Original Manuscript

Clinical Child Psychology and Psychiatry
1–23

DOI: 10.1177/1359104518822673
journals.sagepub.com/home/ccp
SAGE

Validation of short instruments assessing parental and caregivers' perceptions on child health and development for personalized prevention

Bernice Doove[1,2], Jolien Feron[3], Frans Feron[2,4], Jim van Os[5,6,7] and Marjan Drukker[5]

[1]Well-Child Care Centre, Preventive Child Health Care Division, Envida, The Netherlands
[2]Department of Social Medicine, CAPHRI School for Public Health and Primary Care, Maastricht University, The Netherlands
[3]Mental Healthcare, MET ggz, The Netherlands
[4]Youth Health Care Division, Regional Public Health Service South Limburg, The Netherlands
[5]Department of Psychiatry and Psychology, MHeNS School for Mental Health and NeuroScience, Maastricht University Medical Centre, The Netherlands
[6]King's Health Partners and Department of Psychosis Studies, Institute of Psychiatry, Psychology & Neuroscience, King's College London, UK
[7]Department of Psychiatry, Brain Center Rudolf Magnus, University Medical Center Utrecht, The Netherlands

Abstract

Systematically exploring parental as well as other caregivers' concerns is a main component in preventive child health care (PCHC) for family-centred practice and personalized health care. To facilitate communication and early identification of emerging mental health problems, a PCHC toolkit based on short instruments was developed. This article investigates the reliability and validity of (1) two visual analogue scales (VAS) to assess parent-reported 'parenting' and 'child behaviour', (2) a professional caregiver-reported VAS to assess 'child competence' and (3) the parents' evaluation of developmental status (PEDS) in Dutch PCHC. Parents as well as child care, kindergarten and preschool teachers completed instruments in a community-based sample of children ($N = 346$) aged 3 years at baseline. The three VAS and PEDS were associated with standardized questionnaires assessing the same constructs. Overall predictive accuracy showed: good to excellent for 'parenting' VAS, fair to good for 'child behaviour' VAS and poor for 'child competence' VAS. The PEDS, 'parenting' VAS and 'child behaviour' VAS, demonstrated high sensitivity at various cut-off points of index test and reference standard. At follow-up, approximately 1 year later, results were similar. Although the 'child competence' VAS scored lower on one aspect of validity, the PEDS and the different VAS are reliable, valid and useful as brief monitoring tools in daily Dutch PCHC practice.

Corresponding author:
Bernice Doove, Well-Child Care Centre, Preventive Child Health Care Division, Envida, P.O. Box 241, Becanusstraat 9, 6200 AE Maastricht, The Netherlands.
Email: bernice.doove@envida.nl; bernice.doove@maastrichtuniversity.nl

Keywords
Child mental health, personalized preventive child health care, screening, short instruments, PEDS, visual analogue scale, validation

Introduction

Delayed or disrupted childhood development may lead to vulnerabilities and mental disorders during life, with a social and economic impact on society (American Academy of Pediatrics (AAP), 2006; Salum, Polanczyk, Miguel, & Rohde, 2010; Tough et al., 2008). Early identification of developmental problems can break the vicious circle of disadvantage. Supporting healthy development and successful social participation is also more cost effective than treating disorders later (Bayer et al., 2011; Karatsoreos & McEwen, 2013; Sroufe, Egeland, Carlson, & Collins, 2005).

From a public mental health perspective, children with symptoms of mental problems below the threshold of a full-blown disorder represent an important group. These children may have significant clinical needs and are at risk of arrested development, not achieving their potential and capabilities (Nussbaum, 2011). The group of children with symptoms of mental problems may be twice as large as the group of children meeting formal diagnostic criteria for a mental disorder (Carter, Briggs-Gowan, & Davis, 2004; Hielkema, de Winter, de Meer, & Reijneveld, 2011).

The Dutch system for preventive child health care (PCHC) has been established to closely monitor the physical and mental development of all children during routine medical assessments in well-child care clinics offered by PCHC professionals (e.g. doctors, nurses) (van Esso et al., 2010). Monitoring enables professionals to intervene in the earliest phase of emerging problems and disabling symptoms, even before formal criteria for diagnostic classifications are met (Berkovits, O'Brien, Carter, & Eyberg, 2010; Carter et al., 2004; Skovgaard, 2010; Snyderman & Langheier, 2006). It has been suggested, however, that the sensitivity for detecting early mental health problems and symptoms is low (AAP, 2006; Jaspers et al., 2012; Reijneveld, Brugman, Verhulst, & Verloove-Vanhorick, 2004; Theunissen, 2013).

From a dynamic transactional developmental perspective, health can be seen as the ability to adapt and self-manage in presence of social, physical and emotional challenges (Huber et al., 2011; Sameroff, 2010). The individual's social and educational environment including interpersonal relationships is hypothesized to be key to PCHC professionals aiming to provide personalized prevention. Parents are a key determinant of influence on their young children's health. Next to PCHC physical examination, periodic eliciting and addressing parental as well as other caregivers' perception and concerns is a main component in a family-centred practice (Cox, Huntington, Saada, Epee-Bounya, & Schonwald, 2010; Glascoe & Robertshaw, 2007; Vreeswijk, Rijk, Maas, & van Bakel, 2015). Considering the fact that parents know their child best, an optimal communication between child health care providers, parents and other caregivers is essential for indicating the child's health, development and behaviour (Glascoe & Trimm, 2014). Thus, Dutch PCHC has a Public Health task to address and influence school readiness before children enter school (Williams, Okamoto, AAP Council on Early Childhood, & AAP Council on School Health, 2016). From a dynamic transactional developmental perspective, school readiness can be seen as an outcome measure for preschool child health (High, 2008).

To improve early identification of health problems, PCHC needs to focus on 'predict it and personalize it' instead of 'find it and fix it'. Signals and problems meeting formal criteria for disorders represent the late stage of a dynamic process that can be identified in a much earlier phase when treatment plasticity is still considerable (High, 2008; Huisman et al., 2008; Marks et al., 2008; McGorry & van Os, 2013). A screening toolkit including short instruments assessing

multiple constructs was developed to use as a short first step in PCHC mental health screening and shared decision making (AAP, 2006; Glascoe, 2005; Marks et al., 2008; Radecki, Sand-Loud, O'Connor, Sharp, & Olson, 2011; Syurina, Hens, & Feron, 2013).

This article aims to validate the Dutch version of the parents' evaluation of developmental status (PEDS). Research has shown that parent-completed screening tools are highly accurate in detecting true problems, are relatively inexpensive and promote a dialogue about concerns, needs and demands between parents and other caregivers (Kerstjens et al., 2009; Potijk, de Winter, Bos, Kerstjens, & Reijneveld, 2016; Theunissen, Vogels, de Wolff, Crone, & Reijneveld, 2015). In addition, visual analogue scales (VAS) were developed as short instruments on child upbringing (hereafter: parenting), child behaviour and child competence.

In a PCHC setting, screening tests with high sensitivity are important to identify children in the earliest phase of emerging problems and disabling symptoms to prevent 'growing into deficit' (Syurina et al., 2013). A high negative predictive value (NPV) is preferable, it ensures that most children who pass the screening are truly healthy. Over-referrals in a first-stage PCHC screening are no problem, they can benefit from additional preventive monitoring (Glascoe, 2001). Furthermore, PCHC screening and monitoring instruments should (1) easily obtain information in every day PCHC setting; (2) carry out dimensional assessment of symptoms and behaviour; (3) measure the progress of development of young children and their possible determinants of influence; (4) identify general signals and symptoms indicating a possible disruption or imbalance of the educational/parent–child system, not yet related to a specific diagnosis; (5) support communication between PCHC, parents and teachers/employees from child care, kindergarten, preschool or primary school (hereafter: professional caregivers) about their perceptions on health and development; (6) connect to needs and demands of the child and social system around the child and (7) promote shared decision making (Marks, Glascoe, & Macias, 2011; Syurina et al., 2013).

Among the available validated parent-completed screening tools, the PEDS is designed to elicit concerns and facilitate communication between professionals and parents in addressing developmental and behavioural problems in children (Glascoe, 2013). Parents as well as professional caregivers can complete the questionnaire. Ten questions explore several concerns and answers are multiple choice: 'no', 'a little', 'yes'. Subsequently, an open-ended field provides more information. PCHC professionals can use the PEDS in two ways: (1) as an informal means to elicit and to respond to parental concerns and (2) as a developmental screening test. It needs clinical judgement but takes only 5 minutes to categorize the parental or other caregivers' concerns in different developmental domains on the 'PEDS score sheet': global/cognitive, expressive language and articulation, receptive language, fine motor, gross motor, behaviour, social-emotional, self-help, school and health. The 'PEDS Score Form' enables PCHC professionals to make decisions about high, moderate or low risk using the 'PEDS Interpretation Form'. The original PEDS as a screening tool has a sensitivity of 91–97% and a specificity of 73–86% for accuracy of parental concerns in detecting children at high and/or moderate developmental risk (Woolfenden et al., 2014). The PEDS is less time consuming than other instruments, emphasis is on parental and other professional caregivers' opinions, it is validated for clinical samples and general population samples aged between 0 and 8 years and is available in multiple languages. This suggests that the PEDS is an accurate tool for use as an initial screening and monitoring tool in Dutch PCHC, where professionals have to deal with the time constraints of daily practice (Limbos & Joyce, 2011).

Furthermore, to obtain more insight into risk and protective factors for developmental arrest, assessment of parental perception on 'parenting' and 'child behaviour', and professional caregivers' perception on 'child competence' is crucial (NICHD, 2004; Sameroff, 2010). It is known that short instruments like VAS can support communication between health providers and patients about their perceptions on health (Hawker, Mian, Kendzerska, & French, 2011). The authors therefore

developed three different VAS as possible PCHC 'toolkit' short instruments. Original questions were in Dutch. Each VAS is a single-item and continuous scale consisting of a horizontal line 100 mm in length anchored by two verbal descriptors, one for each symptom extreme. The respondent is asked to place a line perpendicular to the VAS line at the point that represents the intensity of the symptom. Using a ruler, the score is determined by measuring the distance (mm) on the 100 mm line between the symptoms extremes, providing a range of scores between 0 mm and 100 mm.

The 'parenting' VAS gives parents the opportunity to express the degree to which they feel competent, secure and happy with raising their child ranging from 0 to 100 (parenting VAS: 0 = I do not manage to raise my child as I wish, 100 = raising my child is up to my expectations). On the 'child behaviour' VAS, parents can evaluate their child's behaviour (child behaviour VAS: 0 = my child is difficult and badly behaved, 100 = my child is very obedient and easy to handle). To address the issue of the child's general functional adaptation to social, physical and emotional challenges, professional caregivers can indicate on the 'child competence' VAS, the degree of competence of the child in general (child competence VAS: 0 = not competent, 100 = very competent). The 'child competence' VAS is related to school readiness and three child outcomes: children have positive social relationships, children acquire and use knowledge and skills and children take appropriate action to meet their needs. These outcomes are based on the assumptions that children of different ages demonstrate these outcomes in different ways with many pathways leading to competence.

Because these VAS were new, validation was required. In addition, the PEDS needed validation after translation into Dutch by the authors (see Methods).

In this article, it is hypothesized that the total amount of parental and professional caregiver concerns is a valid signal for early detection of health problems. To detect problems and signals below the threshold of a full-blown disorder, continuous total problem score of the reference standard was used rather than the frequently used dichotomous clinical-borderline outcomes. To facilitate interpretability of the PEDS and the different VAS in daily PCHC practice, diverse cut-off points of both index test and reference standard were used to calculate different NPVs and receiver operating characteristic curves (ROCs). Eventually, the choice of an appropriate range along horizontal (false positives, 1-specificity) or vertical (sensitivity) axis depends upon the clinical setting. The different VAS cut-off points were set at 10th, 25th and 50th percentile. Higher VAS scores mean less-perceived difficulties. Parental and professional caregiver PEDS cut-off points were 1, 2 or 3 and more concerns. All values above cut-off point were coded as high-perceived difficulty. Total amount and dichotomous 'parental concerns' and 'professional caregivers concerns' variables (any concern yes/no) and a 'PCHC professional decision about developmental risk' variable (high, moderate and low/no developmental risks) were constructed for use in the analyses as well.

The aim of this study is to define psychometric properties of the Dutch PEDS and three VAS about 'parenting', 'child behaviour' and 'child competence' at the age of 3 and 4 years, as a first step in validation of a potential future Dutch PCHC 'toolkit' with parent and other caregiver completed short instruments for multi-axial and multi-informant screening and monitoring of general emerging problems and disabling symptoms.

Methods

Population and procedure

This study was performed as part of the Monitoring Outcome Measurements of child development study (MOM), a prospective observational study within PCHC practice. A community-based sample of children is systematically assessed using multisource and cross-informant repeated measurements designed to identify developmental pathways impacting school readiness as an outcome of

social participation. Children were aged 3 years at baseline. MOM obtained information from parents, professional caregivers and PCHC professionals on baseline risk and environmental factors to track predictive risk indicators for making multi-axial health profiles.

The Maastricht University Medical Centre Medical Ethics Committee (METC) approved the MOM study protocol under registration number MEC 09-04-018/PL. All persons involved provided written informed consent.

Parents of children born in the fourth quarter of 2006, 2007 and 2008 and living in Maastricht and surrounding municipalities were asked to participate in the MOM study. They received written information about the MOM study. A 'parent questionnaire' was included so that parents could see the questions that would be asked. After parents returned the signed consent form and the completed parent questionnaire, the professional caregiver of the participating child received a 'professional caregiver questionnaire' including the name of the participating child. The PCHC doctor of the participating child was then informed and asked to complete (1) a 'PCHC questionnaire' based on the information from the PCHC file and PCHC consultations and (2) the 'PEDS score sheet' and 'PEDS interpretation form'.

Before including participants, managing boards of all PCHC, preschool and day-care organizations in Maastricht and the various surrounding municipalities were contacted, informed and asked for their consent to participate the MOM study. All organizations and their professionals received written information. In total, 46 day-care and 58 preschool classes agreed to participate if parents were willing to participate in the MOM study.

The PEDS was translated into Dutch in collaboration with the original author F. P. Glascoe, by a process of forward translation by a PCHC professional and backward translation by an independent English native speaker (Glascoe, 2013). Content validity and cultural appropriateness of the PEDS were discussed in both parents and professional caregivers. To prevent wording problems, further adaptation of the Dutch PEDS was performed after a pilot with parents and professional caregivers. Ten parents were asked to comment and discuss the different VAS, and the usefulness and feasibility of the Dutch PEDS after translation in Dutch. In total, 36 professional caregivers from 18 different day-care and preschool locations participated in a pilot to investigate the usefulness and feasibility of the MOM study questionnaire including the Dutch PEDS and the different VAS. All PCHC doctors received further training on the PEDS methodology and other instruments and in how to complete the 'PCHC questionnaire'.

Reference standards

To study various types of validity, different validated screening tools for emerging problems and disabling symptoms at the age of 3 and 4 years were included in the parental and professional caregivers MOM questionnaires, to be used as reference standard: Child Behaviour Checklist (CBCL) total score for the child behaviour VAS and parental PEDS concerns, Caregiver-Teacher's Report Form (C-TRF) total score for the child competence VAS and professional caregiver PEDS concerns and Parenting Stress Index Short Form (in Dutch abbreviated as NOSIK) for the parenting VAS. For all index tests, the parental and professional caregivers Strengths and Difficulties Questionnaire (SDQ) were used as a reference standard for perceived impact of distress.

To screen children's behavioural, emotional and social functioning, the CBCL, Dutch version 1.5–5, 2001 (Achenbach, Howell, Quay, & Conners, 1991), consists of 100 parent-reported problem items. Based on the behaviour of the child in the preceding 2 months, each item is scored on a 3-point Likert-type scale: 0 = not true, 1 = somewhat true and 2 = certainly true (Achenbach et al., 2008; Achenbach & Ruffle, 2000; Verhulst, van der Ende, & Koot, 1996). The total problem score is computed by adding the sum of all 0, 1 and 2 scores. Syndrome scales are part of the Internalizing

and Externalizing broadband scales. The Internalizing scale consists of four scales: emotionally reactive (e.g. 'rapid shifts between sadness and excitement', 'disturbed by any change in routine'); anxious/depressed (e.g. 'Looks unhappy without good reason', 'nervous, high strung or tense'); somatic complaints (e.g. 'headaches', 'nausea, feels sick') and withdrawn (e.g. 'refuses to play active games', 'seems unresponsive to affection'). The Externalizing scale contains two scales: Attention Problems (e.g. 'can't concentrate', 'wanders away') and Aggressive Behaviour (e.g. 'angry moods', 'defiant'). Good reliability and validity have been reported for the CBCL/1.5–5, also in the Netherlands (Ivanova et al., 2010).

The C-TRF Dutch version 1.5–5 (Verhulst, van der Ende, & Koot, 1997) is used for professional caregivers-reported children's behavioural, emotional and social functioning. It is the professional caregiver version of the CBCL and consists of almost the same 100 items.

CBCL and C-TRF cut-off points differ between populations: between countries as well as within countries and age groups (Tick, van der Ende, Koot, & Verhulst, 2007). If the clinical and borderline total score cut-off points are set to high, a substantial part of the MOM study children would not be identified as 'at risk' children, suggesting an underestimation. Therefore, in this study, CBCL en C-TRF continuous total problem scores were used and total score cut-off points were set at 10th, 25th and 50th percentile of the MOM study reference standard scores. To replace the clinical cut-off points, the percentile numbers are based on the current sample and were chosen arbitrary to cover a range.

The NOSIK is a 25-item self-report instrument to assess parental perceived difficulty in child rearing. Parental-stress-related statements are provided and parents can answer on a 5-point Likert-type scale ranging from 1 (totally disagree) to 5 (totally agree). The short form includes items with the best performance in the complete version of the Parenting Stress Index. The Cronbach's alpha of the NOSIK total score is between 0.92 and 0.95 (De Brock, Verhulst, Gerris, & Abidin, 1992). The NOSIK total score was obtained by adding the sum of all 1–5 scores, with a scoring range of 25–150. In this article, NOSIK continuous total problem score was used and total problem score cut-off points were set at 7th, 12th and 31st percentile. Scores were calculated using non clinical cut-off scores for the NOSIK: 74 (high–very high), 62 (above average) and 43 (average and below) (De Brock et al., 1992).

The Dutch version of the extended SDQ was included to assess the child's behaviour and the impact of distress (Goodman, 2001; van Widenfelt, Goedhart, Treffers, & Goodman, 2003). It is a brief behavioural screening questionnaire for children aged 3–16 years with 25 items on strengths and difficulties. These 25 items are divided in 5 scales: (1) emotional symptoms, (2) conduct problems, (3) hyperactivity/inattention, (4) peer relationship problems and (5) prosocial behaviour. The SDQ total difficulties sumscore is generated using the items of all subscales except prosocial behaviour. The extended SDQ includes an SDQ impact of distress supplement with eight questions to identify the impact of the behavioural problems of the child. The first question asks whether the informant thinks the child has a problem, the remaining questions assess chronicity, distress, social impairment and burden for others. From the SDQ, three dimensions can be inferred: perceived difficulties (is there a problem), impact score (distress and social incapacity on the child) and burden rating (do symptoms impose a burden). The SDQ is considered acceptable as a research instrument in community samples (Stone, Otten, Engels, Vermulst, & Janssens, 2010). Multi-informant SDQs (parents, professional caregivers) at the age of 5–17 years have a specificity of 80% and a sensitivity of 85% for the detection of child mental disorders (Goodman, Ford, Corbin, & Meltzer, 2004). A more recent study in including parent SDQ only showed a lower but still acceptable validity (sensitivity of 0.76 at a cut-off point with specificity of 0.90) in a Dutch community population of children at the age of 3–4 years (Theunissen et al., 2015). In this study, the parental and the professional caregiver SDQ continuous total problem score, SDQ total score, SDQ conduct score and SDQ impact score are used. Parental SDQ total score cut-off points were set at 12 (10th percentile)

and 9 (21st percentile), professional caregiver SDQ total score at 15 (11th percentile) and 11 (23rd percentile). These norm scores were determined from a sample of Dutch native children, representative for the Dutch population (Theunissen, 2016). The 'SDQ-impact score' refers to the impairment supplement and the 'probe question' refers to the first dimension: perceived difficulties (is there a problem). If any of the parents or professional caregivers scored 'yes' on the impact probe question in this study the dichotomous overall distress variable was set at 'yes'.

Statistical analysis

For this article, data of parents, professional caregivers and PCHC professionals were analysed at the age of 3 years and at the next follow-up almost a year later. Data analysis was performed using Stata 13 (Statistical Software Package) (StataCorp, 2013). Background characteristics and mean scores of used instruments were calculated. In order to test representativeness, parental education of a random sample of 40% non-responders was manually collected from the medical PCHC files.

Validation of the Dutch PEDS and different VAS was performed based on criteria of the COnsensus-based Standards for the selection of health Measurement INstruments (COSMIN) criteria for evidence (Mokkink et al., 2010).

Test–retest reliability of the three VAS and parental PEDS was examined in a separate data set. In total, 17 parents and 20 professional caregivers completed the measurement twice within a period that no real change had occurred between sessions (6–7 days). The intraclass correlation coefficient (ICC) and 95% confidence interval (CI) were calculated using the continuous variables (oneway random effect model, individual measurements) (StataCorp, 2015). In addition, Kappa was calculated using various cut-off points (Bouter & Van Dongen, 2005).

Parental concerns and professional caregivers' concerns were cross tabulated and a chi-square statistic was calculated to assess agreement and disagreement between informants.

Internal consistency of the PEDS items was tested as a measure of reliability by computing Cronbach's alpha coefficients. In addition, factor analysis assessed whether items load on the same factor.

Validity of the different VAS and PEDS, with respect to children's mental health, behaviour and parental perceived difficulty in child rearing, were examined. For construct validity (whether instrument has associations with other constructs as expected), Spearman correlations between these instruments and validated instruments were calculated. In addition, associations between the different VAS and PEDS on the one hand, and the SDQ, CBCL, C-TRF and the NOSIK on the other hand, were analysed using linear (continuous outcomes) and logistic (dichotomous outcomes) regression analysis, in separate regression models. Results of linear regression analysis were standardized, that is, all continuous variables were converted to variables with standard deviation of one. Normal distribution of the residuals and heteroscedasticity were checked after all linear regression analyses. Assumptions were violated, however, violations were minor. Because regression analysis is relatively robust, results are assumed valid.

To compare the index instruments with well-established instruments assessing the same construct, criterion validity (sensitivity and specificity) and convergent validity (correlation between the two continuous variables) were assessed. Each of the VAS had a different reference standard: the NOSIK for the parenting VAS, the CBCL total score for the child behaviour VAS and the C-TRF total score for the child competence VAS. Concerning general emerging problems and disabling symptoms, reference standards of the parental and professional caregiver PEDS were the CBCL and the C-TRF respectively. Sensitivity, specificity and NPV of all index instruments were assessed and ROC curves were generated. Subsequently, the area under the ROC curve (AUROC) was calculated for the different VAS. To examine convergent and construct validity, each variable

was entered in a separate regression model to measure the association between each index/test measures and reference standards. It was hypothesized that there would be a stronger association between the reference standard and the index test and also an association between the index tests and other validated instruments, but not as strong.

Finally, external responsiveness of the different VAS was tested by comparing changes in VAS score between age 3 and 4 years with changes in their reference standard score, to detect change over time in the construct to be measured. Changes in the PEDS were related to changes in the SDQ, NOSIK, CBCL and C-TRF total score.

Results

Descriptives of the study sample

At baseline, questionnaires of 346 children were completed, 166 (48%) boys and 180 (52%) girls. In over 85% of the children, both parent and professional caregiver filled in the questionnaire. The mean age of the children at baseline was 3.0 ($SD \pm 0.2$, Table 1). More than half of the parents had received higher education (63%). At baseline, parents scored on average 70.3 on the parenting VAS and 66.3 on the child behaviour VAS. Professional caregivers scored on average 63.7 on the child competence VAS. Parents of 147 children (43%) and professional caregivers of 162 children (54%) indicated any PEDS concern. Based on the PCHC 'PEDS interpretation form', 33 of 325 children (10%) were scored at 'high risk' at the age of 3 years, another 100 children (31%), were scored at 'moderate risk'. The follow-up was approximately 10 months later (mean age 3.8; $SD \pm 0.2$, Table 1).

Reliability and validity

The test–retest reliability of the different VAS showed a strong and significant correlation between two consecutive assessments (ICC: Parenting VAS = 0.8, 95% CI = [0.6, 0.9]; Child behaviour VAS = 0.9, 95% CI = [0.8, 1.0]; Child competence VAS = 0.9, 95% CI = [0.8, 1.0]). The test–retest reliability of the parental PEDS was strong as well (ICC: 0.8, 95% CI = [0.5, 0.9]). Various cut-off points of the different VAS and PEDS showed different Kappa's (Table 2).

There was a statistically significant association between parental concerns and professional caregivers' PEDS concerns both at baseline and at follow-up (Chi-square = 34.8; df = 1, $p < .001$ and 8.1; df = 1, $p = .004$, respectively).

Internal consistency of the PEDS was tested and factor analysis revealed that all parental PEDS items loaded on one factor. Professional caregiver PEDS items loaded on two factors, but the second factor provided no added value. In both parental PEDS and professional caregivers' PEDS, loadings of individual items were not strong (loadings between 0.33 and 0.70 and between 0.09 and 0.64, respectively). Cronbach's alpha of the PEDS was 0.7 in parents and 0.6 in professional caregivers.

For construct validity, the three VAS were correlated with the SDQ total score, the SDQ conduct subscale, the CBCL, the C-TRF and the NOSIK. Higher VAS scores means less-perceived difficulties on parenting, child behaviour or child competence, therefore, the higher the VAS, the lower the scores on these validated instruments. Table 3 presents correlations including those between the different VAS and their reference standard at the age of 3 years and replicated with data collected approximately a year later (Spearman correlation: Parenting VAS–NOSIK = −0.57 and −0.48, $p < .001$; Child behaviour VAS–CBCL = −0.56 and −0.57, $p < .001$; Child competence VAS–C-TRF = −0.40 and −0.32, $p < .001$; Table 3).

Linear regression analyses showed associations between the VAS, PEDS and various validated outcomes (Table 4). For example, parents who indicated more parenting difficulties also showed

Table 1. Descriptive statistics at base line (T1) and follow-up (T2).

Variable	*N*		Mean (*SD*)		Range	
	T1	T2	T1	T2	T1	T2
Age in years	346	293	3.0 (0.2)	3.8 (0.2)	1.8–3.5	3.5–4.8
Parenting VAS[a]	327	281	70.3 (18.8)	72.5 (16.5)	6–100	14–99
Child's behaviour VAS[b]	329	279	66.3 (17.8)	67.8 (16.5)	1–97	11–100
Child competence VAS[c]	290	251	63.7 (19.7)	69.7 (16.3)	4–100	10–99
SDQ (parents)	338	293	6.8 (4.9)	6.1 (4.2)	0–28	0–27
SDQ (prof.[d])	294	256	6.1 (5.0)	5.0 (5.0)	0–27	0–29
NOSIK	336	289	41.4 (18.7)	38.9 (17.0)	25–140	25–134
CBCL	331	285	21.3 (19.2)	17.8 (15.2)	0–117	0–109
C-TRF	289	257	13.2 (17.1)	13.2 (15.7)	0–92	0–98
			No concerns		Concerns	
PEDS (parents)	339	293	192 (57%)	167 (57%)	147 (43%)	126 (43%)
PEDS (prof.[d])	300	257	138 (46%)	144 (56%)	162 (54%)	113 (44%)
	T1	T2	T1	T2	T1	T2	T1	T2
			Low risk		Moderate risk		High risk	
PEDS (PCHC)	325	312	192 (59%)	206 (66%)	100 (31%)	84 (27%)	33 (10%)	22 (7%)
			Normal		Borderline		Abnormal	
SDQ impact (parents)	340	292	307 (90%)	271 (93%)	6 (2%)	8 (3%)	27 (8%)	13 (4%)
SDQ impact (prof.[d])	292	254	248 (85%)	223 (88%)	0	0	44 (15%)	31 (12%)

VAS: visual analogue scale; SDQ: Strengths and Difficulties Questionnaire; NOSIK: Parenting Stress Index Short Form; CBCL: Child Behaviour Checklist; C-TRF: Caregiver-Teacher's Report Form; PEDS: parents' evaluation of developmental status; PCHC: preventive child health care.
[a]A higher VAS score means parent judges parenting more positive.
[b]A higher VAS score means parent judges child behaviour more positive.
[c]A higher VAS score means professional caregiver judges child competence more positive.
[d]Professional caregivers.

Table 2. Test–retest reliability: Kappa at different cut-off points different VAS and parental PEDS.

	Parenting VAS			Child behaviour VAS			Child competence VAS			Parental PEDS concerns		
Cut-off point[a]	42[b]	64	74	44	55	69	39	50	67	3	2	1
Kappa		0.42	0.75	0.43	0.82	0.85	0.47	1.00	0.52	0.64	0.82	0.64

VAS: visual analogue scale; PEDS: parents' evaluation of developmental status.
Kappa: 0−0.20 = 'slight'; 0.21−0.40 = 'fair'; 0.41−0.60 = 'moderate'; 0.61−0.80 = 'substantial'; above 0.81 = 'almost perfect'.
[a]When dichotomizing the different VAS cut-off point and amount of PEDS concerns, all values below are coded as high-perceived difficulty.
[b]No Kappa available.

significantly higher problem scores on the NOSIK ($\beta = -0.7$, $p < .001$) and the SDQ parents total score (parents $\beta = -0.5$, $p < .001$). The association between child behaviour VAS and professional caregivers SDQ showed a small effect size according to Cohen (Cohen, 1988; Nakagawa & Cuthill,

Table 3. Spearman rho correlations between three VAS scales and validated instruments at base line (T1) and follow-up (T2).

		Parents						Professional caregivers			
		Parenting VAS[a]	Child's behaviour VAS[b]	NOSIK Total score	SDQ total score	SDQ conduct	CBCL total score	Child competence VAS[c]	SDQ total score	SDQ conduct	C-TRF total score
Parents											
Parenting VAS[a]		1									
Child's behaviour VAS[b]	T1	0.64†	1								
	T2	0.63†									
NOSIK total score	T1	−0.57†	−0.43†	1							
	T2	−0.48†	−0.46†								
SDQ total score	T1	−0.47†	−0.46†	0.49†	1						
	T2	−0.45†	−0.49†	0.45†							
SDQ conduct	T1	−0.49†	−0.54†	0.43†	0.59†	1					
	T2	−0.44†	−0.56†	0.37†	0.58†						
CBCL total score	T1	−0.57†	−0.56†	0.68†	0.69†	0.54†	1				
	T2	−0.51†	−0.57†	0.62†	0.71†	0.46†					
Professional caregivers											
Child competence VAS[c]	T1	0.11	0.17**	−0.12	−0.29†	−0.13*	−0.21†	1			
	T2	0.15	0.11	−0.01	−0.21**	−0.09	−0.16*				
SDQ total score	T1	−0.14*	−0.20**	0.23†	0.35†	0.19**	0.32†	−0.47†	1		
	T2	−0.23†	−0.19**	0.16*	0.31†	0.25†	0.31†	−0.32†			
SDQ conduct	T1	−0.18**	−0.20**	0.20**	0.18**	0.25†	0.24†	−0.17**	0.60†	1	
	T2	−0.22†	−0.25†	0.16*	0.19**	0.28†	0.22**	−0.07	0.60†		
C-TRF total score	T1	−0.24†	−0.33†	0.25†	0.29†	0.21†	0.31†	−0.40†	0.65†	0.46†	1
	T2	−0.26†	−0.21**	0.22**	0.27†	0.29†	0.32†	−0.32†	0.73†	0.51†	

VAS: visual analogue scale; NOSIK: Parenting Stress Index Short Form; SDQ: Strengths and Difficulties Questionnaire; CBCL: Child Behaviour Checklist; C-TRF: Caregiver-Teacher's Report Form.

[a]A higher VAS score means parent judges parenting more positive.

[b]A higher VAS score means parent judges child behaviour more positive.

[c]A higher VAS score means professional caregiver judges child competence more positive.

*$p < .05$; **$p < .01$; †$p < .001$.

Table 4. Linear regression analysis and logistic regression analyses to analyse the association between 3 VAS scales and PEDS (in separate regression models) and various validated outcomes (such as NOSIK, SDQ, CBCL, C-TRF) at base line (T1) and follow-up (T2).

Total score		Parents			Professional caregivers		SDQ impact (parents) OR (95% CI)	SDQ impact (prof.[d]) OR (95% CI)
		SDQ β (95% CI)	CBCL β (95% CI)	NOSIK β (95% CI)	SDQ β (95% CI)	C-TRF β (95% CI)		
Parenting VAS (parents)[a]	T1	−0.5 [−0.6, −0.4]†	−0.6 [−0.7, −0.5]†	−0.7 [−0.7, −0.6]†	−0.2 [−0.3; −0.8]**	−0.3 [−0.4, −0.2]†	0.3 [0.2, 0.4]†	0.6 [0.5, 0.8]**
	T2	−0.5 [−0.6, −0.4]†	−0.6 [−0.7, −0.5]†	−0.5 [−0.6, −0.4]†	−0.4 [−0.5, −0.3]†	−0.4 [−0.5, −0.2]†	0.3 [0.2, 0.5]†	0.5 [0.4, 0.8]**
Child behaviour VAS (parents)[b]	T1	−0.5 [−0.6, −0.4]†	−0.5 [−0.6, −0.5]†	−0.5 [−0.6, −0.4]†	−0.2 [−0.3, −0.1]†	−0.4 [−0.5, −0.3]†	0.4 [0.2, 0.5]†	0.6 [0.4, 0.8]**
	T2	−0.5 [−0.6, −0.4]†	−0.5 [−0.6, −0.4]†	−0.5 [−0.6, −0.4]†	−0.2 [−0.4, −0.1]†	−0.2 [−0.4, −0.1]†	0.4 [0.3, 0.6]†	0.7 [0.5, 1.0]
Child competence VAS (prof.[d])[c]	T1	−0.3 [−0.4, −0.2]†	−0.3 [−0.4, −0.2]†	−0.3 [−0.4, −0.1]†	−0.6 [−0.7, −0.5]†	−0.6 [−0.7, −0.5]†	0.4 [0.3, 0.6]†	0.2 [0.1, 0.3]†
	T2	−0.3 [−0.4, −0.2]†	−0.3 [−0.4, −0.2]†	−0.2 [−0.3, −0.1]**	−0.5 [−0.6, −0.4]†	−0.4 [−0.5, −0.3]†	0.4 [0.3, 0.7]†	0.2 [0.1, 0.3]†
PEDS concerns yes/no (parents)	T1	0.5 [0.4, 0.7]†	0.6 [0.5, 0.8]†	0.5 [0.4, 0.7]†	0.3 [0.1, 0.4]†	0.3 [0.2, 0.5]†	24.8 [5.8, 106]†	4.7 [2.2, 9.8]†
	T2	0.6 [0.4, 0.9]†	0.8 [0.6, 1.0]†	0.7 [0.4, 0.9]†	0.5 [0.3, 0.8] †	0.5 [0.3, 0.8]†	14.8 [3.4, 64.8]†	3.9 [1.6, 9.8]**
PEDS concerns yes/no (prof.[d])	T1	0.5 [0.3, 0.7]†	0.5 [0.3, 0.7]†	0.4 [0.2, 0.7]†	0.9 [0.7, 1.0]†	0.6 [0.5, 0.8]†	14.5 [3.4, 62.3]†	52.2 [7.1, 385]†
	T2	0.4 [0.2, 0.7]**	0.4 [0.2, 0.7]**	0.1 [−0.1, 0.4]	1.0 [0.7, 1.1]†	0.8 [0.6, 1.0]†	3.2 [1.2, 8.7] *	25.4 [5.9, 109]†
PEDS interpretation high-moderate /low risk (PCHC)	T1	0.7 [0.4, 0.9]†	0.7 [0.5, 0.9]†	0.6 [0.3, 0.8]†	0.7 [0.5, 0.9]†	0.6 [0.4, 0.8]†	5.2 [2.3, 12.0]†	5.1 [2.5, 10.7] †
	T2	0.5 [0.2, 0.7]†	0.4 [0.2, 0.7]**	0.3 [0.1, 0.6]**	0.5 [0.3, 0.8]†	0.5 [0.3, 0.8]†	3.4 [1.3, 8.4]*	2.5 [1.2, 5.4] *

NOSIK: Parenting Stress Index Short Form; SDQ: Strengths and Difficulties Questionnaire; CBCL: Child Behaviour Checklist; C-TRF: Caregiver-Teacher's Report Form; VAS: visual analogue scale; PEDS: parents' evaluation of developmental status; CI: confidence interval; OR: odds ratio (obtained from logistic regression).
Cells in grey present analysis with an index test and the reference standard of that specific index test; β = standardized regression coefficient.
[a]A higher VAS score means parent judges parenting more positive.
[b]A higher VAS score means parent judges child behaviour more positive.
[c]A higher VAS score means professional caregiver judges child competence more positive.
[d]Professional caregivers.
*$p < .05$; **$p < .01$; †$p < .001$.

Table 5. Parenting VAS with different AUROC and a variation of sensitivity, specificity and negative predictive values (NPVs) as a result of different VAS and reference standard cut-off points, at base line (T1) and follow-up (T2).

Various cut-off points of the NOSIK[a] (Perceived difficulties)	AUROC (%) (95% CI)		Various cut-off points of the parenting VAS[b]		Sensitivity/ specificity (%)		NPV (%)	
	T1	T2	T1	T2	T1	T2	T1	T2
74 (high–very high)	94.0 [90.1, 97.8]	87.8 [76.1, 99.5]	<42	<49	77.3/95.0	62.5/92.4	98.3	97.6
			<64	<66	90.9/78.4	87.5/75.2	99.2	99.0
			<74	<75	100/52.5	93.8/51.9	100	99.3
62 (above average)	87.8 [81.5, 94.2]	85.9 [76.2, 95.5]	<42	<49	55.0/96.5	52.4/92.6	93.8	96.0
			<64	<66	77.5/80.9	81.0/75.9	96.2	98.0
			<74	<75	90.0/54.1	90.5/52.5	97.5	98.5
43 (average and below)	79.8 [74.6, 85.0]	80.7 [75.9, 86.6]	<42	<49	26.5/97.7	31.3/95.7	74.2	81.5
			<64	<66	54.9/86.9	62.7/82.5	80.7	87.4
			<74	<75	81.4/62.4	82.1/59.2	87.9	91.2

VAS: visual analogue scale; AUROC: area under the ROC curve; NPV: negative predictive value; CI: confidence interval. AUROC: > 90% = excellent; 80–90% = good; 70–80% = fair; 60–70% = poor.

[a]When dichotomizing the reference standard cut-off point all values above are coded as difficult parenting from the perspective of the parent.

[b]A higher VAS score means parent judges parenting more positive; the different VAS cut-off points were set at 10%, 25% and 50%.

2007), but the association was statistically significant ($p < .001$). The association between the competence VAS and outcomes assessed by the parents was also small, but significant. All other associations between continuous variables (first 3 rows of Table 4) showed intermediate or large effect sizes. In addition, logistic regression showed strong associations between all three VAS and the dichotomous SDQ impact variables. The linear and logistic regression analyses showed strong associations as well between parental PEDS concerns, professional caregivers PEDS concerns and PCHC PEDS interpretation and various validated instruments (Table 4). When analyses were repeated in the follow-up data, results were similar. Some associations were weaker but remained significant, with one exception. There was no significant association between professional caregiver PEDS concerns and the NOSIK.

In relation to criterion validity, diagnostic accuracy of the Parenting VAS, including AUROC, sensitivity, specificity and NPV is shown in Table 5 and Figure 1. At baseline, when dichotomizing the reference standard cut-off point and all values above are codes as high-perceived difficulty, the different VAS are performing excellent to poor. Different AUROC's are as follows: Parenting VAS – NOSIK 94.0, 87.8, 79.8; Child behaviour VAS – CBCL 84.4, 80.8, 76.9; Child competence VAS – C-TRF 89.5, 72.8, 66.2 (Tables 5 to 7). At the follow-up moment almost a year later, results were similar. Tables 8 and 9 and Figures 2 and 3 show the diagnostic accuracy of the parental and professional caregiver PEDS, depending on various cut-off points of the CBCL and C-TRF as reference standard and the PEDS at baseline. At the follow-up moment almost a year later, results were similar. (Other figures are available on request).

Changes in VAS were associated with changes in reference standard, but correlations were relatively low (Table 10, external responsiveness). Potential outliers were checked manually and twice a parent reversed the VAS scoring which was corrected. Changes in parental and professional caregiver PEDS concerns were associated with changes in SDQ impact scores.

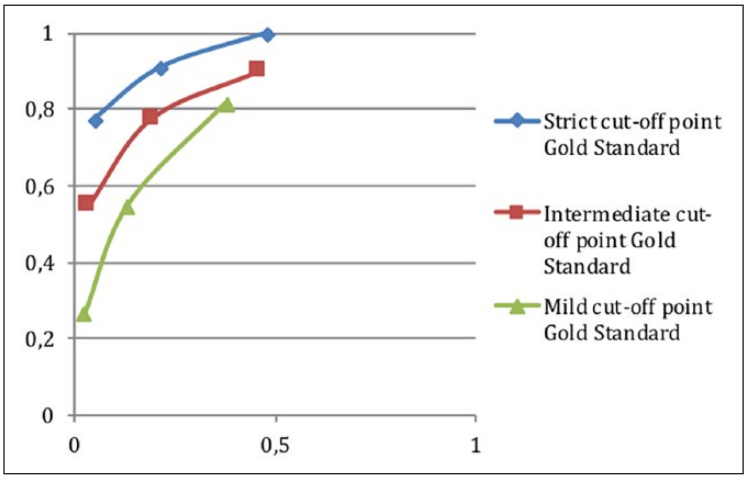

Figure 1. Different cut-off points parenting VAS and NOSIK: strict-intermediate and mild, at baseline (T1).
NOSIK: Parenting Stress Index Short Form; VAS: visual analogue scale.

Table 6. Child behaviour VAS with different AUROC and a variation of sensitivity, specificity and negative predictive values (NPVs) as a result of different VAS and reference standard cut-off points, at base line (T1) and follow-up (T2).

Various cut-off points of the CBCL[a,b]		AUROC (%) (95% CI)		Various cut-off points of the child behaviour VAS[c]		Sensitivity/ specificity (%)		NPV (%)	
T1	T2	T1	T2	T1	T2	T1	T2	T1	T2
43 + >	36 + >	84.4 [76.6; 92.3]	82.0 [73.2, 90.7]	<44	<47	60.0/94.8	44.4/93.1	95.1	93.8
				<55	<60	71.4/80.8	70.4/78.8	95.9	96.0
				<69	<69	88.6/51.8	88.9/51.8	97.4	97.7
28 + >	23 + >	80.8 [75.4, 86.3]	79.8 [74.1, 85.5]	<44	<47	35.4/97.1	27.5/95.1	81.4	79.4
				<55	<60	57.3/86.2	53.6/83.3	85.5	84.1
				<69	<69	82.9/57.7	81.2/57.6	90.8	90.0
16 + >	14 + >	76.9 [71.8, 82.0]	75.5 [69.7, 81.2]	<44	<47	20.5/98.1	19.6/99.2	55.1	52.7
				<55	<60	41.6/91.9	39.2/88.4	60.1	56.7
				<69	<69	71.4/66.3	70.6/68.2	69.7	67.7

VAS: visual analogue scale; AUROC: area under the ROC curve; NPV: negative predictive value; CBCL: Child Behaviour Checklist; CI: confidence interval.
AUROC: > 90% = excellent; 80–90% = good; 70–80% = fair; 60–70% = poor.
[a]When dichotomizing the reference standard cut-off point all values above are coded as bad child behaviour from the perspective of the parent.
[b]The different CBCL cut-off points were set at 10%, 25% and 50%; the higher the score, the more perceived problems.
[c]A higher VAS score means parent judges child behaviour more positive; the different VAS cut-off points were set at 10%, 25% and 50%.

Discussion

Thus, the Dutch PEDS, a 'child behaviour' VAS, a 'parenting' VAS and a 'child competence' VAS, at the age of 3 and 4 years demonstrate good psychometric properties.

Table 7. Child competence VAS with different AUROC and a variation of sensitivity, specificity and negative predictive values (NPVs) as a result of different VAS and reference standard cut-off points, at base line (T1) and follow-up (T2).

Various cut-off points of the C-TRF[a,b]		AUROC (%) (95% CI)		Various cut-off points of the child competence VAS[c]		Sensitivity/specificity (%)		NPV (%)	
T1	T2	T1	T2	T1	T2	T1	T2	T1	T2
31 + >	28 + >	89.5 [82.0, 97.0]	76.4 [65.7, 87.1]	<39	<48	58.6/96.4	44.8/92.8	95.3	92.8
				<50	<61	82.8/82.8	62.1/77.9	97.6	94.0
				<67	<71	93.1/54.4	79.3/53.6	98.6	95.2
16 + >	14 + >	72.8 [65.3, 80.3]	65.5 [57.8, 73.2]	<39	<48	28.8/97.6	27.0/94.9	79.5	75.5
				<50	<61	49.3/85.0	41.9/79.7	82.6	76.6
				<67	<71	74.0/57.8	60.8/54.2	86.2	76.8
7 + >	6 + >	66.2 [59.8, 72.5]	63.4 [56.3, 70.4]	<39	<48	17.5/100	15.7/95.7	51.4	39.6
				<50	<61	34.9/88.5	32.7/83.7	54.3	41.8
				<67	<71	60.4/60.8	56.0/59.8	57.3	44.0

VAS: visual analogue scale; AUROC: area under the ROC curve; NPV: negative predictive value; C-TRF: Caregiver-Teacher's Report Form; CI: confidence interval.
AUROC: > 90% = excellent; 80–90% = good; 70–80% = fair; 60–70% = poor.
[a]When dichotomizing the reference standard cut-off point all values above are coded as low competence from the perspective of the professional caregiver.
[b]The different C-TRF cut-off points were set at 10%, 25% and 50%; the higher the score, the more perceived problems.
[c]A higher VAS score means professional caregiver judges child competence more positive; the different VAS cut-off points were set at 10%, 25% and 50%.

Table 8. Parental PEDS concerns with a variation of sensitivity, specificity and negative predictive values (NPVs) as a result of different amount of PEDS concerns and reference standard cut-off points, at base line (T1) and follow-up (T2).

Various cut-off points of the CBCL[a,b]		Various cut-off points amount of parental PEDS concerns		Sensitivity/specificity (%)		NPV (%)	
T1	T2	T1	T2	T1	T2	T1	T2
43 + >	36 + >	3 + >	3 + >	94.4/63.5	89.7/62.9	98.9	98.2
		2 + >	2 + >	80.6/80.2	72.4/82.8	97.1	96.4
		1 + >	1 + >	58.3/90.4	58.6/93.8	94.6	95.2
28 + >	23 + >	3 + >	3 + >	79.8/69.8	75.0/68.5	91.0	89.0
		2 + >	2 + >	59.5/84.9	50.0/86.4	86.0	83.6
		1 + >	1 + >	41.7/94.3	31.9/95.3	82.5	80.5
16 + >	14 + >	3 + >	3 + >	64.5/79.1	56.6/73.7	68.6	59.8
		2 + >	2 + >	43.4/90.8	34.9/91.0	61.2	55.0
		1 + >	1 + >	26.0/96.3	18.4/96.2	56.1	50.8

PEDS: parents' evaluation of developmental status; NPV: negative predictive value; CBCL: Child Behaviour Checklist.
[a]When dichotomizing the reference standard cut-off point all values above are coded as high-perceived difficulties.
[b]The different CBCL cut-off points were set at 10%, 25% and 50%; the higher the score, the more perceived problems.

Table 9. Professional caregiver PEDS concerns with a variation of sensitivity, specificity and negative predictive values (NPVs) as a result of different amount of PEDS concerns and reference standard cut-off points, at base line (T1) and follow-up (T2).

Various cut-off points of the C-TRF[a,b]		Various cut-off points amount of professional caregiver PEDS concerns		Sensitivity/specificity (%)		NPV (%)	
T1	T2	T1	T2	T1	T2	T1	T2
31 + >	28 + >	3 + >	3 + >	100/51.9	93.5/62.8	100	98.6
		2 + >	2 + >	100/72.9	87.1/80.1	100	97.8
		1 + >	1 + >	96.6/83.1	71.0/89.4	99.5	95.7
16 + >	14 + >	3 + >	3 + >	89.5/59.6	73.7/68.5	94.1	86.1
		2 + >	2 + >	77.6/79.8	61.8/86.2	90.9	84.3
		1 + >	1 + >	64.5/89.2	44.7/93.4	87.5	80.1
7 + >	6 + >	3 + >	3 + >	78.1/75.4	55.2/75.5	74.8	49.3
		2 + >	2 + >	56.8/89.6	38.7/90.4	64.2	45.9
		1 + >	1 + >	43.2/96.3	26.4/96.8	59.4	43.1

PEDS: parents' evaluation of developmental status; NPV: negative predictive value; C-TRF: Caregiver-Teacher's Report Form.
[a]When dichotomizing the reference standard cut-off point all values above are coded as high-perceived difficulties.
[b]The different C-TRF cut-off points were set at 10%, 25% and 50%; the higher the score, the more perceived problems.

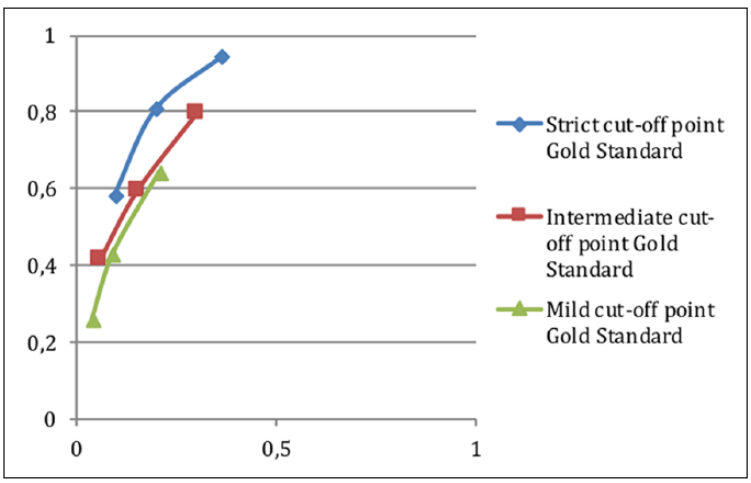

Figure 2. Different cut-off points PEDS parent and CBCL: strict-intermediate and mild, at baseline (T1).
CBCL: Child Behaviour Checklist; PEDS: parents' evaluation of developmental status.

Test–retest reliability

This study showed a good test–retest reliability of the different VAS and PEDS. Internal consistency on the PEDS was weak. This was expected because of the characteristics of the PEDS with the underlying factor structure and results in previous studies (Woolfenden et al., 2014); answers to each PEDS questions contribute uniquely to the overall result.

Both versions of the PEDS concerns were strongly associated but overlap was far from 100%, indicating important contextual effects. This supports the usefulness of the multi-informant procedure as applied in the MOM study.

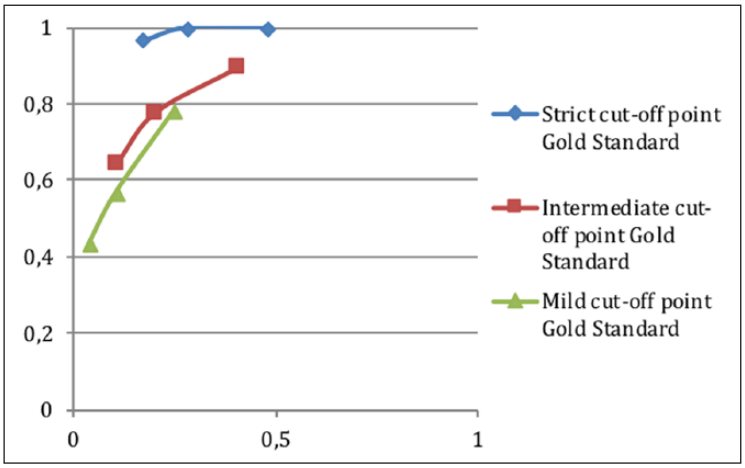

Figure 3. Different cut-off points PEDS professional caregiver and C-TRF: strict-intermediate and mild, at baseline (T1).
C-TRF: Caregiver-Teacher's Report Form; PEDS: parents' evaluation of developmental status.

Table 10. Spearman correlations between change scores in VAS scales and PEDS and changes in their reference standard (external responsiveness).

Changes between baseline and follow-up in VAS and PEDS:	Changes in reference standard	Spearman correlation
Parenting VAS (parents)[b]	NOSIK	−0.35†
Child behaviour VAS (parents)[c]	SDQ (parents)	−0.13*
Child behaviour VAS (parents)[c]	CBCL	−0.38†
Child competence VAS (prof.[a])[d]	SDQ (prof.[a])	−0.36†
Child competence VAS (prof.[a])[d]	C-TRF	−0.26†
PEDS concerns yes/no (parents)	Impact SDQ (parents)	0.26†
PEDS concerns yes/no (prof.[a])	Impact SDQ (prof.[a])	0.27†

NOSIK: Parenting Stress Index Short Form; SDQ: Strengths and Difficulties Questionnaire; CBCL: Child Behaviour Checklist; C-TRF: Caregiver-Teacher's Report Form; VAS: visual analogue scale; PEDS: parents' evaluation of developmental status.
[a]Professional caregivers.
[b]A higher VAS score means parent judges parenting more positive.
[c]A higher VAS score means parent judges child behaviour more positive.
[d]A higher VAS score means professional care giver judges child competence more positive.
*$p < .05$; **$p < .01$; †$p < .001$.

Construct validity

All three VAS as well as the PEDS were also moderately but significantly correlated with the NOSIK as well as with the SDQ, CBCL and C-TRF in the expected direction. Standardized effect sizes showed a relatively strong association between the parenting VAS and the NOSIK, the child behaviour VAS and the CBCL and the child competence VAS and the C-TRF.

Criterion and convergent validity

Table 4 shows the associations between index test and different validated instruments. The parenting VAS and the child behaviour VAS were moderately but significantly correlated with their

reference standards and other parental validated instruments. There was a significant but less strong correlation between the parental VAS and the professional caregiver outcomes. The same was true for the child competence VAS, a stronger correlation with the professional caregiver outcomes SDQ and C-TRF but less stronger correlations with the parental outcomes. As a global indicator of the diagnostic performance of the index test, the diverse AUROC of the different VAS were good, especially with strict and intermediate reference standard cut-off points. In addition, at various cut-off points of the reference standard, both parenting VAS and behaviour VAS show a high sensitivity and NPV at cut-off points of <74 and <69, respectively, with similar results almost a year later (Tables 5 and 6). Tables 5 to 9 are an example of an overview to support PCHC professionals to make a choice of an appropriate combination of sensitivity, specificity, NPV in relation to cut-off points of the index test. For the PCHC professional, for example, a specific parental VAS cut-off point in daily practice can be reason to ask more about the parental representation of parenting. Figure 1 shows that there is more between a 'strict clinical' cut-off point and 'no concerns'. The child competence VAS showed moderate but significant correlation with the C-TRF (Table 3). Only with two VAS cut-off points (<50 and <67) related with a strict cut-off point of the reference standard, the competence VAS showed a high sensitivity/specificity and NPV (Table 7). Thus, these results suggest that for early identification of emerging concerns and disabling symptoms, both parental and child behaviour VAS are sufficiently valid and the child competence VAS is less valid.

Furthermore, both parental and professional caregivers' PEDS concerns as well as PCHC PEDS interpretation were strongly associated with their reference standard. Depending on various cut-off points of the reference standard, three or more PEDS concerns showed a high sensitivity and NPV for high-perceived mental health problems. Figures 2 and 3 show the diversity of sensitivity between a 'strict clinical' cut-off point and 'no concerns'.

It is noteworthy that associations and regression coefficients testing convergent validity (association with reference standard) are not stronger than associations and regression coefficients testing construct validity (association with a similar but non-identical construct). No obvious explanation is available, yet. The difference between parental and professional caregivers' outcomes as seen in Table 4 emphasizes the contextual influences.

PEDS in previous research

In this study, 10% and 7% of the children were indicated as high developmental risk at 3 and 4 years of age respectively, and 31% and 27% were associated with moderate risk. A review study identified a prevalence of parental concerns on the PEDS indicating 13.8% (95% CI=[10.9%, 16.8%]) high-risk children and 19.8% (95% CI=[16.7%, 22.9%]) moderate-risk children. The high amount of children at moderate risk in this study is possibly related with the research population of children aged 3 and 4 years of age with relatively highly educated parents. Studies parental PEDS concerns conducted in high-income countries reported a significantly higher rate of moderate-risk concerns (Woolfenden et al., 2014).

Thus, the PEDS may also be valid for use in routine Dutch PCHC practice. The PEDS reveals questions and subjects that parents and professional caregivers want to discuss which are not yet related to changes in the SDQ, NOSIK, CBCL and C-TRF total score. The PEDS facilitates communication concerning parental and professional caregivers concerns about and across different child developmental domains. The answers always need to be checked in case parents do not use screening instruments properly, for example, because of literacy and language barriers or in case parents do not raise concerns when they should. In addition to these short instruments, other PCHC developmental instruments, for example, the Van Wiechen Developmental screen, Family Centred Method, Spark, Ages and Stages Questionnaire (ASQ) and SDQ can be used to assess and monitor

parental and child mental health, their interactions, family history, risk and resilience factors and developmental milestones (Hielkema et al., 2011; Jacobusse, van Buuren, & Verkerk, 2006; Staal, Hermanns, Schrijvers, & van Stel, 2013; Theunissen et al., 2015).

Methodological issues

This study has some limitations. A rather high proportion of parents did not give consent to participation. The response ratio was 20%. In all, parents of 1702 children were asked to participate in the MOM study. These children were within the caseload of the PCHC professionals participating in the MOM study. However, not all PCHC professionals participated in MOM. Consequently, parents of non-participating PCHC doctors were asked to participate by another PCHC doctor (B.D.), who did not know these families. During the baseline inclusion, the number of participating PCHC professionals increased. Response from one PCHC doctor who participated from the beginning (B.D.) was 70%. In order to test representativeness, 40% of non-responders were randomly sampled to manually collect data on parental education from the medical files. The distribution in non-responders was minimally different from distribution in responders (responders baseline 63%, 27%, 10%, follow-up 64%, 26%, 10% and non-responders 55%, 33%, 12% having high, intermediate and low parental education, respectively). The overrepresentation of parents with a high level of education suggests the presence of selection and the possibility of limited representativeness. Findings suggest that there is a tendency to meaningful differences in effect sizes between different levels of parental education (Tables available on request). Further research is required with a greater numbers of participants with lower parental educational status.

Because response in participating doctors was relatively high and because distribution in socio-economic status was comparable, results presented in this article can be considered approximately representative for the general population. If the PEDS and different VAS are implemented in general PCHC practice, a response rate higher than 70% is expected because a possible barrier for parents to participate in MOM was the number of questions added for research purposes (e.g. additional instruments for the purpose of validation of PEDS and VAS). Usually, short form questionnaires collected in PCHC have response rates between 80% and 90% (de Wolff, Theunissen, Vogels, & Reijneveld, 2013).

In addition, although in only two occasions a parent reversed the VAS scoring, instructions for use of the VAS has to be improved to prevent mistakes in scoring.

Another limitation is the reference standard which should ideally be an established measure accurately measuring the same construct as the index test should measure. In this article, the same kind of reference tests, CBCL and C-TRF, are used to establish the validity of two different index tests, the parenting VAS and the competence VAS. Notwithstanding, parental CBCL and professional caregivers' C-TRF are the true reference standards for parenting VAS and competence VAS, respectively. These reference tests were the best available and applicable validated instruments for routine preventive child health assessments to measure signs and symptoms of early distress and impairment in a general population of children.

Furthermore, it is recommended to repeat the test–retest with at least 50 participants and to validate other aspects of the Dutch PEDS. The PEDS includes separate domains such as language and motor skills and these constructs also have a reference standard. Although each type of parental concern can be associated with validated tests on the same developmental domain, studies about the validity of the PEDS showed that parents often have concerns in seemingly unrelated domains, that is, parents often reflect on not just the apparent problem but also its impact on other aspects of development (Glascoe, 2013). However, assessing the validity of the language and motor developmental domains of the Dutch PEDS is beyond the scope of this article. The focus of this article is

on obtaining early signals to facilitate a dialogue on overall concerns across several developmental domains like behavioural, social-emotional and mental health.

Conclusion

One of the most basic activities in PCHC is the focus on early identification of developmental and behavioural problems. In this respect, PCHC professionals have to deal with emerging problems and symptoms at a stage where signs and symptoms do not yet meet diagnostic criteria, but already give rise to early impairment and distress for both the children and their families. In order to screen emerging problems, short but valid monitoring tools are required. The PEDS as well as the parenting VAS and the child behaviour VAS have shown to be valid in Dutch PCHC standard practice. The child competence VAS is less valid. Further research is needed. The experiences in the context of this study stress the challenge to use a multi-informant approach with a PCHC 'toolkit' of short validated instruments for multi-axial information. It can improve communication and 'shared decision making' for personalized PCHC (Huber et al., 2016; Rutter, 2013).

Authors' note

Underlying research materials related to this paper are available on request by the corresponding author.

Acknowledgements

All authors conceptualized and designed the study. The first two authors drafted the initial manuscript. The last three authors reviewed and revised the manuscript. All authors approved the final manuscript as submitted and agreed to be accountable for all aspects of the work.

Declaration of conflicting interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship and/or publication of this article.

Funding

The author(s) disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: ZonMw, the Netherlands Organization for Health Research and Development, supported this study (No 7125.0001).

ORCID iD


Bernice Doove https://orcid.org/0000-0002-3189-4276

References

American Academy of Pediatrics. (2006). Identifying infants and young children with developmental disorders in the medical home: An algorithm for developmental surveillance and screening. *Pediatrics*, *118*, 405–420. doi:10.1542/peds.2006-1231

Achenbach, T. M., Becker, A., Dopfner, M., Heiervang, E., Roessner, V., Steinhausen, H. C., & Rothenberger, A. (2008). Multicultural assessment of child and adolescent psychopathology with ASEBA and SDQ instruments: Research findings, applications, and future directions. *Journal of Child Psychology and Psychiatry*, *49*, 251–275. doi:10.1111/j.1469-7610.2007.01867.x

Achenbach, T. M., Howell, C. T., Quay, H. C., & Conners, C. K. (1991). National survey of problems and competencies among four- to sixteen-year-olds: Parents' reports for normative and clinical samples. *Monographs of the Society for Research in Child Development*, *56*(3), 1–131.

Achenbach, T. M., & Ruffle, T. M. (2000). The Child Behavior Checklist and related forms for assessing behavioral/emotional problems and competencies. *Pediatrics in Review*, *21*, 265–271.

Bayer, J. K., Rapee, R. M., Hiscock, H., Ukoumunne, O. C., Mihalopoulos, C., & Wake, M. (2011). Translational research to prevent internalizing problems early in childhood. *Depress Anxiety*, *28*, 50–57. doi:10.1002/da.20743

Berkovits, M. D., O'Brien, K. A., Carter, C. G., & Eyberg, S. M. (2010). Early identification and intervention for behavior problems in primary care: A comparison of two abbreviated versions of parent-child interaction therapy. *Behavior Therapy*, *41*, 375–387. doi:10.1016/j.beth.2009.11.002

Bouter, L. M., & Van Dongen, M. C. J. M. (2005). *Epidemiological research, design and interpretation*. Houten, The Netherlands: Bohn Stafleu van Loghum.

Carter, A. S., Briggs-Gowan, M. J., & Davis, N. O. (2004). Assessment of young children's social-emotional development and psychopathology: Recent advances and recommendations for practice. *Journal of Child Psychology and Psychiatry*, *45*, 109–134.

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Hillsdale, NJ: Lawrence Erlbaum.

Cox, J. E., Huntington, N., Saada, A., Epee-Bounya, A., & Schonwald, A. D. (2010). Developmental screening and parents' written comments: An added dimension to the parents' evaluation of developmental status questionnaire. *Pediatrics*, *126* (Suppl. 3), S170–S176. doi:10.1542/peds.2010-1466M

De Brock, A. J. L. L., Verhulst, A. A., Gerris, J. R. M., & Abidin, R. R. (1992). *NOSI-Nijmegen Parenting Stress Index, Manual experimental version*. Lisse, The Netherlands: Swets en Zweitlinger.

de Wolff, M. S., Theunissen, M. H., Vogels, A. G., & Reijneveld, S. A. et al. (2013). Three questionnaires to detect psychosocial problems in toddlers: A comparison of the BITSEA, ASQ:SE, and KIPPPI. *Academic Pediatrics*, *13*, 587–592. doi:10.1016/j.acap.2013.07.007

Glascoe, F. P. (2001). Are overreferrals on developmental screening tests really a problem? *Archives of Pediatrics and Adolescent Medicine*, *155*, 54–59. doi:poa00202

Glascoe, F. P. (2005). Screening for developmental and behavioral problems. *Mental Retardation and Developmental Disabilities Research Reviews*, *11*, 173–179. doi:10.1002/mrdd.20068

Glascoe, F. P. (2013). *Collaborating with parents: Using parents' evaluation of developmental status (PEDS) to detect and address developmental and behavioural problems* (2nd ed.). Nolensville, TN: Ellsworth & Vandermeer Press.

Glascoe, F. P., & Robertshaw, N. S. (2007). New AAP policy on detecting and addressing developmental and behavioral problems. *Journal of Pediatric Health Care*, *21*, 407–412. doi:10.1016/j.pedhc.2007.08.008

Glascoe, F. P., & Trimm, F. (2014). Brief approaches to developmental-behavioral promotion in primary care: Updates on methods and technology. *Pediatrics*, *133*, 884–897. doi:10.1542/peds.2013-1859

Goodman, R. (2001). Psychometric properties of the strengths and difficulties questionnaire. *Journal of the American Academy of Child and Adolescent Psychiatry*, *40*, 1337–1345. doi:10.1097/00004583-200111000-00015

Goodman, R., Ford, T., Corbin, T., & Meltzer, H. (2004). Using the Strengths and Difficulties Questionnaire (SDQ) multi-informant algorithm to screen looked-after children for psychiatric disorders. *European Child & Adolescent Psychiatry*, *13*(Suppl. 2), II25–II31. doi:10.1007/s00787-004-2005-3

Hawker, G. A., Mian, S., Kendzerska, T., & French, M. (2011). Measures of adult pain: Visual Analog Scale for Pain (VAS Pain), Numeric Rating Scale for Pain (NRS Pain), McGill Pain Questionnaire (MPQ), Short-Form McGill Pain Questionnaire (SF-MPQ), Chronic Pain Grade Scale (CPGS), Short Form-36 Bodily Pain Scale (SF-36 BPS), and Measure of Intermittent and Constant Osteoarthritis Pain (ICOAP). *Arthritis Care & Research*, *63*(Suppl. 11), S240–S252. doi:10.1002/acr.20543

Hielkema, M., de Winter, A. F., de Meer, G., & Reijneveld, S. A. (2011). Effectiveness of a family-centered method for the early identification of social-emotional and behavioral problems in children: A quasi experimental study. *BMC Public Health*, *11*, Article 636. doi:10.1186/1471-2458-11-636

High, P. C. (2008). School readiness. *Pediatrics, 121*, e1008–e1015. doi:10.1542/peds.2008-0079

Huber, M., Knottnerus, J. A., Green, L., van der Horst, H., Jadad, A. R., Kromhout, D., & Smid, H. (2011). How should we define health? *British Medical Journal*, *343*, Article d4163. doi:10.1136/bmj.d4163

Huber, M., van Vliet, M., Giezenberg, M., Winkens, B., Heerkens, Y., Dagnelie, P. C., & Knottnerus, J. A. (2016). Towards a 'patient-centred' operationalisation of the new dynamic concept of health: A mixed methods study. *BMJ Open*, *6*(1), e010091. doi:10.1136/bmjopen-2015-010091

Huisman, M., Oldehinkel, A. J., de Winter, A., Minderaa, R. B., de Bildt, A., Huizink, A. C., & . . . Ormel, J. (2008). Cohort profile: The Dutch 'TRacking Adolescents' Individual Lives' Survey'; TRAILS. *International Journal of Epidemiology*, *37*, 1227–1235. doi:10.1093/ije/dym273

Ivanova, M. Y., Achenbach, T. M., Rescorla, L. A., Harder, V. S., Ang, R. P., Bilenberg, N., & . . . Verhulst, F. C. (2010). Preschool psychopathology reported by parents in 23 societies: Testing the seven-syndrome model of the child behavior checklist for ages 1.5-5. *Journal of the American Academy of Child and Adolescent Psychiatry*, *49*, 1215–1224. doi:10.1016/j.jaac.2010.08.019

Jacobusse, S., van Buuren, S., & Verkerk, P. H. (2006). An interval scale for development of children aged 0–2 years. *Statistics in Medicine*, *25*, 2272–2283.

Jaspers, M., de Winter, A. F., Huisman, M., Verhulst, F. C., Ormel, J., Stewart, R. E., & Reijneveld, S. A. (2012). Trajectories of psychosocial problems in adolescents predicted by findings from early well-child assessments. *Journal of Adolescent Health*, *51*, 475–483. doi:10.1016/j.jadohealth.2012.02.007

Karatsoreos, I. N., & McEwen, B. S. (2013). Annual research review: The neurobiology and physiology of resilience and adaptation across the life course. *Journal of Child Psychology and Psychiatry*, *54*, 337–347. doi:10.1111/jcpp.12054

Kerstjens, J. M., Bos, A. F., ten Vergert, E. M., de Meer, G., Butcher, P. R., & Reijneveld, S. A. (2009). Support for the global feasibility of the Ages and Stages Questionnaire as developmental screener. *Early Human Development*, *85*, 443–447. doi:10.1016/j.earlhumdev.2009.03.001

Limbos, M. M., & Joyce, D. P. (2011). Comparison of the ASQ and PEDS in screening for developmental delay in children presenting for primary care. *Journal of Developmental and Behavioral Pediatrics*, *32*, 499–511. doi:10.1097/DBP.0b013e31822552e9

Marks, K., Glascoe, F. P., Aylward, G. P., Shevell, M. I., Lipkin, P. H., & Squires, J. K. (2008). The thorny nature of predictive validity studies on screening tests for developmental-behavioral problems. *Pediatrics*, *122*, 866–868. doi:10.1542/peds.2007-3142

Marks, K. P., Glascoe, F. P., & Macias, M. M. (2011). Enhancing the algorithm for developmental-behavioral surveillance and screening in children 0 to 5 years. *Clinical Pediatrics*, *50*, 853–868. doi:10.1177/0009922811406263

McGorry, P., & van Os, J. (2013). Redeeming diagnosis in psychiatry: Timing versus specificity. *The Lancet*, *381*, 343–345. doi:10.1016/S0140-6736(12)61268-9

Mokkink, L. B., Terwee, C. B., Patrick, D. L., Alonso, J., Stratford, P. W., Knol, D. L., & . . . de Vet, H. C. (2010). The COSMIN study reached international consensus on taxonomy, terminology, and definitions of measurement properties for health-related patient-reported outcomes. *Journal of Clinical Epidemiology*, *63*, 737–745. doi:10.1016/j.jclinepi.2010.02.006

Nakagawa, S., & Cuthill, I. C. (2007). Effect size, confidence interval and statistical significance: A practical guide for biologists. *Biological Reviews of the Cambridge Philosophical Society*, *82*, 591–605. doi:10.1111/j.1469-185X.2007.00027.x

NICHD. (2004). Affect dysregulation in the mother-child relationship in the toddler years: Antecedents and consequences. *Development and Psychopathology*, *16*, 43–68.

Nussbaum, M. C. (2011). *Creating capabilities: The human development approach*. Cambridge, MA: The Belknap Press of Harvard University Press.

Potijk, M. R., de Winter, A. F., Bos, A. F., Kerstjens, J. M., & Reijneveld, S. A. (2016). Co-occurrence of developmental and behavioural problems in moderate to late preterm-born children. *Archives of Disease in Childhood*, *101*, 217–222. doi:10.1136/archdischild-2015-308958

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. doi:10.1542/peds.2010-2180

Reijneveld, S. A., Brugman, E., Verhulst, F. C., & Verloove-Vanhorick, S. P. (2004). Identification and management of psychosocial problems among toddlers in Dutch preventive child health care. *Archives of Pediatrics and Adolescent Medicine*, *158*, 811–817. doi:10.1001/archpedi.158.8.811

Rutter, M. (2013). Annual research review: Resilience – Clinical implications. *Journal of Child Psychology and Psychiatry*, *54*, 474–487. doi:10.1111/j.1469-7610.2012.02615.x

Salum, G. A., Polanczyk, G. V., Miguel, E. C., & Rohde, L. A. (2010). Effects of childhood development on late-life mental disorders. *Current Opinion in Psychiatry*, *23*, 498–503. doi:10.1097/YCO.0b013e32833ead33

Sameroff, A. (2010). A unified theory of development: A dialectic integration of nature and nurture. *Child Development*, *81*, 6–22. doi:10.1111/j.1467-8624.2009.01378.x

Skovgaard, A. M. (2010). Mental health problems and psychopathology in infancy and early childhood: An epidemiological study. *Danish Medical Bulletin*, *57*(10), Article B4193.

Snyderman, R., & Langheier, J. (2006). Prospective health care: The second transformation of medicine. *Genome Biology*, *7*(2), 104. doi:10.1186/gb-2006-7-2-104

Sroufe, L. A., Egeland, B., Carlson, E., & Collins, W. A. (2005). *The development of the person: The Minnesota study of risk and adaptation from birth to adulthood*. New York, NY: Guilford Publications.

Staal, I. I., Hermanns, J. M., Schrijvers, A. J., & van Stel, H. F. (2013). Risk assessment of parents' concerns at 18 months in preventive child health care predicted child abuse and neglect. *Child Abuse & Neglect*, *37*, 475–484. doi:10.1016/j.chiabu.2012.12.002

StataCorp. (2013). *Stata statistical software: Release 13*. College Station, TX: StataCorp LP.

StataCorp. (2015). *Statabase reference manual* (pp. 1029–1051). College Station, TX: Stata Press.

Stone, L. L., Otten, R., Engels, R. C., Vermulst, A. A., & Janssens, J. M. (2010). Psychometric properties of the parent and teacher versions of the strengths and difficulties questionnaire for 4- to 12-year-olds: A review. *Clinical Child and Family Psychology Review*, *13*, 254–274. doi:10.1007/s10567-010-0071-2

Syurina, E., Hens, K., & Feron, F. (2013). Literature review: Use of family history for primary paediatric care as the next step towards use of genomics in healthcare. *Current Pediatric Reviews*, *94*, 353–372.

Theunissen, M. H. C. (2013). *The early detection of psychosocial problems in children aged 0–6 years by Dutch preventive child health care: Professionals and their tools*. Groningen:s.n.

Theunissen, M. H. C. (2016). *Manual for the use of the SDQ within the Youth Health Care. Questionnaire for signaling psychosocial problems among 3-17 years old*. The Netherlands, Leiden: TNO.

Theunissen, M. H. C., Vogels, A. G., de Wolff, M. S., Crone, M. R., & Reijneveld, S. A. (2015). Comparing three short questionnaires to detect psychosocial problems among 3 to 4-year olds. *BMC Pediatrics*, *15*, Article 84. doi:10.1186/s12887-015-0391-y

Tick, N. T., van der Ende, J., Koot, H. M., & Verhulst, F. C. (2007). 14-year changes in emotional and behavioral problems of very young Dutch children. *Journal of the American Academy of Child and Adolescent Psychiatry*, *46*, 1333–1340. doi:10.1097/chi.0b013e3181337532

Tough, S. C., Siever, J. E., Leew, S., Johnston, D. W., Benzies, K., & Clark, D. (2008). Maternal mental health predicts risk of developmental problems at 3 years of age: Follow up of a community based trial. *BMC Pregnancy Childbirth*, *8*, Article 16. doi:10.1186/1471-2393-8-16

van Esso, D., del Torso, S., Hadjipanayis, A., Biver, A., Jaeger-Roman, E., Wettergren, B., . . . Primary-Secondary Working Group of European Academy of Paediatrics. (2010). Paediatric primary care in Europe: Variation between countries. *Archives of Disease in Childhood*, *95*, 791–795. doi:10.1136/adc.2009.178459

van Widenfelt, B. M., Goedhart, A. W., Treffers, P. D., & Goodman, R. (2003). Dutch version of the Strengths and Difficulties Questionnaire (SDQ). *European Child & Adolescent Psychiatry*, *12*, 281–289. doi:10.1007/s00787-003-0341-3

Verhulst, F. C., van der Ende, J., & Koot, H. M. (1996). *Handleiding voor de Child Behaviour Checklist (CBCL)/4-18 (Manual for the CBCL/4-18)*. Rotterdam, The Netherlands: Department of Child and Adolescent Psychiatry, Sophia Children's Hospital, Academisch Ziekenhuis Rotterdam/Erasmus Universiteit Rotterdam, Erasmus University.

Verhulst, F. C., van der Ende, J., & Koot, H. M. (1997). *Handleiding voor de Teachers's Report Form (TRF)* (Manual for the TRF). Rotterdam, The Netherlands: Department of Child and Adolescent Psychiatry, Sophia Children's Hospital, Academisch Ziekenhuis Rotterdam/ Erasmus Universiteit Rotterdam, Erasmus University.

Vreeswijk, C. M., Rijk, C. H., Maas, A. J., & van Bakel, H. J. (2015). Fathers' and mothers' representations of the infant: Associations with prenatal risk factors. *Infant Mental Health Journal*, *36*, 599–612. doi:10.1002/imhj.21541

Williams, P. G., Okamoto, J., AAP Council on Early Childhood, & AAP Council on School Health. (2016). The pediatrician's role in optimizing school readiness. *Pediatrics*, *138*(3), e20162293. doi:10.1542/peds.2016-2293

Woolfenden, S., Eapen, V., Williams, K., Hayen, A., Spencer, N., & Kemp, L. (2014). A systematic review of the prevalence of parental concerns measured by the Parents' Evaluation of Developmental Status (PEDS) indicating developmental risk. *BMC Pediatrics*, *14*, Article 231. doi:10.1186/1471-2431-14-231

Author biographies

Bernice Doove is a Medical Specialist in Community Health and Social Medicine in the field of Child and Youth Health Care. She works as a medical doctor in Maastricht and participates as a PhD candidate in the Academic Collaborative Centre for Public Health Limburg.

Jolien Feron is a Psychiatrist and and works in a Flexible Assertive Community Treatment (FACT) team of MET ggz in Roermond, providing care for people with severe mental illness, of which the majority is diagnosed with a psychotic disorder.

Frans Feron is Professor of Social Medicine, in particular Child and Adolescent Health, and chair of the Department of Social Medicine, Maastricht University. Areas of interest include: neurodevelopmental disorders, public health genomics and child maltreatment.

Jim van Os is Professor of Psychiatric Epidemiology and Public Mental Health, and chair of the Division Neuroscience at Utrecht University Medical Centre. His focus is on the clinical, developmental and genetic epidemiology of (early states of) mental disorders.

Marjan Drukker is a Senior Epidemiologist at the Department of Psychiatry, School for Mental Health and NeuroScience (MHeNS), Maastricht University. Her areas of interest are gene-environment interactions and networks of psychological symptoms in psychiatry.