating that as closeness increased the chance of being classified as a false positive decreased. More specifically, the fully adjusted analysis revealed that for each year of age, the risk of being classified as a false positive increased by 68%, and a one-unit increase in internalizing or externalizing problem score was associated with a 55% and 21% increased risk of being classified as false positive, respectively. A one unit increase on the closeness scale was associated with an 8% decrease in the risk for false positive classification. Male gender was associated with a 1% decrease in the risk of being classified as a false positive, while a one-unit increase on the conflict scale was associated with a 2% increase in the risk of being classified as a false positive. Neither gender nor conflict were significant covariates in the fully adjusted analysis.

Discussion

The current study aimed to investigate, in essence, how false a false positive teacher nomination is for preschool children at risk for mental health problems and to compare the characteristics of those classified as false positive with those classified as true negative. In support of our initial hypotheses, our findings indicate that children in the normal range of the C-TRF who were nominated by preschool teachers with developmental concerns (false positive) were reported to have significantly higher internalizing and externalizing problem scores compared with children who not were nominated. Children identified as false positive cases were also perceived by teachers to have poorer teacher-child relationships, have higher levels of conflict, and have lower levels of teacher-child closeness. Neither child's gender nor degree of conflictual teacher-child relationship was significantly associated with false positive classification when adjusting for other factors. Age, internalizing problems, and externalizing problems increased the risk of false positive classification, while increased closeness. Closeness in the teacher-child relationship reduced the risk of false positive classification.

Children's Internalizing and Externalizing Problems and Teacher-Child Relationships

The findings of the current study support results from previous research on school-aged children, showing that teacher-nominated children differ significantly from their non-nominated peers (Layne et al. 2006; Ollendick et al. 1990;

Table 2 Frequency overview of children's risk status and teacher nominations ($N = 1430$)

		Teacher Nomination Screening	
		Positive	Negative
Criterion C-TRF Clinical Range	Positive	TP, 119 (8%)	FN, 42 (3%)
	Negative	FP, 127 (9%)	TN, 1142 (80%)

Note: TP = true positive, FP = false positive, FN = false negative, TN = true negative

Table 3 Continuous covariates

Covariate (min–max)	True negative (n=1142)	False positive (n=127)
Conflict (8–44)	16.35 (4.17)	20.17 (6.48)
Closeness (22–55)	42.71 (5.29)	39.17 (5.95)
Internalizing problems (0–17)	2.00 (2.30)	5.76 (3.60)
Externalizing problems (0–23)	3.50 (4.33)	8.54 (5.71)

All values are mean (SD); the group difference was significant at $p < .001$ for all covariates

Roeser and Midgley 1997). Although not reaching the clinical cutoff, the teacher-nominated preschool children still received higher scores for internalizing and externalizing problems, and lower scores for teacher-child relationship quality compared to their non-nominated peers (Table 4). These results suggest that even when the teachers’ concerns are classified as false positives and referrals for a more thorough assessment appear to be unnecessary, children may still be in need of extra monitoring and support to ensure that development proceeds normally. Contrary to previous research (Loades and Mastroyannopoulou 2010), internalizing rather than externalizing problems made teachers more likely to report concerns about children’s development, at least for children in the non-clinical range of the C-TRF. Although preschool teachers tend to underreport internalizing problems (Berg-Nielsen et al. 2012), they may be more vigilant in raising concerns when asked specifically to report this type of problem. If they have more experience dealing with externalizing problems than internalizing problems, a lower threshold for deviancy from what they perceive as normal behavior in the internalizing domain may occur, thus causing them to raise concerns more readily when dealing with children with internalizing problems.

As seen in Table 4, characteristics of the children themselves (i.e., age, internalizing problems, externalizing problems) were the main factors that increased the odds of teacher nomination, while teachers’ perception of teacher-child relationships did not have the same impact. This may be because positive relationships can and do develop even in the presence

of problem behavior (Myers and Pianta 2008). For each one-unit increase in the closeness scale, the odds of false positive classification were reduced significantly by 8%, indicating that teachers’ proximity is important when teachers raise concerns. It seems plausible that with increased closeness, teachers are more able to accurately assess development and thus provide more reliable information. Children with externalizing problems are more likely to develop conflicting relationships with their teachers, which may lead to a maladaptive spiral (Sabol and Pianta 2012). This reciprocal relationship has been found in several studies (e.g., Skalická et al. 2015; Zhang and Sun 2011). Internalizing problems also exhibit the same reciprocal relationship with conflict, but neither externalizing nor internalizing problems show this bidirectional effect for teacher-child closeness (Zhang and Sun 2011). These findings may indicate that two distinct mechanisms are involved in these two relationship dimensions. It may be that preschool teachers’ perception of conflict is child-driven, while the perception of closeness is more teacher-driven (Silver et al. 2005). This could explain the stronger link between problem behavior and conflict, while factors such as teachers’ self-efficacy may play a more important role in teacher-child closeness. In this study, the false positive group received significantly higher scores in internalizing and externalizing problems and conflict as well as lower scores in closeness than the true negative group. Even though the reciprocity of children’s problem behaviors and teacher-child conflict has been demonstrated in previous studies (Skalická et al. 2015; Zhang and Sun 2011),

Table 4 Associations of covariates with false positive vs. true negative classification based on teacher nominations of children with developmental concerns

Covariate	Unadjusted OR (CI)	Age- and gender-adjusted OR (CI)	Fully adjusted OR (CI) ^a
Age (years)	1.39 (1.18–1.64) $p < .001$	1.39 (1.19–1.64) $p < .001$	1.68 (1.30–2.17) $p < .001$
Male gender	1.49 (1.01–2.20) $p = .04$	1.52 (1.03–2.24) $p = .04$.99 (.58–1.71) $p = .98$
Conflict	1.17 (1.12–1.22) $p < .001$	1.18 (1.13–1.23) $p < .001$	1.02 (.96–1.09) $p = .55$
Closeness	.87 (.84–.91) $p < .001$.87 (.84–.91) $p < .001$.92 (.87–.97) $P = .002$
Internalizing problems	1.64 (1.49–1.79) $p < .001$	1.63 (1.49–1.79) $p < .001$	1.55 (1.40–1.72) $p < .001$
Externalizing problems	1.22 (1.17–1.27) $p < .001$	1.24 (1.19–1.30) $p < .001$	1.21 (1.14–1.29) $p < .001$

Note: OR = odds ratio (significant associations in bold); CI = 95% confidence interval

^aAdjusted for all covariates in first column

conflictual relationships do not seem to be a source of concern for teachers, although closeness reduced slightly but significantly the odds for stating concern.

Children's Age and Gender

As seen in Table 4, even though boys are more prone to false positive classification, this association was not significant when adjusting for other covariates, such as internalizing and externalizing problems. Although preschool teachers tend to report boys with more problem behaviors than girls on the C-TRF (Achenbach and Rescorla 2000; Kristensen et al. 2010), it does not seem to bias teacher concerns when children fall in the non-clinical range. As mentioned previously, different teacher thresholds for internalizing and externalizing problems may be in play when teachers raise developmental concerns and the same mechanism may also be at play regarding child gender. If teachers' perception of normal behavior differs for boys and girls and they operate with different thresholds for raising concern, it may result in girls and boys having the same odds for teacher nomination, even when boys display more symptoms. Future studies should investigate the interaction effect between gender and type of problems regarding teacher concerns and false positive classification.

Children's age was a significant predictor of teacher nomination for children in the non-clinical range of the C-TRF. Preschool teachers were more likely to raise concerns for older children than for younger children, making a false positive classification more likely with increased age. One explanation may be that teachers feel more capable of discriminating between normal and abnormal behavior for older children, as symptom expression in older children can be more distinguishable than the more subtle expressions in younger children. Thus, teachers may have a better reference base for normal rather than abnormal behavior when reporting concern. Another explanation may be that teachers perceive younger children to have more transient problems that are more likely to normalize before school entry. As children age and school entry approaches, teachers may grow increasingly concerned if there is a dissonance in the perception of developmental skills and school readiness. As the current study shows, false positive cases received significantly higher problem scores than the true negative cases, indicating that teachers' concerns should not be disregarded out-of-hand.

Some behaviors are more likely to be considered deviant or unusual as they raise concerns more easily. However, most behaviors are displayed on a continuum and depend on context, which makes the discrimination of normal and abnormal behavior more difficult to establish (Carter et al. 2004). In addition, internalizing and externalizing problems exhibit a low to modest correlation to functional impairment (Gordon et al. 2006; McKnight and Kashdan 2009; McKnight et al. 2016). Thus, some children may display an elevated symptom

load without significant impairment, while other children may display few symptoms and significant impairment. Consequently, blindly relying on categorical criteria may hinder the identification of children with emerging psychopathology. As children with sub-clinical levels of symptoms continue to display impairment later in life (Finsaas et al. 2018; Shankman et al. 2009), it is important to identify relevant cases below clinical cutoffs and among children not meeting diagnostic criteria. As symptoms of psychiatric disorders reflect normal behaviors that change phenotypically as children develop, approaches that capture the full range of behaviors relevant to psychopathology are needed (Dougherty et al. 2015). A probabilistic approach to clinical cutoff values was proposed by Sheldrick et al. (2015) in which children with a very high symptom score are assumed to be more likely to have some psychopathology than children with a low score. In addition, children approaching a clinical cutoff would also be more likely to display some psychopathology than children with low scores. Thus, this approach assumes that increases in symptom scores indicate an increased probability of psychopathology. In support of Sheldrick et al. (2015), the current study reports that increases in internalizing and externalizing problems were associated with an increased likelihood of nomination by preschool teachers. Further, the nominated children were reported to have significantly more symptoms of problematic behavior compared to the children not nominated, indicating that children classified as false positive through preschool teachers' nominations should be developmentally monitored rather than regarded as completely behaviorally healthy. Our results and those of previous studies indicate that, although both categorical and dimensional approaches to psychopathology are capable of discriminating between normal and abnormal behaviors in preschoolers (Moreland and Dumas 2008), dimensional approaches are better suited for monitoring developmental trajectories because of their flexibility (Coghill and Sonuga-Barke 2012). Consequently, dimensional approaches to identify and intervene for sub-clinical problems should be initiated before more stable patterns of psychopathology emerge.

Clinical Implications

It is reasonable to assume that a false positive classification may be more correct for some children and less correct for others. Further, better training for preschool teachers so that they can recognize what constitutes developmental concerns, might be expected to lower teacher nomination-associated misclassification rates. Findings from this study suggest that the term *false positive* may in some cases be misleading or inaccurate, which could potentially hinder at-risk children getting appropriate help in a timely manner. When a preschool teacher raises concerns that a child may potentially have clinical problems, it is important and worthwhile to scrutinize the

concern in a prompt manner even for children who are found to be in the non-clinical range following further tests. Further, preschool teachers should be encouraged to express their concerns once they arise, preferably in a forum which includes colleagues, parents, and other mental health professionals. The Norwegian Kindergarten Act (2017) states that preschool teachers are responsible for following up concerns they might have regarding a child's development. Thus, one approach may be to apply teacher concerns as a pre-screening method to identify children of interest, supplemented by a psychometrically sound screening instrument, which will either confirm or dismiss the concern. It may also be beneficial to monitor child development dimensionally by scrutinizing scores and/or establishing a symptom profile rather than blindly using categorical cutoffs. This may ensure that children with elevated (but non-clinical) problem scores are monitored and get help for their problems, which in turn may increase the likelihood of healthy development. The importance of this task is underlined by findings indicating that young children who display internalizing and externalizing problems at a sub-clinical level will continue to exhibit poorer functioning throughout childhood and adolescence (Finsaas et al. 2018). Since few preschool children are referred and receive treatment for existing mental health problems (Egger and Angold 2006; Horwitz et al. 2003; Horwitz et al. 2007; Wichstrøm et al. 2014), scrutinizing preschool teachers' concerns may lead to an increase in referral rate. Currently, it is known that parents are a strong trigger for the initiation of contact with health services for young children not meeting developmental expectations within the family context (Ellingson et al. 2004). The results of the current study and previous studies suggest that preschool teachers may play a similar role within the context of childcare centers, as their concerns seem to capture a considerable portion of children with a clinical symptom load (Stensen et al. 2021), as well as those with a sub-clinical symptom load.

Strength and Limitations

To the best of our knowledge, this study is the first to investigate false positive classification rates and the factors that predict teachers' developmental concerns in a large sample which includes the full age range of preschool children. Although this study was strengthened by its large sample size, an even larger sample to obtain more false positive cases would have been beneficial, allowing more covariates to be investigated. Future studies should investigate the interaction of the covariates examined in the current study to further illuminate the factors leading to teachers being concerned about children without obvious clinical problems. One limitation of the current study is that the ordering of measures in the survey may have increased the susceptibility to a

confirmation bias. Preschool teachers nominated children prior to completing the C-TRF, which could have primed them to respond to the survey differently than if the nomination item was dropped or maybe located elsewhere in the survey. Although the C-TRF might be regarded as a "gold standard" for measuring children's mental health problems based on its psychometric properties, it does not exhibit perfect accuracy compared with diagnostic interviews (Lavigne et al. 2016b). Future studies could investigate preschool teachers' concerns against classifications from a diagnostic interview, thus not solely relying on one informant. The inclusion of other informants, such as parents, would have been beneficial when investigating the nomination method. As preschool teachers provide one perspective, future studies should include other informants to investigate the psychometric indices of the nomination method, as the "gold standard" for clinical decisions in developmental psychopathology use multi-informant reporting. It would also be valuable to investigate the relationship between preschool teachers' concerns and measures of functional impairment, rather than symptom scales isolated. Finally, as this study is cross-sectional and represents only a snapshot, an important focus for future research should be to longitudinally investigate whether scrutinizing preschool teachers' concerns, preferably in combination with a psychometrically sound screening instrument and in collaboration with parents and other mental health professionals, actually leads to increased support for children displaying an elevated level of mental health problems. As children's mental health problems may be a precursor for emerging psychopathology, early identification and support can be of great importance to ensure their healthy development.

Conclusion

The results of the current study emphasize the importance of seriously considering preschool teachers concerns about children in their care, as even children classified as false positive by teacher nomination displayed significantly poorer outcomes compared with true negative cases. Because an elevated level of mental health problems and a decreased quality of teacher-child relationship may be precursors to emerging psychopathology and later maladjustment, it is important to scrutinize concerns when they arise and supplement with a psychometrically sound screening instrument that may confirm or dismiss the initial concern. Even in cases where follow-up instruments dismiss the initial concern, scrutinizing the preschool teachers' concern in collaboration with others may still reveal that the child needs support to reduce the level of problem behavior and thus increase the likelihood of healthy development.

Funding Open access funding provided by NTNU Norwegian University of Science and Technology (incl St. Olavs Hospital - Trondheim University Hospital).

Declarations

Conflict of Interests None of the authors have declared any competing or potential conflicts of interest.

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).

Informed Consent Informed consent was obtained from all parents and teachers of included children.

Experiment Participants All procedures performed in studies involving human participants were conducted in accordance with the ethical standards of the institutional and/or national research committee and with the 1964 Helsinki Declaration and its later amendments or comparable ethical standards.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

- Achenbach, T. M., & Rescorla, L. A. (2000). *Manual for the ASEBA Preschool Forms & Profiles*. Burlington: University of Vermont, Research Center for Children, Youth, & Families.
- Angold, A., Costello, E. J., Farmer, E. M. Z., Burns, B. J., & Erkanli, A. (1999). Impaired but undiagnosed. *Journal of the American Academy of Child & Adolescent Psychiatry*, 38, 129–137.
- Basten, M., Tiemer, H., Althoff, R. R., van de Schoot, R., Jaddoe, V. W. V., Hofman, A., et al. (2016). The stability of problem behavior across the preschool years: An empirical approach in the general population. *Journal of Abnormal Child Psychology*, 44, 393–404.
- Belfer, M. L. (2008). Child and adolescent mental disorders: The magnitude of the problem across the globe. *Journal of Child Psychology and Psychiatry*, 49, 226–236.
- Bellina, M., Brambilla, P., Garzitto, M., Negri, G. A. L., Molteni, M., & Nobile, M. (2013). The ability of CBCL DSM-oriented scales to predict DSM-IV diagnoses in a referred sample of children and adolescent. *European Child and Adolescent Psychiatry*, 22, 235–246.
- Berg-Nielsen, T. S., Solheim, E., Belsky, J., & Wichstrøm, L. (2012). Preschoolers' psychosocial problems: In the eyes of the beholder? Adding Teacher Characteristics as Determinants of Discrepant Parent-Teacher Reports. *Child Psychiatry and Human Development*, 43, 393–413.
- Berkhout, L., Hoekman, J., & Goorhuis-Brouwer, S. M. (2012). Psychosocial health of 4-year-old boys and girls, as observed by parents and teachers. *European Journal of Developmental Psychology*, 9, 500–507.
- Briggs-Gowan, M. J., & Carter, A. S. (2008). Social-emotional screening status in early childhood predicts elementary school outcomes. *Pediatrics*, 121, 957–962.
- Carter, A. S., Briggs-Gowan, M. J., & Davis, N. O. (2004). Assessment of young children's development and psychopathology: Recent advances and recommendations for practice. *Journal of Child Psychology and Psychiatry*, 45, 109–134.
- Center on the Developing Child at Harvard University (2010). The Foundations of Lifelong Health Are Built in Early Childhood. Retrieved from <http://www.developingchild.harvard.edu>
- Coghill, D., & Sonuga-Barke, E. J. S. (2012). Annual research review: Categories versus dimensions in the classification and conceptualisation of child and adolescent mental disorders – Implications of recent empirical study. *Journal of Child Psychology and Psychiatry*, 53, 469–489.
- Council on Children With Disabilities. (2006). Identifying infants and Young children with developmental disorders in the medical home: An algorithm for developmental surveillance and screening. *Pediatrics*, 118, 405–420.
- Dadds, M. R., Spence, S. H., Holland, D. E., Barrett, P. M., & Laurens, K. R. (1997). Prevention and early intervention for anxiety disorders: A controlled trial. *Journal of Consulting and Clinical Psychology*, 65, 627–635.
- Dougherty, L. R., Leppert, K. A., Merwin, S. M., Smith, V. C., Bufferd, S. J., & Kushner, M. R. (2015). Advances and directions in preschool mental Health Research. *Child Development Perspectives*, 9, 14–19.
- Drugli, M. B., & Hjemdal, O. (2013). Factor structure of the student-teacher relationship scale for Norwegian school-aged children explored with confirmatory factor analysis. *Journal of Educational Research*, 57, 457–466.
- Dwyer, S. B., Nicholson, J. M., & Battistutta, D. (2006). Parent and teacher identification of children at risk of developing internalizing or externalizing mental health problems: A comparison of screening methods. *Prevention Science*, 7, 343–357.
- Ebesutani, C., Bernstein, A., Nakamura, B. J., Chorpita, B. F., Higa-McMillan, C. K., Weisz, J. R., & the Research Network on Youth Mental Health. (2010). Concurrent validity of the child behavior checklist DSM-oriented scales: Correspondence with DSM diagnoses and comparison to syndrome scales. *Journal of Psychopathology and Behavioral Assessment*, 32, 373–384.
- Egger, H. L., & Angold, A. (2006). Common emotional and behavioral disorders in preschool children: Presentation, nosology, and epidemiology. *Journal of Child Psychology and Psychiatry*, 47, 313–337.
- Eklund, K., Renshaw, T. L., Dowdy, E., Jimerson, S. R., Hart, S. R., Jones, C. N., & Earhart, J. (2009). Early identification of behavioral and emotional problems in youth: Universal screening versus teacher-referral identification. *The California School Psychologist*, 14, 89–95.
- Ellingson, K. D., Briggs-Gowan, M. J., Carter, A. S., & Horwitz, S. M. (2004). Parent Identification of Early Emerging Child Behavior Problems. *Archives of Pediatric and Adolescent Medicine*, 158, 766–772.
- Essex, M. J., Kraemer, H. C., Slattery, M. J., Burk, L. R., Boyce, W. T., Woodward, H. R., & Kupfer, D. J. (2009). Screening for childhood mental health problems: Outcomes and early identification. *Journal of Child Psychology and Psychiatry*, 50, 562–570.
- Fanti, K. A., & Henrich, C. C. (2010). Trajectories of pure and co-occurring internalizing and externalizing problems from age 2 to age 12: Findings from the National Institute of Child Health and

- Human Development study of early child care. *Developmental Psychology*, 46, 1159–1175.
- Finsaas, M. C., Kessel, E. M., Dougherty, L. R., Bufferd, S. J., Danzig, A. P., Davila, J., Carlson, G. A., & Klein, D. N. (2018). Early childhood psychopathology prospectively predicts social functioning in early adolescence. *Journal of Clinical Child & Adolescence Psychology*, 49, 353–364. <https://doi.org/10.1080/15374416.2018.1504298>.
- Gilliom, M., & Shaw, D. S. (2004). Codevelopment of externalizing and internalizing problems in early childhood. *Development and Psychopathology*, 16, 313–333.
- Glascoc, F. P. (2001). Are Overreferrals on developmental screening tests really a problem? *Archives of Pediatric and Adolescent Medicine*, 155, 54–59.
- Glascoc, F. P., & Marks, K. P. (2011). Detecting children with developmental-behavioral problems: The value of collaborating with parents. *Psychological Test and Assessment Modeling*, 53, 258–279.
- Gordon, M., Antshel, K., Faraone, S., Barkley, R., Lewandowski, L., Hudziak, J. J., Biederman, J., & Cunningham, C. (2006). Symptoms versus impairment: The case of respecting DSM-IV's criterion D. *Journal of Attention Disorders*, 9, 465–475.
- de Groot, A., Koot, H. M., & Verhulst, F. C. (1994). Cross-cultural generalizability of the child behavior checklist cross-informant syndromes. *Psychological Assessment*, 6, 225–230.
- Horwitz, S. M., Gary, L. C., Briggs-Gowan, M. J., & Carter, A. S. (2003). Do needs drive service use in Young children? *Pediatrics*, 112, 1373–1378.
- Horwitz, S. M., Kelleher, K. J., Stein, R. E. K., Storfer-Isser, A., Youngstrom, E. A., Park, E. R., et al. (2007). Barriers to the identification and Management of Psychosocial Issues in children and maternal depression. *Pediatrics*, 119, 208–218.
- Ivanova, M. Y., Achenbach, T. M., Rescorla, L. A., Dumenci, L., Almqvist, F., Bathiche, M., Bilenberg, N., Bird, H., Domuta, A., Erol, N., Fombonne, E., Fonseca, A., Frigerio, A., Kanbayashi, Y., Lambert, M. C., Leung, P., Liu, X., Minaei, A., Roussos, A., Simsek, Z., Weintraub, S., Wolanczyk, T., Zubrick, S., Zukauskienė, R., & Verhulst, F. (2007). Testing the Teacher's report form syndromes in 20 societies. *School Psychology Review*, 36, 468–483.
- Ivanova, M. Y., Achenbach, T. M., Rescorla, L. A., Harder, V. S., Ang, R. P., Bilenberg, N., et al. (2010). Preschool psychopathology reported by parents in 23 societies: Testing the seven-syndrome model of the child behavior checklist for ages 1.5–5. *New Research*, 49, 1215–1224.
- Ivanova, M. Y., Achenbach, T. M., Rescorla, L. A., Bilenberg, N., Bjamadottir, G., Denner, S., et al. (2011). Syndromes of preschool psychopathology reported by teachers and caregivers in 14 societies using the caregiver-teacher report form (C-TRF). *Journal of Early Childhood and Infant Psychology*, 7, 87–103.
- Jensen, P. S., & Watanabe, H. (1999). Sherlock Holmes and child psychopathology assessment approaches: The case of the false-positive. *Journal of American Academy of Child and Adolescent Psychiatry*, 38, 138–146.
- Keenan, K., Shaw, D., Delliquadri, E., Giovannelli, J., & Walsh, B. (1998). Evidence for the continuity of early problem behaviors: Application of a developmental model. *Journal of Abnormal Child Psychology*, 26, 441–454.
- Koot, H., van den Oord, J., Verhulst, F. C., & Boomsma, D. (1997). Behavioral and emotional problems in young preschoolers: Cross-cultural testing of the validity of the child behavior checklist/2-3. *Journal of Abnormal Child Psychology*, 25, 183–196.
- Kristensen, S., Henriksen, T. B., & Bilenberg, N. (2010). The child behavior checklist for ages 1.5–5 (CBCL/1½–5): Assessment and analysis of parent and caregiver-reported problems in a population-based sample of Danish preschool children. *Nordic Journal of Psychiatry*, 64, 203–209.
- Krol, N. P. C. M., De Bruyn, E. E. J., Coolen, J. C., & van Aarle, E. J. M. (2006). From CBCL to DSM: A comparison of two methods to screen for DSM-IV diagnoses using CBCL data. *Journal of Clinical Child and Adolescent Psychology*, 35, 127–135.
- Lavigne, J. V., Feldman, M., & Meyers, K. M. (2016a). Screening for mental health problems: Addressing the base rate fallacy for a sustainable screening program in integrated primary care. *Journal of Pediatric Psychology*, 41, 1081–1090.
- Lavigne, J. V., Meyers, K. M., & Feldman, M. (2016b). Systematic review: Classification accuracy of behavioral screening measures for use in integrated primary care settings. *Journal of Pediatric Psychology*, 41, 1091–1109.
- Lavigne, J. V., Arend, R., Rosenbaum, D., Binns, H. J., Christoffel, K. K., & Gibbons, R. D. (1998). Psychiatric disorders with onset in the preschool years: I. Stability of Diagnoses. *Journal of American Academy of Child and Adolescent Psychiatry*, 37, 1246–1254.
- Layne, A. E., Bernstein, G. A., & March, J. S. (2006). Teacher awareness of anxiety symptoms in children. *Child Psychiatry and Human Development*, 36, 383–392.
- Loades, M. E., & Mastroyannopoulou, K. (2010). Teachers' recognition of Children's mental health problems. *Child and Adolescent Mental Health*, 15, 150–156.
- Masten, A. S., & Cicchetti, D. (2010). Developmental cascades. *Development and Psychopathology*, 22, 491–495.
- McKnight, P. E., & Kashdan, T. B. (2009). The importance of functional impairment to mental health outcomes: A case for reassessing our goals in depression treatment research. *Clinical Psychology Review*, 29, 243–259.
- McKnight, P. E., Monfort, S. S., Kashdan, T. B., Blalock, D. V., & Calton, J. M. (2016). Anxiety symptoms and functional impairment: A systematic review of the correlation between the two measures. *Clinical Psychology Review*, 45, 115–130.
- Moor, S., Sharrock, G., Scorr, J., McQueen, H., Wrate, R., Cowan, J., & Blair, C. (2000). Evaluation of a teaching package designed to improve teachers' recognition of depressed pupils - a pilot study. *Journal of Adolescence*, 23, 331–342.
- Moreland, A. D., & Dumas, J. E. (2008). Categorical and dimensional approaches to the measurement of disruptive behavior in the preschool years: A meta-analysis. *Clinical Psychology Review*, 28, 1059–1070.
- Myers, S. S., & Pianta, R. C. (2008). Developmental commentary: Individual and contextual influences on student-teacher relationships and Children's early problem behaviors. *Journal of Clinical Child & Adolescent Psychology*, 37, 600–608.
- Njoroge, W. F. M., & Bernhart, K. P. (2011). Assessment of behavioral disorders in preschool-aged children. *Current Psychiatry Reports*, 13, 84–92.
- Norwegian Kindergarten Act (2017). Framework Plan for Kindergartens. Retrieved from: <https://www.udir.no/globalassets/filer/barnehage/rammeplan/framework-plan-for-kindergartens2-2017.pdf>
- Ollendick, T. H., Greene, R. W., Weist, M. D., & Oswald, D. P. (1990). The predictive validity of teacher nominations; A Five-Year Followup of At-Risk Youth. *Journal of Abnormal Child Psychology*, 18, 699–713.
- de la Osa, N., Granero, R., Trepate, E., Domenech, J. M., & Ezpeleta, L. (2016). The discriminative capacity of CBCL/1½–5-DSM5 scales to identify disruptive and internalizing disorders in preschool children. *European Child and Adolescent Psychiatry*, 25, 17–23.
- Pianta, R. C. (2001). *STRS student-teacher relationship scale (professional manual)*. Lutz: Psychological Assessment Resources, Inc.
- Polanczyk, G. V., Salum, G. A., Sugaya, L. S., Caye, A., & Rohde, L. A. (2015). Annual research Review: A meta-analysis of the worldwide prevalence of mental disorders in children and adolescents. *Journal of Child Psychology and Psychiatry*. doi:<https://doi.org/10.1111/jcpp.12381>.

- Poulou, M. S. (2015). Emotional and Behavioural difficulties in pre-school. *Journal of Child and Family Studies*, *24*, 225–236.
- Radecki, L., Sand-Loud, N., O'Connor, K. G., Sharp, S., & Olson, L. M. (2011). Trends in the use of standardized tools for developmental screening in early childhood: 2002–2009. *Pediatrics*, *128*, 14–19.
- Rescorla, L. A., Achenbach, T. M., Ivanova, M. Y., Bilenberg, N., Bjarnadottir, G., Denner, S., Dias, P., Dobrea, A., Döpfner, M., Frigerio, A., Gonçalves, M., Guðmundsson, H., Jusiene, R., Kristensen, S., Lecannelier, F., Leung, P. W. L., Liu, J., Löbel, S. P., Machado, B. C., Markovic, J., Mas, P. A., Esmaeili, E. M., Montiroso, R., Plück, J., Pronaj, A. A., Rodriguez, J. T., Rojas, P. O., Schmeck, K., Shahini, M., Silva, J. R., van der Ende, J., & Verhulst, F. C. (2012). Behavioral/emotional problems of preschoolers: Caregiver/teacher reports from 15 societies. *Journal of Emotional and Behavioral Disorders*, *20*, 68–81.
- Rescorla, L. A., Boichichio, L., Achenbach, T. M., Ivanova, M. Y., Almqvist, F., Begovac, I., et al. (2014). Parent-teacher agreement on Children's problems in 21 societies. *Journal of Clinical Child and Adolescent Psychology*, *43*, 627–642.
- Roeser, R. W., & Midgley, C. (1997). Teachers' views of issues involving students' mental health. *The Elementary School Journal*, *2*, 115–133.
- Sabol, T. J., & Pianta, R. C. (2012). Recent trends in research on teacher-child relationships. *Attachment and Human Development*, *14*, 213–231.
- Sawyer, A. C. P., Chittleborough, C. R., Lynch, J. W., Baghurst, P., Mittinty, M. N., Kaim, A. L. E., & Sawyer, M. G. (2013). Can screening 4–5 year olds accurately identify children who will have teacher reported mental health problems when children are aged 6–7 years? *Australian & New Zealand Journal of Psychiatry*, *00*, 1–10.
- Shankman, S. A., Lewinsohn, P. M., Klein, D. N., Small, J. W., Seeley, J. R., & Altman, S. E. (2009). Subthreshold conditions as precursors for full syndrome disorders: A 15-year longitudinal study of multiple diagnostic classes. *Journal of Child Psychology and Psychiatry*, *50*, 1485–1494.
- Sheldrick, R. C., Benneyan, J. C., Kiss, I. G., Briggs-Gowan, M. J., Copeland, W., & Carter, A. S. (2015). Threshold and accuracy in screening tools for early detection of psychopathology. *The Journal of Child Psychology and Psychiatry*, *56*, 936–948.
- Silver, R. B., Measelle, J. R., Armstrong, J. M., & Essex, M. J. (2005). Trajectories of classroom externalizing behavior: Contributions of child characteristics, family characteristics, and the teacher-child relationship during the school transition. *Journal of School Psychology*, *43*, 39–60.
- Skalická, V., Stenseng, F., & Wichstrøm, L. (2015). Reciprocal relations between student-teacher conflict, children's social skills and externalizing behavior: A three wave longitudinal study from pre-school to third grade. *International Journal of Behavioral Development*, *39*, 413–425.
- Skogen, J. C., Kjeldsen, A., Knudsen, A. K., Myklestad, I., Nesvåg, R., Reneflot, A., & Major, E. (2014). *Psykisk helse hos barn og unge (Folkehelse rapporten 2014:4)*. Oslo: Norwegian Institute of Public Health.
- Soles, T., Bloom, E. L., Heath, N. L., & Karagiannakis, A. (2008). An exploration of teachers' current perceptions of children with emotional and behavioural difficulties. *Emotional and Behavioural Difficulties*, *13*, 275–290.
- Solheim, E., Berg-Nielsen, T. S., & Wichstrøm, L. (2012). The three dimensions of the student-teacher relationship scale: CFA validation in a preschool sample. *Journal of Psychoeducational Assessment*, *30*, 250–263.
- Stensen, K., Lydersen, S., Stenseng, F., Wallander, J. L., Tveit, H., & Drugli, M. B. (2021). Childcare providers' nominations of preschool children at risk for mental health problems: Does it discriminate well compared to the caregiver-teacher report form (C-TRF)? *Scandinavian Journal of Educational Research*, 1–15. <https://doi.org/10.1080/00313831.2020.1869089>.
- Sveen, T. H., Berg-Nielsen, T. S., Lydersen, S., & Wichstrøm, L. (2013). Detecting psychiatric disorders in preschoolers: Screening with the strengths and difficulties questionnaire. *Journal of American Academy of Child and Adolescent Psychiatry*, *52*, 728–736.
- Verhulst, F. C., & Koot, H. M. (1992). *Child psychiatric epidemiology: Concepts, methods and findings*. Beverly Hills: Sage Publications.
- Wichstrøm, L., Berg-Nielsen, T. S., Angold, A., Egger, H. L., Solheim, E., & Sveen, T. H. (2012). Prevalence of psychiatric disorders in preschoolers. *Journal of Child Psychology and Psychiatry*, *53*, 695–705.
- Wichstrøm, L., Belsky, J., Jozefiak, T., Sourander, A., & Berg-Nielsen, T. S. (2014). Predicting service use for mental health problems among Young children. *Pediatrics*, *133*, 1054–1060.
- Young, N. D., & Takala, C. R. (2018). Sequential screening to improve behavioral health needs in primary care. *Journal of American Child and Adolescent Psychiatry*, *57*, 603–609.
- Zhang, X., & Sun, J. (2011). The reciprocal relations between teachers' perception of Children's behavior problems and teacher-child relationships in the first preschool year. *The Journal of Genetic Psychology*, *172*, 176–198.

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.