Preschool children with ODD, CD and ADHD

Psychiatric assessment and stability of diagnosis

Tessa L. Bunte-Rosingh

Author: Tessa L. Bunte-Rosingh
Lay-out en drukwerk: Ridderprint BV, Ridderkerk
Lay-out and printing by: Ridderprint BV, Ridderkerk, the Netherlands

ISBN: 978-90-5335-850-4

© 2014, T.L. Bunte-Rosingh
All rights reserved. No part of this thesis may be reproduced or transmitted in any form or by any means without the prior written permission of the author, or, where appropriate, of the publisher of the articles.

Preschool children with ODD, CD and ADHD

Psychiatric assessment and stability of diagnosis

Kleuters met ODD, CD en ADHD

Psychiatrisch onderzoek en stabiliteit van diagnose

(met een samenvatting in het Nederlands)

Proefschrift

ter verkrijging van de graad van doctor aan de Universiteit Utrecht
op gezag van de rector magnificus, prof. dr. G.J. van der Zwaan,
ingevolge het besluit van het college voor promoties
in het openbaar te verdedigen op
dinsdag 17 juni 2014 des ochtends te 10.30 uur

door

Tessa Leonoor Bunte-Rosingh
geboren op 31 juli 1974
te Curaçao

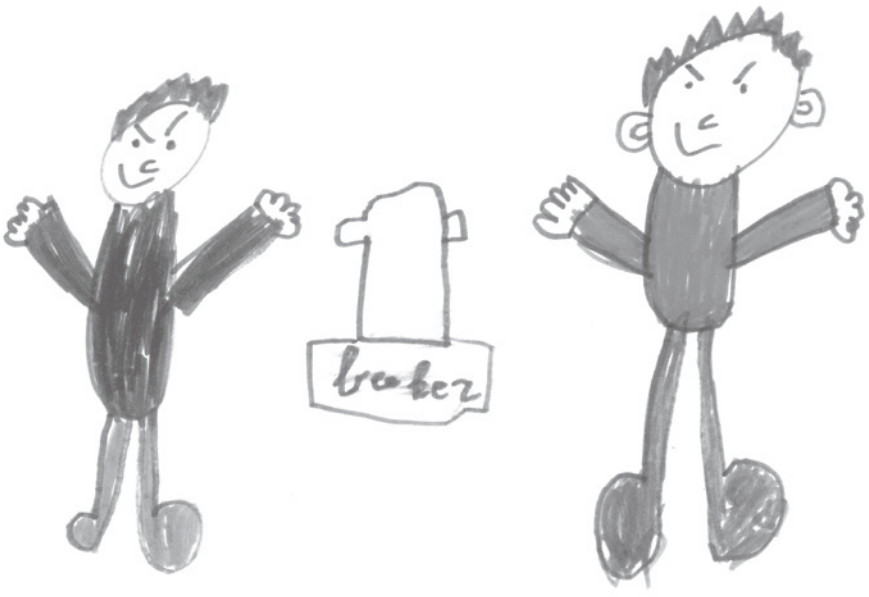
Promotoren: Prof. dr. W. Matthys
Prof. dr. P.G.M. van der Heijden

Voor mijn ouders

TABLE OF CONTENTS

CHAPTER 1

General Introduction



ODD, CD AND ADHD WITHIN THE PRESCHOOL AGE

There is an increasing number of young children referred to mental health clinics (Howell & Teich, 2008). Among these children attention deficit and disruptive behavior problems are the most common reason for referral (Wilens et al., 2002). Epidemiological studies have demonstrated that the prevalence rates of Oppositional Defiant Disorder (ODD), Conduct Disorder (CD) (also called disruptive behavior disorder; DBD) and Attention-Deficit/Hyperactivity Disorder (ADHD) (American Psychiatric Association, 2000) in preschool children are similar to those later in childhood (Egger & Angold, 2006). Moreover, it has been shown that the symptoms and associated impairment of preschool children diagnosed with ODD, CD and ADHD are likely to persist into school age (Keenan et al., 2011; Lahey et al., 2004).
However, clinicians have long been reluctant to diagnose young children with these disorders. It has indeed been questioned how to differentiate clinical from normative transient disruptive behaviors within the preschool age as most preschoolers exhibit at least some of the behaviors that fall under the rubric of disruptive behavior, for example losing temper and physical aggression (Keenan et al., 2007). A point of concern is the possibility of over-identification of preschool children as having ODD, CD or ADHD. In summary, the clinical assessment of preschool behavioral problems is particularly subtle and complex (Wakschlag et al., 2005).

PSYCHIATRIC ASSESSMENTS IN PRESCHOOL CHILDREN

Recently, several assessment procedures have been developed to improve the diagnostic process for preschoolers with DBD and ADHD. Diagnosing behavioral problems in preschool children requires a combination of information about the child's behavior from parents and teachers/caregivers, and clinical observation. Parent reports are particularly important as they are based on children's behaviors across a long period of time. Teacher/caregiver reports are important as well for they are based on information of children's behavior in another setting than the home. Clinical observation, which is not filtered through the perceptions of the parent, may provide a different window on the child's functioning (Le Couteur & Gardner, 2008) and is essential to distinguish normative maladaptive behavior from DBD and ADHD symptoms.

DB-DOS

Wakschlag et al. (2008a, 2008b) developed the Disruptive Behavior Diagnostic Observation Schedule (DB-DOS), i.e., a standardized observational assessment of the preschool

child for diagnosing ODD and CD. The DB-DOS is specifically designed to examine multiple, qualitative and quantitative aspects of preschool children's behavior and socio-emotional functioning hypothesized to distinguish normative and disruptive behavior. Child behavior is assessed in three interactional contexts: one parent context and two examiner contexts; (Wakschlag et al., 2008a, 2008b). The DB-DOS has been shown to be a reliable and valid tool to distinguish normative maladaptive behaviors from DBD symptoms (Wakschlag et al., 2008a, 2008b). However, for clinical use, i.e., to identify individual clinical cases, it requires further research. Furthermore, in research on the DB-DOS, no attention has been paid to ADHD symptoms while the correlation between DBD symptoms and hyperactivity-impulsivity in preschool children is 0.79 (Sterba, Egger, & Angold, 2007). Therefore, while assessing DBD symptoms in referred young children with behavior problems, it seems evident to also include the assessment of ADHD symptoms in the DB-DOS

K-DBDS

In addition to observational assessment for the preschool child, (semi-)structured interviews with the parents are developed as well for preschoolers, such as the Kiddie-Disruptive Behavior Disorder Schedule (K-DBDS; Keenan et al., 2007). The K-DBDS has the advantage of eliciting information necessary to make clinical diagnoses, based on DSM-IV-TR criteria (American Psychiatric Association, 2000). The K-DBDS has been shown to have satisfactory interrater and test-retest reliability with regard to ODD and CD symptoms (Keenan et al., 2007). Validity was demonstrated via associations with impairment and the differentiation between referred and nonreferred children (Keenan et al., 2007). Thus, most DSM-IV-TR ODD and CD symptoms can be reliably and validly assessed in preschool children using the K-DBDS. In contrast, the reliability and validity of ADHD symptoms have not been investigated thus far using the K-DBDS.

Another issue in the diagnosis of ODD, CD and ADHD, is the total number of symptoms required for a diagnosis. Among preschoolers this issue is even more important as most preschoolers exhibit at least some of the behaviors that fall under the rubric of disruptive behavior. The DSM-5 (American Psychiatric Association, 2013) proposed to keep the same thresholds for ODD, CD and ADHD as used with the DSM-IV-TR (American Psychiatric Association, 2000). However, for ODD, Angold & Costello (1996) have shown that children with two or three symptoms (plus impairment) are similar to those with four or more symptoms in terms of outcome 1 year later. For ADHD, although results of the field trial suggested a threshold of five hyperactivity-impulsivity symptoms, the DSM-IV Child Disorder Workgroup decided on a six out of nine symptom threshold for the ADHD symptoms of hyperactivity-impulsivity to reduce false positives. With regard to ADHD, the issue is whether, in the definition of symptoms, pervasiveness across contexts should be included or not. The DSM-IV-TR criteria for ADHD require that some impairment from the

symptoms is present in two or more settings (e.g., at school/work and at home) (American Psychiatric Association, 2000). It is, however, unclear why children who meet all diagnostic criteria for ADHD but display impairment in only one setting (e.g., seriously impaired academic functioning) should not be given the diagnosis of ADHD. Although functional impairment is a diagnostic criterion for all mental disorders, it has been questioned why impairment in multiple settings should be required only for ADHD (Lahey et al., 2004).

STABILITY AND CHANGE IN DIAGNOSES

Evidence for the stability of ODD, CD and ADHD diagnoses is critical for refuting the claim that these disorders pathologize normative behavior in young children (Keenan et al., 2011). At the preschool age there are a few studies that have shown evidence for the stability of these disorders from preschool to elementary school age, but change as well (Keenan et al., 2011; Lahey et al., 2004; Riddle et al., 2013). Results of these studies support the stability of ODD, CD and ADHD diagnosis in preschool children, but they also point to remission of diagnosis from preschool to school age. The latter, however, seems to be underestimated in the literature. Moreover, the studies did not examine whether preschool children who were not diagnosed at baseline reach the symptom threshold at later assessments. In summary, further study is needed into the changes of diagnosis in the preschool and early elementary school years.

FACTORS PREDICTING STABILITY OF DIAGNOSIS

If diagnosis of ODD, CD and ADHD change in the preschool and early elementary school years, it is important to find out which factors predict stability of diagnosis. Several non-clinical studies examined which individual and family factors are associated with the persistence of behavior problems and hyperactivity/attention problems. Individual factors have been identified such as initial symptom level, gender, IQ, behavioral inhibition, and the occurrence of symptoms at daycare or school. Likewise, family factors such as SES, stress, maternal depression, and harsh and inconsistent discipline appear to be associated with the persistence of behavior problems and hyperactivity/attention problems. In view of clinical decision making, in particular with regard to starting treatment, it is important to take into account which factors are associated with stability of diagnosis.

OUTPATIENT CLINIC

The aim of this dissertation was, first, to study the clinical usefulness of an observational schedule, the DB-DOS, and a semi-structured interview, the K-DBDS, at the preschool age, to diagnose young children with behavioral problems. The second aim was to study the stability and change of ODD, CD and ADHD diagnosis at the preschool age, and the factors that predict stability of ODD, CD and ADHD diagnosis in the preschool and early elementary school years.

To address these aims, an outpatient clinic for preschool children with externalizing behavioral problems was started at the Department of Child and Adolescent Psychiatry, University Medical Center Utrecht in 2007. Children were assessed by a multi-disciplinary team in one morning, and in order to evaluate their development they were re-assessed after 9 and 18 months. Children were referred by general practitioners, well-baby clinics and pediatricians. Inclusion criteria were age between 3.5 and 5.5 years and a score at or above the 90th percentile either on the Attention Problems scale or the Aggressive Behavior scale of the Child Behavior Checklist completed by parents (CBCL/1.5–5) or the Child Teacher Report Form completed by teachers or day-care caregivers (C-TRF/1.5–5); (both: Achenbach & Rescorla, 2000; Dutch version by Verhulst & Van der Ende, 2000). Furthermore, children with an IQ below 70 were excluded, estimated with the Raven colored progressive matrices (Raven, Court, & Raven, 1998) and the Peabody Picture Vocabulary-III-NL (Dunn & Dunn, 2005; Dutch translation by Schlichting, 2005).

Children were evaluated in a single, morning session. First, an interview with the parents regarding the child's behavior took place with the child psychiatrist or clinical child psychologist. Second, two measures of intellectual functioning were administered, i.e, the Raven colored progressive matrices (Raven, Court, & Raven, 1998) and the Peabody Picture Vocabulary-III-NL (Dunn & Dunn, 2005; Dutch translation by Schlichting, 2005), followed by the EF tasks. We included EF tasks that were designed to measure working memory, inhibition and cognitive flexibility (Go-No-Go task, Snack Delay and Shape School; Schoemaker et al., 2012). At the same time, parents completed questionnaires including questionnaires regarding parenting practices (Parenting Practices Interview; Webster-Stratton, 2001), parental stress (Parenting Stress Index; Abadin, 1983), parental depressive symptoms (Beck Depression Inventory; Beck, Ward, Mendelson, Mock & Erbaugh, 1961) and ADHD symptoms (Kooij & Buitelaar, 1997).

The tasks were administered in a fixed order and lasted about two hours, including breaks. After another break the child's behavior was observed using the Disruptive Behavior Diagnostic Observation Schedule (DB-DOS; Wakschlag et al., 2008a, 2008b). Finally, the K-DBDS (K-DBDS; Keenan et al., 2007) and the Children's Global Assessment Scale (Schaffer et al., 1983) were administrated. The children were re-assessed after 9 and 18 months.

The second assessment was shorter and did not include the IQ tasks, some questionnaires (regarding maternal depression and parental ADHD symptoms) and the DB-DOS. The assessment after 18 months was similar to the first assessment and did not include the IQ tasks and some questionnaires (regarding maternal depression and parental ADHD symptoms).
The TD group was recruited from regular primary schools and daycare centers. Children with a score at or above the 90th percentile on the Attention Problems scale or the Aggressive Behavior scale of the CBCL and C-TRF were excluded. The same procedure was followed but without the initial interviews with a child psychiatrist or clinical child psychologist.

AIMS AND OUTLINE

The first aim of this dissertation was to study the clinical usefulness of the DB-DOS (Chapter 2) and the K-DBDS (Chapter 3) at the preschool age to diagnose young children with behavioral problems. Second, the stability and change of ODD, CD and ADHD diagnosis at the preschool age (chapter 4) was studied as well as the factors predicting stability of ODD, CD and ADHD diagnosis and persistence of symptoms in preschool children (chapter 5). In Chapter **two**, we focus on the clinical usefulness of the DB-DOS in the diagnosis of DBD and ADHD in preschool children in a European sample of referred preschool children with behavioral problems; a typically developing (TD) control group was also included in the study. We then focus (Chapter **three**) on the clinical usefulness of the K-DBDS. Two issues with regard to the operationalization of symptoms are crucial here: 1. whether the definition of ODD symptoms should be based either on the threshold “often” or on a more specific frequency threshold; 2. whether, in the definition of ADHD symptoms, pervasiveness across contexts should be included or not.
With regard to the stability and change of ODD, CD and ADHD diagnosis, Chapter **four** reports the study of the stability and change of ODD, CD and ADHD HI diagnosis in terms of percentages of children being diagnosed again at 9 months and 18 months follow-up. We also anticipated that referred preschool children with externalizing behavior problems not diagnosed at the first assessment may be diagnosed at 9 months and/or 18 months follow-up. The second aim was to examine the predictive validity of an ODD, CD and ADHD HI diagnosis. The third aim of that study was to examine impairment and number of symptoms in groups that differed in stability of diagnosis (chronic, partial remission, full remission, new onset, no diagnosis).
Chapter **five**, reports the study of individual factors (initial symptom level, gender, IQ, behavioral inhibition, occurrence of symptoms at daycare or school) and which family

factors (SES, stress, maternal depression, harsh and inconsistent discipline) as predictors of stability of ODD, CD and ADHD diagnosis in the preschool and early elementary school period. We also investigated which factors predict persistence of ODD, CD and ADHD symptoms. Therefore, we investigated whether the above mentioned factors predicted 1. ODD, CD and ADHD diagnosis according to the Kiddie-Disruptive Behavior Diagnostic Schedule (K-DBDS; Keenan et al., 2007, Bunte et al., 2013b) at 18 months follow-up, 2. the total symptom score of ODD, CD and ADHD symptoms on the K-DBDS.
Chapter **six** summarizes and discusses the findings of the studies.

REFERENCES

Abidin, R.R. (1983). Parental Stress Index: Manual. Charlottesville: Pediatric Psychology Press.

Achenbach, T.M., & Rescorla, L.A. (2000). *Manual for the ASEBA preschool forms and profiles.* Burlington: Research Center for Children, Youth & Families VT: University of Vermont.

Angold, A. & Costello, E.J. (1996). Toward establishing an empirical basis for the diagnosis of oppositional defiant disorder. *Journal of the American Academy of Child and Adolescence Psychiatry, 35*, 1205-12.

American Psychiatric Association. (2000). *Diagnostic and statistical manual of mental disorders (4th Edition, text-revised)*. Washington, DC: Author.

American Psychiatric Association. (2013). *Diagnostic and statistical manual of mental disorders (5th Edition)*. Washington, DC: Author.

Beck, A., Ward, C., Mendelson, M., Mock, J., & Erbaugh, J. (1961). An inventory for measuring depression. *Archives of general psychiatry, 4,* 561-571.

Dunn, L.M., & Dunn, L.M. (2005). *Peabody picture vocabulary test (III-NL)* [Dutch version by L. Schlichting]. Amsterdam: Harcourt Assessment.

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.

Howell, E.M., & Teich, J. (2008). Variations in medicaid mental health service use and cost for children. *Administration and Policy in Mental Health, 35,* 220-8.

Keenan, K., Boeldt, D., Chen, D., Coyne, C., Donald, R., Duax J., …Humphries, M. (2011). Predictive validity of DSM-IV oppositional defiant and conduct disorders in clinically referred preschoolers. *Journal of Child Psychology and Psychiatry, 52,* 47-55.

Keenan, K., Wakschlag, L.S., Danis, B., Hill, C., Humphries, M., Duax, J., & Donald, R. (2007). Further evidence of the reliability and validity of DSM-IV ODD and CD in preschool children. *Journal of the American Academy of Child and Adolescent Psychiatry, 26,* 457-468.

Kooij, J.J.S., & Buitelaar, J.K. (1997). Zelfrapportagelijst aandachtsproblemen en hyperactiviteit voor volwassenheid en kindertijd. In: J.J.S. Kooij, ADHD bij volwassenen. Diagnostiek en behandeling. Pearson Assessment and Information BV.

Lahey, B., Pelham, W.E., Loney, D., Kipp, H., Ehrhardt, A., Lee, S, …Massetti, G. (2004). Three-Year predictive validity of DSM-IV attention deficit hyperactivity disorder in children diagnosed at 4-6 years of age. *American Journal of Psychiatry, 161,* 2014-20.

Le Couteur, A., & Gardner, F. (2008). Use of structured interviews and observational methods in clinical settings. In M. Rutter, D. Bishop, D. Pine, S. Scott, J. Stevenson, E. Taylor & A. Thapar (Eds.), Rutter's Child and Adolescent Psychiatry 5th Ed. (pp. 271-288). Oxford: Blackwell.

Raven, J.C., Court, J.H., & Raven, J. (1998). *Raven coloured progressive matrices*. Oxford: Oxford Psychologist Press.

Riddle, M.A., Yershova, K., Lazzaretto, D., Paykina, N., Yenokyan, G. Greenhill, L. …Posner, K. (2013). The preschool Attention-Deficit/Hyperactivity Disorder treatment study (PATS) 6-year follow-up. *Journal of the American Academy of Child and Adolescent Psychiatry, 52,* 264-278.

Schaffer, D., Goud, M.S., Brasic, J., Ambrosini, P., Fisher, P., Bird, H., & Aluwahlia, S. (1983). A children's global assessment scale (CGAS). *Archives of General Psychiatry*, *40*, 1228-1231.

Schlichting, L. (2005). Peabody picture vocabulary test III-NL. Amsterdam: Hartcourt Assessment.

Schoemaker, K., Bunte, T., Wiebe, S.A., Espy, K.A., Dekovic, M., & Matthys, W. (2012). Executive function deficits in preschool children with ADHD and DBD. *Journal of Child Psychology and Psychiatry*, *53(2)*, 111-119.

Sterba, S., Egger, H.L., & Angold, A. (2007). Diagnostic specificity and nonspecificity in the dimensions of preschool psychopathology. *Journal of Child Psychology and Psychiatry, 48*, 1005-1013.

Wakschlag, L. S., Hill, C., Carter, A. S., Danis, B., Egger, H. L., Keenan, K., …Briggs -Gowan, M.J. (2008a). Observational Assessment of Preschool Disruptive Behavior, Part I: Reliability of the Disruptive Behavior Diagnostic Observation Schedule (DB-DOS). *Journal of the American Academy of Child and Adolescent Psychiatry, 47,* 622-631.

Wakschlag, L. S., Briggs-Gowan, M. J., Hill, C., Danis, B., Leventhal, B.D., Keenan, K., …Carter, A.S. (2008b). Observational Assessment of Preschool Disruptive Behavior, Part II: Validity of the Disruptive Behavior Diagnostic Observation Schedule (DB -DOS). *Journal of the American Academy of Child and Adolescent Psychiatry, 47,* 632-640.

Wakschlag, L.S., Leventhal, B.L., Briggs-Gowan, M.J., Danis, B., Keenan, K., Hill, C., …Carter, A.S. (2005). Defining the "disruptive" in preschool behavior: what diagnostic observation can teach us. *Clinical Child and Family Psychology Review, 8,* 183-201.

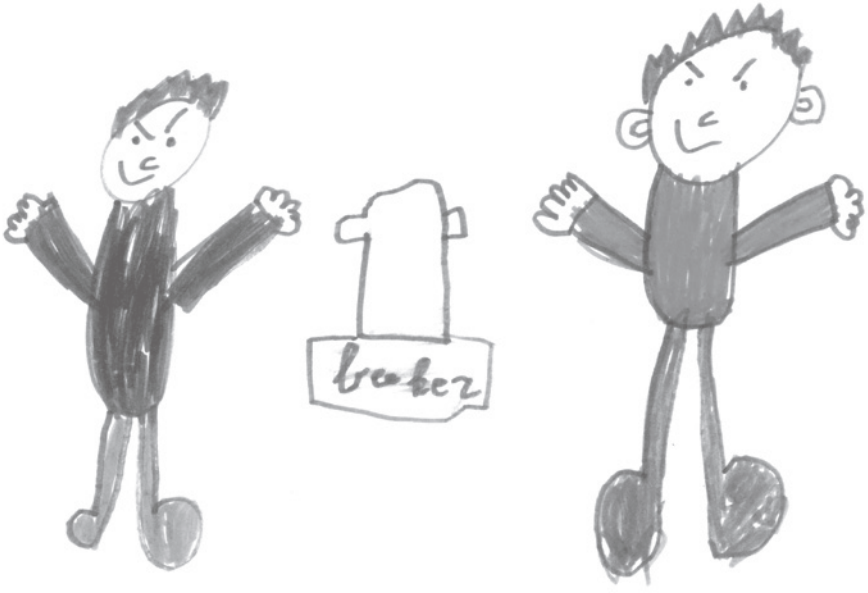
Wilens, T.E., Biederman, J., Brown, S., Monuteaux, M., Prince, J., & Spencer, T.J. (2002). Patterns of psychopathology and dysfunction in clinically referred preschoolers. *Journal of Developmental & Behavioral Pediatrics, 23*, S31-6.

Webster-Stratton, C. (2001). Parenting practices interview. Unpublished assessment instrument. (http://www.son.washington.edu/centers/parenting-clinic/forms.asp).

CHAPTER 2

Clinical usefulness of observational assessment in the diagnosis of DBD and ADHD in preschoolers

Tessa L. Bunte, M.D., Sarah Laschen, M.S., Kim Schoemaker, M.S., David J. Hessen, Ph.D., Peter G.M. van der Heijden, Ph.D., Walter Matthys, M.D., Ph.D.



Bunte, T.L., Laschen, S., Schoemaker, K., Hessen, D.J., van der Heijden, P.G.M., Matthys, W. (2013). Clinical usefulness of observational assessment in the diagnosis of DBD and ADHD in preschoolers. *Journal of Clinical Child & Adolescent Psychology*, *42,* 749-61.

ABSTRACT

The aim of the present study was to investigate the clinical usefulness of an observational tool, i.e., the Disruptive Behavior Diagnostic Observation Schedule (DB-DOS), in the diagnosis of disruptive behavior disorders (DBD) and attention-deficit/hyperactivity disorder (ADHD) in preschoolers. We hypothesized that the DB-DOS may help support the presumption of a diagnosis generated by the information from parents and teachers (or other caregivers). Method: Participants were referred preschool children with externalizing behavioral problems (N=193; 83% male) and typically developing (TD) children (N=58; 71% male). In view of the clinical validity study each child was given a diagnosis of either DBD (N=40), or ADHD (N=54) or comorbid (DBD+ADHD) (N=66) based on best-estimate diagnosis. Results: The DB-DOS demonstrated good interrater and test-retest reliability for DBD and ADHD symptom scores. Confirmatory factor analysis demonstrated an excellent fit of the DB-DOS multi-domain model of DBD symptom scores and a satisfactory fit of ADHD symptom scores. The DB-DOS demonstrated good convergent validity, moderate divergent validity and good clinical validity on a diagnostic group level for DBD and ADHD symptom scores. The Receiver Operating Characteristic curve analyses revealed that for DBD the sensitivity and specificity are moderate and for ADHD good to excellent. The presumption of a diagnosis based on information from parents, teachers and cognitive assessment was supported by the DB-DOS in 60% for DBD and 75% for ADHD. Conclusion: The DB-DOS can be used to help support a presumption of a DBD and/or ADHD diagnosis in preschool children.

Keywords: preschoolers, DBD, ADHD, assessment, behavioral observation

INTRODUCTION

There is an increasing number of young children referred to mental health clinics (Howell & Teich, 2008). Among these children attention deficit and disruptive behavior problems are the most common reason for referral (Wilens et al., 2002). Epidemiological studies have demonstrated that the prevalence rates of Oppositional Defiant Disorder (ODD), Conduct Disorder (CD) and Attention-Deficit/Hyperactivity Disorder (ADHD) (American Psychiatric Association, 2000) in preschool children are similar to those later in childhood (Egger & Angold, 2006). Furthermore, there is evidence of the predictive validity of ODD and CD in clinically referred preschool children (Keenan et al., 2011). Likewise, the validity of the Diagnostic and Statistical Manual of Mental Disorders (4th ed. [DSM-IV-TR]; American Psychiatric Association, 2000) diagnosis of ADHD in younger children has been supported by a study demonstrating that the symptoms and associated impairment are likely to persist well into elementary school (Lahey et al., 2004).

Yet clinicians have long been reluctant to diagnose young children with these disorders. It has indeed been questioned how to differentiate clinical from normative transient disruptive behaviors within the preschool age, as most preschoolers exhibit at least some of the behaviors that fall under the rubric of disruptive behavior. This makes clinical assessment of preschool behavioral problems particularly subtle and complex (Wakschlag et al., 2005). In connection with this, a point of concern is the possibility of over-identification of preschool children as having ODD, CD or ADHD as the manifestations of most of the symptoms of these disorders are not atypical per se when observed in a preschooler, for example losing his temper and physical aggression (Keenan et al., 2007). Clearly, diagnosing a young child with, for example, ODD comorbid with ADHD, and starting treatment, e.g., pharmacotherapy with methylphenidate, in the absence of these disorders is inappropriate.

Diagnosing young children with ODD, CD and ADHD is a challenge. Recently, however, assessment procedures have been developed, including (semi)structured interviews with the parents such as the Kiddie-Disruptive Behavior Disorder Schedule (K-DBDS; Keenan et al., 2007) and the Preschool Age Psychiatric Assessment (PAPA; Egger et al., 2006). Although these newly developed parent-report diagnostic interviews have advanced identification of preschool psychopathology, the limitations of sole reliance on parents as informants are amplified during the preschool years because young children cannot serve as informants about their own behavior (Wakschlag et al., 2008a). Furthermore, although parent reports are based on children's behaviors on a daily basis parent observations may be biased due to a number of factors such as personality characteristics (Collishaw, Goodman, Ford, Rabe-Hesketh & Pickels, 2009). Therefore, direct observation which is not filtered through the perceptions of the parent may provide a different window on the

child's functioning (Le Couteur & Gardner, 2008). Our clinical experience suggests that many parents desire the clinician to observe the child in order to arrive at a diagnosis as parents, who consult a clinician would not accept that the decision on the presence or absence of a disorder would be based only on parent and teacher reports; parents want clinicians to "look at their child" themselves.

For clinicians who need to make a decision whether a child is diagnosed with a disorder or not, a crucial issue in the use of direct observation, however, is how to weigh information from observation against information from parents and teachers. Clearly, a limitation of observation is its brevity and contextual dependency; in other words, the behaviors observed need not be representative of all problem behaviors (Le Couteur & Gardner, 2008). Even though clinicians do not expect to observe each symptom, they may want to see at least "the tip of the iceberg". Thus, while the presumption of the presence of ODD or ADHD may be generated based on information from parents and teachers, direct observation may be used to support this presumption or not.

Various observational tools have been developed such as The Dyadic Parent–Child Interaction Coding System (DPICS; Robinson & Eyberg, 1981). This is a short (fifteen minutes) observational assessment of parent-child interaction which is sensitive to treatment change and therefore has often been used to evaluate behavioral parent training for young children with conduct problems (e.g., Posthumus et al., 2012). The DPICS is an observational assessment, which is parent-based, not examiner based. The DPICS is an observational assessment focusing not only on child behavior but also on parent behavior; the DPICS is parent-based, not examiner-based. Examiner-based assessments are designed to be clinically sensitive by standardizing adult responses in a manner that presses for a range of clinically salient behaviors in the child; however, they lack the ecological validity of parent-child assessments. Therefore, combining examiner- and parent-based behavioral observation paradigms, as the DB-DOS does, provides complementary methods for incorporating the interactive nature of social behavior into the assessment of clinical significance (Wakschlag et al., 2008a). Besides, Wakschlag et al., (2008a, 2008b) developed the more extended (sixty minutes) Disruptive Behavior Diagnostic Observation Schedule (DB-DOS), i.e., a standardized observational assessment of the preschool child for diagnosing ODD and CD (also called disruptive behavior disorder; DBD). The DB-DOS is specifically designed to examine multiple, qualitative and quantitative aspects of preschool children's behavior and socio-emotional functioning hypothesized to distinguish normative and disruptive behavior. Child behavior is assessed in three interactional contexts: one parent context and two examiner contexts; (Wakschlag et al., 2008a, 2008b).

The DB-DOS has been shown to be a reliable and valid tool to distinguish normative maladaptive behaviors from DBD symptoms (Wakschlag et al., 2008a, 2008b). However, for clinical use, i.e., to identify individual clinical cases, it requires further research. Indeed,

Wakschlag et al. (2008b) investigated the clinical validity of the DB-DOS on a group level and not on an individual level, which is required in order to identify individual clinical cases. Furthermore, in research on the DB-DOS no attention has been paid to ADHD symptoms while the correlation between DBD symptoms and hyperactivity-impulsivity in preschool children is 0.79 (Sterba, Egger, & Angold, 2007). Therefore, while assessing DBD symptoms in referred young children with behavior problems, it seems evident to also include the assessment of ADHD symptoms in the DB-DOS**.**

The main aim of the present study was to examine the clinical usefulness of the DB-DOS in the diagnosis of DBD and ADHD in preschool children. Our main hypothesis was that the DB-DOS may help support the presumption of a diagnosis generated by the information from parents and teachers (or other caregivers). First, we examined the reliability and validity of DBD and ADHD symptom scores in a European sample of referred preschool children with behavioral problems; a typically developing (TD) control group was also included in the study. Second, we examined clinical validity of DB-DOS symptom scores on a group level by comparing the TD group with DBD and ADHD groups. Third, we examined clinical validity on an individual level by determining cutoff points of DBD and ADHD symptom scores which divide clinical from TD cases. Finally, we examined how often a diagnosis of DBD and/or ADHD based on information from parents, teachers and cognitive assessment was confirmed by the DB-DOS.

METHOD

Participants

Participants were 251 children aged 3.5-5.5 years (mean 55 months, SD 7.8); 80% were male, 86% White, 12% Turkish/Moroccans, 2% African American, 0.5% Asian. In terms of parent education Primary accounted for 6%, Secondary for 25%, Intermediate vocational for 34%, Higher vocational for 16% and University for 19%. Children with behavior problems (N=193) were referred by general practitioners, pediatricians and well-baby clinics for clinical assessment at the Outpatient Clinic for Preschool Children with Behavioral Problems, Department of Child and Adolescent Psychiatry, Utrecht, the Netherlands.

Inclusion criteria for the referred patient group were a score above the 90-th percentile either on the Aggressive Behavior scale or on the Attention Problems scale of the CBCL/1.5-5 or C-TRF/1.5-5 (Achenbach & Rescorla, 2000). The typically developing (TD) group (N=58) was recruited from regular elementary schools and daycare centers. Children with a score in the clinical range either on the Attention Problems scale or on the Aggressive Behavior scale of the CBCL/1.5-5 or C-TRF/1.5-5 were excluded. All the children (referred and TD) with an IQ below 70, estimated with the average score on the Raven Coloured Progressive

Table 1 *Means (and SD) for the Demographics and Control Variables in the Four Groups*

	TD (N=58)	DBD (N=40)	ADHD (N=54)	DBD+ADHD (N=66)
	Mean (SD)	Mean (SD)	Mean (SD)	Mean (SD)
Age (months)	56.0 (7.3)	53.1 (8.4)	54.7 (8.8)	53.4 (7.1)
% Boys	71	85	76	86
IQ[a]	112.4 (10.4)	101.8 (9.8)***	101.1 (11.3)***	101.2 (12.7)***
CBCL aggression	50.5 (1.3)	75.0 (10.5)***	63.3 (11.0)***	75.6 (11.1)***
CBCL attention	50.8 (2.1)	63.4 (8.1)***	68.6 (6.7)***	69.7 (6.4)***
TRF aggression	52.0 (3.1)	61.5 (9.7)***	62.4 (9.3)***	67.0 (11.1)***
TRF attention	52.0 (3.8)	58.5 (6.6)***	70.8 (10.9)***	69.4 (11.7)***

Note. TD=typically developing; DBD=disruptive behavior disorder; ADHD=attention-deficit/hyperactivity disorder; CBCL=Child Behavior Checklist; TRF=Teacher Report Form;
IQ[a] estimate based on mean Raven and Peabody;
***p <.001. (compared with the TD Group).

Matrices (Raven, Court, & Raven, 1998) and the Peabody Picture Vocabulary Test-III-NL (Dunn & Dunn, 2005; Schlichting, 2005), were excluded. None of the participants were on medication.

In order to study the reliability and convergent/divergent validity children referred with behavioral problems (N=193) and TD children (N=58) were included. To examine test-retest reliability N=24 children of the original group were re-assessed after eight weeks, mean 58 days, SD 7.6, range 40-72 days, where 20 of these preschoolers were referred for behavior problems and 4 were TD children.

For the purpose of the clinical validity study, each child was given a diagnosis of either DBD, ADHD or comorbid DBD and ADHD. As there is no gold standard by which to diagnose a psychiatric disorder, the procedure of best-estimate diagnosis was used, i.e., an independent child psychiatrist and clinical child psychologist consider the information about the child, discuss all the findings and reach consensus (Lord et al., 2006). Before the consensus meeting each clinician had made a decision on the diagnosis of the particular child (DBD, ADHD or both). Only when there was no agreement on the diagnosis a discussion was needed. In 75% of the children there was agreement on the diagnosis, in 15% there was agreement on one of the two diagnoses, and in 10% there was agreement on neither of both diagnoses.

Both clinicians were unfamiliar with the child. The information about the child consisted of 1. Interview with the parents about the reasons for referral and the development of the child; 2. The symptoms coded on the basis of the Kiddie-Disruptive Behavior Disorder Schedule (K-DBDS) (Keenan et al., 2007), a semi-structured DSM-IV based parent interview for the assessment of ODD, CD and ADHD in preschool children; 3. The scores of the Attention Problem scale and the Aggressive Behavior Scale of the Child Behavior Checklist completed by parents (CBCL/1.5-5) and the Child Teacher Report Form completed by teachers or daycare caregivers (C-TRF/1.5-5); 4. The scores on the Children's Global Assessment Scale (CGAS) (Schaffer et al., 1983): i.e., a measure of the impairment of the functioning of the child's behavior, filled out by the parents as well as the teacher/caregiver; 5. The scores on the Raven Coloured Progressive Matrices (Raven et al., 1998) and the Peabody Picture Vocabulary Test III-NL (Dunn & Dunn 2005; Schlichting, 2005); 6. The results of the executive function (EF) tasks; Three inhibition tasks (GoNoGo task, Snack Delay and Shape School), and two working memory tasks (Delayed Alternation and Nine boxes) (Schoemaker et al., 2012); 7. The videotaped observations of the child's behavior during the psychological tasks**.**

In order to reach a clinical consensus diagnosis the DSM-IV-TR criteria of one of the DBDs (ODD, CD or DBD-NOS i.e., defined as either three ODD symptoms and at least one CD symptom or at least one ODD symptom and two CD symptoms) and ADHD were strictly applied (American Psychiatric Association 2000). On the basis of the consensus best-estimate diagnosis, children were positioned in the DBD, ADHD or comorbid (DBD and ADHD) group. In this way the patients were divided into four groups: TD (N=58), DBD (N=40), ADHD (N=54) and comorbid (N=66). The characteristics of the four groups are displayed in Table 1. In the clinical validity study there were group differences in estimated IQ (p<0.001), with the TD group significantly outperforming the three clinical groups, who performed similarly to each other. The four groups did not differ in age or gender.

Procedure

Children were assessed during one morning session. First, general intellectual functioning was assessed followed by the executive functions (EF) tasks. After a break the DB-DOS and K-DBDS were administered. The tasks were administered in a quiet and non-distractible room with a one-way mirror. Parents received a small financial compensation for participating, children received two small gifts. Written informed consent from the parents was gathered before participating. The Medical Ethical Review Committee of the University Medical Center Utrecht, approved this study.

Measures

DB-DOS. The DB-DOS is a method for observational assessment of disruptive behavior in preschool children (Wakschlag et al., 2008a, 2008b). It is divided into three interactional contexts: two examiner contexts and one parent context. In the first examiner context, the examiner is ostensibly responsive to child behavior, the so called Examiner Engaged context. Then, within the context of minimal support, the child is observed while working independently, with the examiner being busy doing his or her own work (Examiner Busy context). In addition, the examiner also briefly leaves the room in order to probe potential covert rule breaking behaviors. Procedures are explained to the parent before starting the tasks and parents are provided with simply worded instructions on flip cards. The primary caregiver was involved in the Parent context. The different tasks, lasting approximately 5 minutes, are parallel across the modules, including frustration, internalization of rules, compliance and social play tasks. The DB-DOS coding system for behavior problems consists of problems in the domains of Behavior Regulation (15 items), and Anger Modulation (6 items). Although there is a Competence scale, in this study we concentrate on the behavior problems domains.

Furthermore, extending the DB-DOS, we developed 10 ADHD items organized in terms of problem behaviors related to inattention (4 items), hyperactivity (4 items) and impulsivity (2 items). The ADHD domain consists of these 10 items. The items were developed on those DSM IV-TR criteria for ADHD which could be observed as behaviors and scored during the observation (see Appendix).

DB-DOS codes are ratings of child behavior ranging from 0-3 and comprising two categories: typical (code 0=normative behavior and 1=normative misbehavior) and clinically concerning behavior (code 2=of concern and 3=atypical). Each item is rated separately for each context. DB-DOS uses the quantitative and the qualitative dimensions of disruptive behavior for the ratings as in preschool children it is not only the frequency but also the intensity of such behaviors that distinguishes clinical problems from normative behavior.

DB-DOS administration and coding. The study group of Wakschlag trained our study group for administration and coding of the DB-DOS. Training for administration included review of the DB-DOS manual, live and videotaped observations, and practice administrations. For coding, reliability was established via 80% exact item-level agreement with the coders of the study group of Wakschlag. For the present study two criterion coders trained an independent team of research students for reliable coding. Coders were independent and blind to the child's clinical status. Following Wakschlag, approximately 15% of the videotaped sessions were randomly selected for double coding to monitor ongoing interrater reliability (Wakschlag et al., 2008a).

K-DBDS. Parents were administrated the Kiddie-Disruptive Behavior Disorder Schedule (Keenan et al., 2007), a semi-structured clinical interview assessing DSM-IV (American

Psychiatric Association 2000). DBD and ADHD symptoms in preschool children. For the scoring we used the developmentally modified approach; it is designed to adhere as closely as possible to DSM-IV. In the present study we found the following Cronbach alpha's (Cronbach, 1951): ODD: 0.75, DBD 0.79, CD: 0.64, ADHD: 0.83.

CBCL/C-TRF. Disruptive behavior symptoms and attention problems were assessed using the scores of the Attention Problem scale and the Aggressive Behavior Scale of the Child Behavior Checklist (Achenbach & Rescorla, 2000) (CBCL/1.5-5) completed by parents and the Child Teacher Report Form (Achenbach & Rescorla, 2000) (C-TRF/1.5-5) completed by teachers or daycare caregivers;

CGAS. To provide a global assessment of functioning, parents and teachers/caregivers completed the non-clinician version of the Children's Global Assessment Scale (CGAS) (Schaffer et al., 1983). The CGAS is scored from 0-100 with lower scores indicating greater impairment.

IFS. The impact on family scale (Sheeber & Johnson, 1992) completed by the mother assesses the social, financial and personal burden resulting from child behavior problems. In the present study Cronbach alpha was 0.95.

SSRS. Parents reports on the summary score of the Social Skills Rating Scale (Gresham & Elliot, 1990) were used to evaluate children's social skills. This scale was completed during the parent module of the DB-DOS. In the present study Cronbach alpha was 0.91.

ECI. Parents and teachers/caregivers reported on child behavior symptoms with the Early Child Inventory (Gadow & Sprafkin, 1996), a DSM based checklist. The pervasive developmental scale of the ECI was used in view of examining divergent validity of DBD and ADHD scores. In the present study Cronbach alpha was 0.77 (parent), 0.82 (teacher).

Statistical analyses

For all scales used the percentage of missing data was at most 4.1%, with the exception of the ECI teacher questionnaire (4.8%) and the CGAS teacher questionnaire (6.8%). For the *reliability analyses* the internal consistency was examined with Cronbach alpha (Cronbach, 1951). Interrater agreement and test-retest reliability at the domain level and by context were assessed using intraclass correlation coefficients. We also examined mean level differences in domain scores and by context at test and retest. *Confirmatory factor analyses* were used to test the fit of the multidomain model of disruptive behavior as proposed by Wakschlag et al., (Wakschlag et al. 2008a), and an extension of this model including the newly proposed ADHD domain. *Convergent and divergent validity* were examined by computing the correlations of DB-DOS domain scores and multi-method assessment of child functioning as evaluated by parent (K-DBDS, CGAS, CBCL/1.5-5, SSRS, ECI) and teacher/caregivers reports (CGAS, C-TRF/1.5-5, ECI).

To assess *clinical validity,* examined on a diagnostic group level, analyses of (co)variance (AN(C)OVA) were conducted to compare the DBD and ADHD groups (with and without comorbid children) and the comorbid group separately with the TD group, while controlling for age and gender for the DBD total group. Group membership was based on consensus best-estimate diagnosis. The analyses were carried out with and without controlling for IQ, using Bonferroni corrections. Statistical significance was based on two-sided tests. Patients groups were not compared with each other as this was not an aim of the study.

To use the DB-DOS as a clinical assessment tool on an individual level, an appropriate cutoff point of the DB-DOS score is required to divide clinical from nonclinical cases for DBD and/or ADHD. For this, Receiver Operating Characteristic (ROC) curve analysis was used. The optimal cutoffs were determined by finding the values that led to a balance between sensitivity (percentage true positives) and specificity (percentage true negatives). All cutoffs, of course, result in some degree of misclassification. For clinical use a high sensitivity and an acceptable specificity (>70%) is required (Kim & Lord, 2011). On the other hand, diagnosing a child with a psychiatric disorder while there is no diagnosis clearly is inappropriate; from this perspective a high specificity is essential. We note that in clinical practice diagnosis will not be based exclusively on the basis of the direct observation of the child's symptoms in the outpatient clinic. Information on the child's every day's behavior from parents and teachers/caregivers should also be taken into account.

To examine how the identification of the presence or absence of DBD and/or ADHD, according to the DB-DOS, may be used in supporting or rejecting a possible diagnosis of DBD and/or ADHD, based on the information from parents, teachers (or other caregivers) and cognitive assessment, using best-estimate diagnosis (see participants), we computed the percentage agreement between the best-estimate diagnosis DBD and/or ADHD (or no diagnosis) and the DB-DOS diagnosis DBD and/or ADHD (or no diagnosis). For these analyses we used data from the total sample including children who are not diagnosed and checked if a child diagnosed only with ADHD (DBD), based on best-estimate diagnosis, will get an ADHD (DBD) diagnosis based on the DB-DOS, but will not receive a DBD (ADHD) diagnosis on the DB-DOS. We assumed the typically developing children have no diagnosis based on their score beneath the 90-th percentile on the Attention Problems scale and on the Aggressive Behavior scale of the CBCL/1.5-5 and C-TRF/1.5-5. We used the BR/ DBD total score cutoff points for the DBD diagnosis and the ADHD total score cutoff point for the ADHD diagnosis.

RESULTS

Domain Level Reliability

Diagnostic observation measures may include clinically salient items whose occurrence is too low to be evaluated statistically (Lord et al., 1989). Five items (all in the Behavioral Regulation domain) were rated as present fewer than 10 times in the clinical concerning range; verbal aggression (two items), directed aggression, spiteful behavior and sneaky behavior. Because estimates of reliability can be unduly biased by differences between a single pair of raters for items with a very low occurrence, these items were excluded from domain level analyses.

Internal consistency. Overall the three domains (Behavioral Regulation, Anger Modulation and ADHD) exhibited good internal consistency (see Table 2), both in terms of total domain scores (Cronbach alpha mean .82, range .77-.91) and by context (Cronbach alpha mean .82, range .69-.94).

Table 2 *Domain Score Reliability*

Domain	Cronbach alpha	Interrater ICC	Test-retest ICC	Test Mean (SD)	Retest Mean (SD)
	N=251	EA=39/EB=47 PC=28	N=24		
Behavioral Regulation	Mean .79	Mean .95	Mean .63		
Examiner Engaged	.77	.95	.52	3.8 (2.6)	3.0 (2.7)
Examiner Busy	.75	.96	.62	5.6 (4.5)	3.1 (3.4)
Parent	.84	.92	.74	11.1 (5.6)	6.5 (4.1)
Anger Modulation	Mean .91	Mean .88	Mean .59		
Examiner Engaged	.86	.86	.53	1.6 (1.8)	2.2 (2.3)
Examiner Busy	.92	.91	.68	3.8 (3.3)	2.2 (2.3)
Parent	.94	.88	.57	7.7 (5.2)	3.5 (4.0)
ADHD	Mean .77	Mean .92	Mean .71		
Examiner Engaged	.82	.94	.68	11.0 (4.0)	11.5 (4.2)
Examiner Busy	.81	.94	.80	12.3 (4.8)	10.3 (4.0)
Parent	.69	.87	.64	14.8 (6.0)	12.7 (5.1)

Note. ICC=Intraclass Correlation Coefficient, SD=Standard Deviation, ADHD=attention-deficit/hyperactivity disorder.

Interrater reliability. It showed excellent interrater reliability as Table 2 illustrates, both in terms of total domain scores (intraclass correlation coefficients mean .92, range .88-.95) and by context (intraclass correlation coefficients mean .92, range .86-.97)

Test-retest reliability of domain scores*.* Test-retest analyses indicated moderate reliability across domains and contexts both (see Table 2) in terms of total domain scores (intraclass correlation coefficients mean .64, range .59-.71) and by context (intraclass correlation coefficientsmean .64, range .52-.80). Across the contexts, domain scores tended to decrease slightly in the second testing.

Domain intercorrelations across contexts. Pearson correlations among the DB-DOS domains across context indicated modest to substantial consistency in child behavior across the different DB-DOS contexts. Cross-context associations within a domain were significantly higher across the two examiner contexts (Pearson correlation ranged from .44-.58) relative to associations between each of the examiner contexts and the Parent context (Pearson correlation ranged from .18-.40).

Confirmatory factor analyses. They were conducted to test the fit of two models.

Model 1 was the multidomain model of disruptive behavior as proposed by Wakschlag et al. (2008a). Model 2 was an extended model that includes the ADHD domain. In both models, three of the observed domain scores were regressed on the latent Anger Modulation construct, three other observed domain scores were regressed on the latent Behavioral Regulation construct, and the latent Anger Modulation and Behavioral Regulation constructs were regressed on the latent Disruptive Behavior construct. Following Wakschlag et al. (2008a), the factor loadings of Anger Modulation and Behavioral Regulation onto Disruptive Behavior were constrained to be equal. The errors of observed domain scores from the same context were allowed to correlate. Model 2 extends model 1; here three new observed domain scores were regressed on the latent Inattention construct, three other new observed domain scores were regressed on the latent Hyperactivity/Impulsivity construct, and the latent Inattention and Hyperactivity/Impulsivity constructs were regressed on the latent ADHD construct. The errors of these domain scores were again allowed to correlate with the errors of other observed domain scores from the same context. The latent ADHD construct was allowed to correlate with the latent Disruptive Behavior construct. The parameter estimation results of the two models are shown in Figure 1.

As in the study of Wakschlag et al. (2008a), the fit of model 1 was excellent ($\chi 2=8.1$, df=5, p=.151, CFI=.992, RMSEA=.051). The fit of model 2 could be considered satisfactory ($\chi 2=74.41$, df=33, p=.000, CFI=.946, RMSEA=.073).

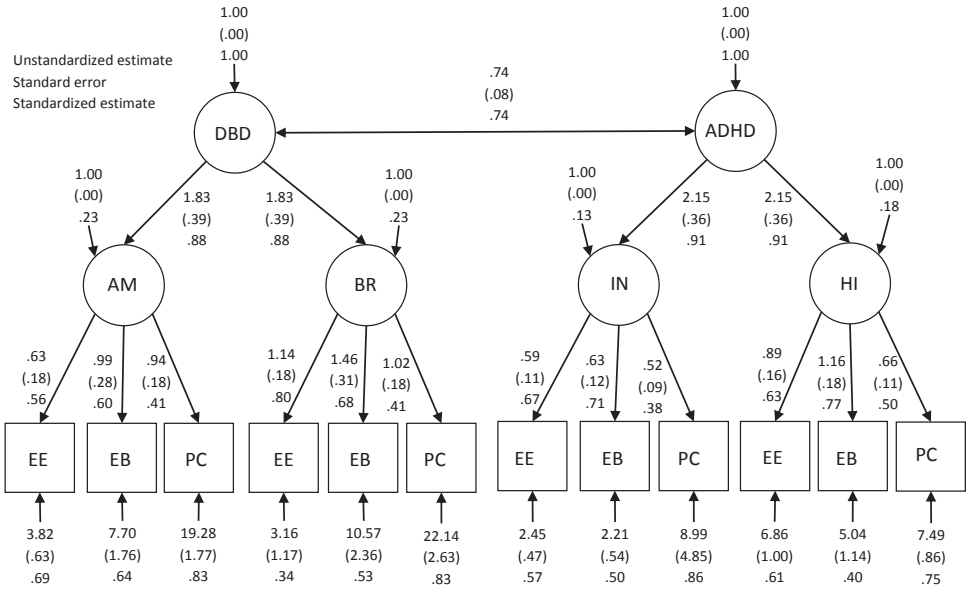


Figure 1. Disruptive Behavior Diagnostic Observation Schedule (DB-DOS) factor structure.
DBD=disruptive behavior disorder; ADHD=attention-deficit/hyperactivity disorder; AM=Anger Modulation; BR=Behavior Regulation; EA=Examiner Active; EB=Examiner Busy; PC=Parent Context; IN=Inattention; HI=Hyperactivity-Impulsivity.

Validity

Convergent validity. As Table 3 illustrates significant correlations were established between DB-DOS domain scores and parent and teacher/caregiver questionnaires of symptoms and impairment (K-DBDS, CGAS, CBCL/1.5-5,C-TRF/1.5-5, IFS). Correlations between the Behavioral Regulation domain and the parent and teacher/caregiver questionnaires, with the exception of the CBCL/1.5-5 Aggressive Behavior scale and Behavioral Regulation domain in the Examiner Engaged context, were significant and as expected. Correlations between the Anger Modulation domain and the questionnaires were slightly lower but significant as well. Correlations between the ADHD scales and the questionnaires were significant and the absolute values were higher than the correlations between the Behavioral Regulation/Anger Modulation domain and the parent and teacher/caregiver questionnaires.

Divergent validity. Small but significant minor correlations were found between the majority of DB-DOS scores and parent and teacher/caregiver questionnaires (SRSS, ECI-PDD) (Table 3).

Clinical validity. To investigate clinical validity on a group level, analyses of (co)variance were conducted to examine group differences in DB-DOS scores between the TD group and the clinical groups without controlling for IQ (See Table 4). The results were similar

Table 3 *Association (Pearson r Correlation) of the Disruptive Behavior Diagnostic Observation Schedule (DB-DOS); Scores by Context and Multi-Informant Assessments of Child Disruptive Behavior and Functioning*

	Behavioral Regulation				Anger Modulation				ADHD			
	EE	EB	PC	Total	EE	EB	PC	Total	EE	EB	PC	Total
Parent reported												
K-DBDS	.19**	.20**	.27**	.28**	.24**	.15*	.17**	.24**	.30**	.32**	.30**	.39**
CGAS	-.28**	-.30**	-.35**	-.41**	-.21**	-.14*	-.21**	-.25**	-.30**	-.35**	-.42**	-.46**
CBCL	.09	.17**	.23**	.22**	.14*	.14*	.15*	.18**	.29**	.34**	.41**	.44**
IFS (M)	.22**	.26**	.25**	.32**	.19**	.14*	.14*	.20**	.28**	.33**	.34**	.31**
SSRS	-.11	-.17**	-.37**	-.30**	-.07	-.11	-.24**	-.21**	-.12	-.16*	-.32**	-.27**
ECI (PDD)	.10	.09	.12	.13*	.18**	.15*	.06	.15*	.13	.08	.10	.11*
Teacher/Caregiver reported												
CGAS	-.40**	-.38**	-.25**	-.45**	-.23**	-.24**	-.17**	-.29**	-.34**	-.34**	-.28**	-.41**
TRF	.34**	.41**	.26**	.44**	.25**	.29**	.24**	.36**	.30**	.37**	.32**	.42**
ECI (PDD)	.26**	.17**	.22**	.28**	.18**	.17**	.24**	.29**	.19**	.18**	.16**	.23**

Note. ADHD=attention-deficit/hyperactivity disorder; EE=Examiner Engaged context;EB=Examiner Busy context; PC=Parent context; K-DBDS=Kiddie-Disruptive Behavior Disorder Schedule; CGAS=Children's Global Assessment Scale; CBCL=Child Behavior Checklist; IFS=Impact on Family Scale; SSRS=Social Skills Rating Scale; IFS (M)=Impact on Family Scale (Mother); ECI (PDD)=Early Child Inventory (Pervasive Developmental Disorder); *p <.05. **p <.01.

Table 4 *Means (and SD) of the DB-DOS Domains in the Six Groups*

	TD (N=58)	DBD (N=40)	ADHD (N=54)	DBD+ ADHD (N=66)	DBD only and DBD+ ADHD (N=106)	ADHD only and DBD+ ADHD (N=120)	
	Mean(SD)	Mean (SD)	Mean (SD)	Mean (SD)	Mean (SD)	Mean (SD)	
BR EE	2.5 (2.1)	3.4 (2.4)	4.7 (2.9)	5.4 (3.4)	4.7 (3.2)	5.1 (3.2)	(bcde)***
BR EB	2.4 (2.5)	5.3 (3.8)	6.4 (4.5)	6.9 (5.2)	6.3 (4.8)	6.7 (4.9)	(abcde)***
BR PC	5.4 (3.9)	9.9 (4.2)	9.2 (4.1)	10.8 (4.5)	10.5 (4.4)	10.1 (4.4)	(abcde)***
BR tot	10.4 (6.0)	18.7 (8.1)	20.3 (8.7)	23.1 (9.8)	21.4 (9.4)	21.8 (9.4)	(abcde)***
AM EE	1.9 (1.7)	3.1 (3.0)	2.2 (1.9)	3.1 (3.1)	3.1 (3.0)	2.7 (2.7)	(ae)* (cd)**
AM EB	2.7 (3.4)	4.2 (4.2)	3.6 (2.9)	4.4 (4.3)	4.3 (4.2)	4.0 (3.7)	(acde)*
AM PC	4.4 (4.4)	6.9 (5.0)	6.3 (4.7)	7.4 (4.8)	7.2 (4.9)	7.0 (4.8)	(ab)*(de)**(c)***
AM tot	9.0 (6.1)	14.1 (9.6)	11.9 (5.7)	15.0 (9.3)	14.6 (9.4)	13.8 (8.0)	(b)*(a)**(cde)***
ADHD EE	8.8 (3.1)	10.9 (4.7)	14.1 (4.2)	13.7 (5.0)	12.6 (5.0)	13.9 (4.6)	(a)*(bcde)***
ADHD EB	9.4 (3.8)	11.4 (3.8)	15.4 (4.1)	14.5 (5.0)	13.3 (4.8)	14.9 (4.6)	(a)*(bcde)***
ADHD PC	9.4 (4.2)	13.7 (4.3)	15.7 (4.5)	16.9 (5.6)	15.7 (5.4)	16.4 (5.2)	(abcde)***
ADHD tot	27.6 (7.2)	35.7 (8.8)	45.3 (9.4)	42.2 (11.1)	41.6 (11.2)	45.3 (10.4)	(abcde)***

Note. TD=typically developing; DBD=disruptive behavior disorder; ADHD=attention-deficit/hyperactivity disorder; DBD total=DBD only and DBD+ADHD; ADHD Total=ADHD only and ADHD+DBD; BR=Behavior Regulation; AM=Anger Modulation; EE=Examiner Engaged context; EB=Examiner Busy context; PC=Parent context; tot=total; a=DBD vs TD, b=ADHD vs TD, c=DBD+ADHD vs TD, d=DBD totaal vs TD, e=ADHD totaal vs TD; e.g. (ae)* means p-value for a and e lower than .05. *p<.05. **p <.01. ***p <.001.

for the analyses when controlling for IQ. All clinical groups differed from the TD group on all DB-DOS scores, with the exception of the DBD group with respect to the Behavior Regulation Examiner Engaged context and the ADHD group with respect to the Anger Modulation Examiner Engaged/Examiner Busy context.

To examine clinical validity on an individual level, an appropriate cutoff point of the DB-DOS score is required. We used the total DBD group (DBD and DBD+ADHD both consensus best estimate diagnosis), and the total ADHD group, (ADHD and DBD+ADHD consensus best estimate diagnosis), with the TD group for the ROC analysis. Overall, the estimated area under curve (AUC) values and their 95% confidence intervals (C.I.) were satisfactory. First we determined a cutoff point of the total (Examiner Engaged, Examiner Busy and the Parent context together) DBD score (Behavior Regulation and Anger Modulation). The ROC analysis yielded an estimated AUC of .81 (95% C.I.=[.74-.88]). Using a total DBD score of 24 yielded a sensitivity of 77% and a specificity of 69% (see Table 5). On the basis of this

DBD cutoff score, sensitivity of 77% means that three out of four children, who according to clinical consensus were diagnosed with DBD, will receive a DBD diagnosis. Specificity of 69% means that 69% of the TD children will not be diagnosed with DBD, according to consensus best estimate diagnosis, and 31% will be diagnosed with DBD while the diagnosis is not given. Second, as this low sensitivity/specificity was unsatisfactory, we explored whether a better cutoff point of the total (Examiner Engaged, Examiner Busy and the Parent context together) Behavior Regulation score could be determined (i.e., part of the total DBD score). The ROC analysis yielded an AUC of .85 (95% C.I.=[.78-.91]). Using a total Behavior Regulation score of 13 yielded a sensitivity of 83% and a specificity of 72%. This cutoff point for Behavior Regulation was slightly better than the one for total DBD score.
We have explored the role of different contexts (Examiner Engaged, Examiner Busy and the Parent context) for the total DBD, the Behavior Regulation and the Anger Modulation scores, but the results of the sensitivity/specificity balance for the different contexts were unsatisfactory. We explored also whether the total (Examiner Engaged, Examiner Busy and the Parent context together) Anger Modulation score would be better; a total Anger Modulation score of 11 yielded a sensitivity of 62% and a specificity of 64%, using a total Anger Modulation score of 13 yielded a sensitivity of 55% and a specificity of 74%. The ROC analysis yielded an AUC of 0.69 (95% C.I.=[.60-.77]). These cutoff points for Anger Modulation were unsatisfactory. Fourth we examined whether the Behavior Regulation cutoff point for younger and older children differed in specificity and sensitivity. Using a total Behavioral Regulation score for the 3.5-4.5 year old children of 14 yielded a sensitivity of 87% and a specificity of 76% (see Table 5), while for 4.5-5.5 years sensitivity was 67% and specificity 76%. The ROC analysis yielded an AUC of .89 (95% C.I.=[.83-.95]).
Fifth we examined whether the cutoff point for referred patients with clinical scores on the TRF 1.5-5 (T score above 70) differed in specificity and sensitivity for the different modules. The referred children with DBD with clinical scores on the TRF 1.5-5 (T score above 70) showed a better sensitivity and specificity balance compared to the total DBD group. Categorization of the total DBD group in sex, IQ (IQ above or below 104) CGAS parents or teachers/caregivers (above or below 60), and CBCL (above or below 70) did not show better results compared to the total DBD group. For the ADHD domain a cutoff point of 34 was suitable (see Table 5); using a total (Examiner Engaged, Examiner Busy and the Parent context together) ADHD score of 34 yielded a sensitivity of 87% and a specificity of 79% (see Table 5). The ROC analysis yielded an AUC of .92 (95% C.I.=[.88-.96]). Thus, on the basis of the ADHD score, sensitivity of 87% means almost nine out of ten patients who according to clinical consensus were diagnosed with ADHD, will receive an ADHD diagnosis. Specificity of 79% means that 79% of the TD children will not be diagnosed with ADHD, according to consensus best-estimate diagnosis, and 21% will be diagnosed with ADHD while the diagnosis is not given. Finally, the percentage agreement was computed between the best-estimate diagnosis

Table 5 *ROC Curve*

DBD total (N=106)	DB-DOS score	Sensitivity	Specificity	Positive Predictive Value	Negative Predictive Value
Total DBD score	20	85	59	67	80
	24	**77**	**69**	**71**	**75**
	28	67	79	76	71
	31	56	83	77	65
Total BR score	12	87	60	69	82
	13	**83**	**72**	**75**	**81**
	16	68	81	78	72
	18	63	86	94	70
Total BR score 3.5 - 4.5 years	13	90	72	76	88
	14	**87**	**76**	**78**	**85**
	15	81	79	79	81
	16	77	81	80	78
ADHD Total (N=120)					
Total ADHD score	32	91	69	76	88
	33	89	72	76	87
	34	**87**	**79**	**81**	**86**
	36	83	83	83	83
	38	79	95	94	82

Note. DBD=disruptive behavior disorder; DBD total=DBD only and DBD+ADHD; ADHD=attention-deficit/hyperactivity disorder; ADHD Total=ADHD only and ADHD+DBD; BR=Behavior Regulation. Bolded numbers indicate maximized sensitivities and specificities depending on criteria used in selecting cutoff scores.

DBD and/or ADHD (or no diagnosis) and the DB-DOS diagnosis DBD and/or ADHD (or no diagnosis). The percentage agreement, in this study group (referred and typically developing children), between the best-estimate DBD diagnosis (or not) and the DB-DOS DBD diagnosis (or not) based on the cutoff point of the BR total score of 13 (see Table 5) is 59% and the percentage agreement based on the cutoff point of the DBD total score of 24 (see Table 5) is 60%. For the best-estimate diagnosis ADHD (or not) and the DB-DOS ADHD diagnosis (or not) based on the cutoff point of 34 (see Table 5), the percentage agreement is 75%.

DISCUSSION

The results of reliability (internal consistency, interrater reliability, test retest reliability) and validity (convergent and divergent) of DBD symptom scores in the current study in a European sample are similar to the findings of the Wakschlag et al. studies (2008a, 2008b) in a US sample. Likewise, the results of reliability and validity of the ADHD symptom scores are satisfactory. Thus, the current study not only replicated but also extended the US study by adding a reliable and valid ADHD domain. This is important, as in the clinical assessment of preschool children with externalizing behavior problems, comorbidity with ADHD should be investigated. The results of confirmatory factor analyses for ADHD domain are satisfactory.

With respect to the clinical validity of the DB-DOS on a group level, all clinical groups differed from the TD group on DB-DOS DBD and ADHD symptom scores, with the exception of the DBD group with respect to the Behavior Regulation Examiner Engaged context, and the ADHD group with respect to the Anger Modulation Examiner Engaged/Examiner Busy context; the exception of the DBD group can be explained by the highly structured context with an unknown examiner, the exception of the ADHD group is in line with expectations. Besides, children with ADHD showed significantly higher DB-DOS symptom scores in all the three Behavior Regulation contexts and the Anger Modulation Parent context compared with the TD group. This can be explained by the high correlation of DBD and ADHD symptoms in preschool children (Sterba et al., 2007).

The main aim of the present study was to examine clinical validity on an individual level using clinical cutoff points for DBD and ADHD symptom scores. While exploring the clinical cutoff point for DBD, the Behavior Regulation total score appeared to have slightly better sensitivity/specificity than the total DBD score. Thus, using the Behavioral Regulation total score, four out of five children, who according to clinical consensus were diagnosed with DBD, will receive a DBD diagnosis. By contrast the Anger Modulation total score was unsatisfactory. Furthermore, sensitivity of the clinical cutoff point for Behavior Regulation appeared to be somewhat better in the younger age group (3.5-4.5 years) than in the older age group (4.5-5.5 years), maybe because older preschoolers can better inhibit their own behavior than younger preschoolers. The better results for referred patients with clinical scores on the TRF 1.5-5 (T score above 70) are in line with evidence that youth with symptoms according to both parent and teacher had more mental health concerns than youth with symptoms according to either parent or teacher (Gadow & Drabick, 2012).

The clinical cutoff point for ADHD symptom scores appeared to be good. Indeed, almost nine out of ten children, who according to clinical consensus were diagnosed with ADHD, will receive an ADHD diagnosis. Based on ROC curve analyses the sensitivity and specificity for the ADHD diagnosis when compared with DBD diagnoses appeared to be better for

ADHD than for DBD. This may be due to the cross-situational character of ADHD symptoms (Gadow & Drabick, 2012). In contrast, ODD symptoms may be present in just one context (at home or at school). Thus, it may be more difficult to elicit DBD symptoms than ADHD symptoms in an observational procedure at an outpatient clinic. Another explanation could be that young children with ADHD are typically unable to control their impulsive behavior in a novel context, whereas a child with ODD or CD may be less likely to display their symptoms in an unfamiliar context.

When investigating the clinical validity of the DB-DOS on an individual level we should keep in mind that in everyday clinical practice the decision of diagnosing a child with ODD, CD or ADHD is not based on the results of a single measure but on the combination of results from multiple measures such as a standardized parent and teacher/caregiver rating scales and a (semi-) structured DSM orientated interview with the parents (Matthys & Lochman, 2010). Our main hypothesis was that the DB-DOS may help support the presumption of a diagnosis generated by the information from parents, teachers (or other caregivers) and cognitive assessment. In this study group the DB-DOS BR/DBD total score supported approximately 60% the DBD diagnosis (or not), generated by the information from parents, teachers (or other caregivers) and cognitive assessment using the best-estimate diagnosis. This means that for approximately six out of ten children, diagnosed with DBD (or not), generated by the information from the parents, teachers (or other caregivers) and cognitive assessment, the DBD diagnosis has been confirmed (or not) by the DB-DOS. The DB-DOS ADHD total score supported nearly 75% of the ADHD diagnosis (or not), generated by the information from parents, teachers (or other caregivers) and cognitive assessment using the best-estimate diagnosis, which means that nearly three out of four children the diagnosis ADHD (or not), generated by the information from the parents and teachers (or other caregivers), has been confirmed (or not) by the DB-DOS. To our knowledge this is the first study that gives a quantitative indication of what may be expected from a standardized direct observation of the child within the context of a broad clinical assessment for preschool children suspected for DBD and ADHD.

The replication of the Wakschlag et al. study (2008a, 2008b) in a European sample of referred children, the newly developed ADHD domain, the determination of cutoff points to distinguish clinical from non-clinical cases and the determination of the importance of the DB-DOS to support (or not) a presumption of a DBD and/or ADHD or no diagnosis, generated by the information from parent, teachers (or other caregivers) and cognitive assessment, are strengths of the present study. In the present study the DB-DOS is examined in the context of DSM-IV-TR-based diagnoses while the use of DBD diagnoses in young children has been questioned; as an alternative, based on a dimensional conceptualization of psychopathology, a multidimensional developmentally-based approach to preschool disruptive behavior has been developed (Wakschlag et al., 2012).

The study also has limitations. First, the test-retest reliability was only moderate for DBD, in line with the Wakschlag study (2008a), and for ADHD. In our study we organized the retest after 8 weeks. Familiarity with the tasks and parents promising presents for their child's good behavior may explain the lower score on the retest. However, a longer period of time for the retest, for example six months, is not an appropriate alternative since preschool children show fast developmental changes. Second, in the course of studying divergent clinical validity, in line with the Wakschlag study (2008b) small but significant minor correlations were found between the majority of DB-DOS scores and parent reports of social skills on the SSRS; one may question whether the SSRS is an appropriate measure to examine divergent clinical validity as social skills are less well developed in children with DBD and ADHD (Ronk, Hund, & Landau, 2011). Therefore we added the ECI-PDD part in view of studying divergent validity. As expected, the correlation between the ECI-PDD part and the Behavioral Regulation/Anger Modulation/ADHD symptom scores was weak. Children with an IQ below 70 were excluded; results of the study thus are not generalizable to children with IQs below 70. Third, the very small number of girls constrains the ability to discuss sex differences. Another limitation of this study is that in this dataset no subscales of ADHD (i.e., Inattention and Hyperactivity/Impulsivity) are used in the clinical validity study because of the low (N=6) ADHD inattention subtype diagnosis, based on consensus best-estimate diagnosis.

Implications for practice and research

First, the DB-DOS may be used as an observational tool to help support (or not) a presumption of a DBD and/or ADHD, generated by the information from parent, teachers (or other caregivers) and cognitive assessment. The complexity of diagnosing DBD and ADHD in preschool children necessitates a combination of information about the child's behavior not only from a parent interview/questionnaire and a teacher/caregiver questionnaire but also from an observation of the child (Le Couteur & Gardner, 2008, Wakschlag et al., 2008a, 2008b). Second, accurate identification of ADHD in preschoolers is crucial as treatment of these children may include pharmacotherapy and concerns have been raised about overuse of psychostimulants to decrease symptoms of ADHD in the absence of a clear diagnosis (Rey & Sawyer, 2003). Third, further research is required to study the predictive validity of the DB-DOS. Fourth, it should be considered whether a more practical and less time-consuming way of coding may be developed in view of using the DB-DOS in clinical practice. Fifth, the incremental knowledge gained from the DBDOS beyond parent interviews and parent/teacher questionnaires is a future research direction as well.

APPENDIX

DB-DOS ADHD items

Inattention

1. Careless/inattention to details
2. Difficulty sustaining attention
3. Easily distracted
4. Easily bored

Hyperactivity

5. Fidgets, squirms (minor movements)
6. Difficulty remaining seated and vigorous motor activity (major movements/hyperactivity)
7. Difficulty playing quietly
8. Talks excessively

Impulsivity

9. Difficulty waiting (non verbal)
10. Interrupts; difficulty waiting (verbal)

Acknowledgments

Disclosure: Ms. Bunte, Laschen, Schoemaker, and Dr. Hessen, van der Heijden, and Matthys report no biomedical financial interests or potential conflicts of interest.

The authors are grateful to the parents and children who participated in this study. We especially thank Lauren Wakschlag, at the Department of Medical Social Sciences, Northwestern University, Feinberg School of Medicine, Chicago, United States, for her cooperation in providing assessment procedures used in this study and in training the Utrecht research group in administering and scoring the DB-DOS. We also thank Eva van der Kleij, Justa Kamstra and Susanne van Reedt Dortland, at the Department of Psychiatry, University Medical Center Utrecht, Utrecht, the Netherlands, for their assistance with the data collection.

REFERENCES

Achenbach, T.M., & Rescorla, L.A. (2000). *Manual for the ASEBA preschool forms and profiles.* Burlington: Research Center for Children, Youth & Families VT: University of Vermont.

American Psychiatric Association. (2000). *Diagnostic and statistical manual of mental disorders (4th Edition, text-revised)*. Washington, DC: Author.

Collishaw, S., Goodman, R., Ford, S., Rabe-Hesketh, S., & Pickels, A. (2009). How far are associations between child, family and community factors and child psychopathology informant-specific and informant-general? *Journal of Child Psychology and Psychiatry, 50,* 571-580.

Cronbach, L. (1951). Coefficient alpha and the internal structure of tests. *Psychometrika*, *16*, 297-334.

Dunn, L.M., & Dunn, L.M. (2005). *Peabody picture vocabulary test (III-NL)* [Dutch version by L. Schlichting]. Amsterdam: Harcourt Assessment.

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.

Egger, H.L., Erkanli, A., Keeler, G., Potts, E., Walter, B.K., & Angold, A. (2006). Test -retest reliability of the Preschool Age Psychiatric Assessment (PAPA). *Journal of the American Academy of Child and Adolescence Psychiatry, 45*, 538-549.

Gadow, K., & Drabick, D. (2012). Anger and irritability symptoms among youth with ODD: Cross-informant versus source-exclusive syndromes. *Journal of Abnormal Child Psychology, 40,* 1073-1085.

Gadow, K., & Sprafkin, J. (1996). *Early childhood inventories manual.* Stony Brook, NY: Checkmate Plus.

Gresham, F., & Elliot, S. (1990). *Social Skills Rating System*. Circle Pines, MN: American Guidance Service.

Howell, E.M., & Teich, J. (2008). Variations in medicaid mental health service use and cost for children. *Administration and Policy in Mental Health, 35*(3), 220-8.

Keenan, K., Boeldt, D., Chen, D., Coyne, C., Donald, R., Duax J., …Humphries, M. (2011). Predictive validity of DSM-IV oppositional defiant and conduct disorders in clinically referred preschoolers. *Journal of Child Psychology and Psychiatry, 52*, 47-55.

Keenan, K., Wakschlag, L.S., Danis, B., Hill, C., Humphries, M., Duax, J., & Donald, R. (2007). Further evidence of the reliability and validity of DSM-IV ODD and CD in preschool children. *Journal of the American Academy of Child and Adolescent Psychiatry, 26*, 457-468.

Kim, S.H., & Lord, C. (2011). New Autism Diagnostic Interview-revised algorithms for toddlers and young pre-schoolers from 12-47 months of age. *Journal of Autism and Developmental Disorders*, Sep1.doi:10.1111/j.1469-7610.2011.02458.x.

Lahey, B., Pelham, W.E., Loney, D., Kipp, H., Ehrhardt, A., Lee, S, …Massetti, G. (2004). Three-Year predictive validity of DSM-IV attention deficit hyperactivity disorder in children diagnosed at 4-6 years of age. *American Journal of Psychiatry, 161*, 2014-20.

Le Couteur, A., & Gardner, F. (2008). Use of structured interviews and observational methods in clinical settings. In M. Rutter, D. Bishop, D. Pine, S. Scott, J. Stevenson, E. Taylor & A. Thapar (Eds.), Rutter's Child and Adolescent Psychiatry 5th Ed. (pp. 271-288). Oxford: Blackwell.

Lord, C., Risi, S., DiLavore, S., Shulman, C., Thurm, A., & Pickles, A. (2006). Autism from 2 to 9 years of age. *Archives of General Psychiatry, 63*, 694-701.

Lord, C., Rutter, M., Goode, S., Heemsbergen, J., Jordan, H., Mawhood, L., & Schopler, E. (1989). Autism diagnostic observation schedule: a standardized observation of communicative and social behavior. *Journal of Autism and Developmental Disorders*, *19*, 185-212.

Matthys, W., & Lochman, J.E. (2010). *Oppositional defiant disorder and conduct disorder in childhood*. Chichester: Wiley-Blackwell.

Posthumus J.A., Raaijmakers, M.A., Maassen, G.H., van Engeland, H., & Matthys, W. (2012). Sustained effects of incredible years as a preventive intervention in preschool children with conduct problems. *Journal of Abnormal Child Psychology, 40,* 487-500.

Raven, J.C., Court, J.H., & Raven, J. (1998). *Raven coloured progressive matrices*. Oxford: Oxford Psychologist Press.

Rey, J.M., & Sawyer, M.G. (2003). Are psychostimulant drugs being used appropriately to treat child and adolescent disorders? *The British Journal of Psychiatry, 182*, 284-286.

Robinson, E.A., & Eyberg, S.M. (1981). The dyadic parent-child interaction coding system: standardization and validation. *Journal Consult Clinical Psychology, 49,* 245-50.

Ronk, M.J., Hund, A.M., & Landau, S. (2011). Assessment of social competence of boys with attention deficit/hyperactivity disorder: problematic peer entry, host responses, and evaluations. *Journal of abnormal Child Psychology*, *39*, 829 -40.

Schaffer, D., Goud, M.S., Brasic, J., Ambrosini, P., Fisher, P., Bird, H., & Aluwahlia, S. (1983). A children's global assessment scale (CGAS). *Archives of General Psychiatry*, *40*, 1228-1231.

Schlichting, L. (2005). Peabody picture vocabulary test III-NL. Amsterdam: Hartcourt Assessment.

Schoemaker, K., Bunte, T., Wiebe, S.A., Espy, K.A., Dekovic, M., & Matthys, W. (2012). Executive function deficits in preschool children with ADHD and DBD. *Journal of Child Psychology and Psychiatry*, *53(2)*, 111-119.

Sheeber, L., & Johnson, J. (1992). Applicability of the Impact on Family Scale for assessing families with behaviorally difficult children. *Journal of Child Psychology and Psychiatry, 17*, 155-159.

Sterba, S., Egger, H.L., & Angold, A. (2007). Diagnostic specificity and nonspecificity in the dimensions of preschool psychopathology. *Journal of Child Psychology and Psychiatry, 48*, 1005-1013.

Wakschlag, L. S., Hill, C., Carter, A. S., Danis, B., Egger, H. L., Keenan, K., …Briggs -Gowan, M.J. (2008a). Observational Assessment of Preschool Disruptive Behavior, Part I: Reliability of the Disruptive Behavior Diagnostic Observation Schedule (DB-DOS). *Journal of the American Academy of Child and Adolescent Psychiatry, 47*, 622-631.

Wakschlag, L. S., Briggs-Gowan, M. J., Hill, C., Danis, B., Leventhal, B.D., Keenan, K., …Carter, A.S. (2008b). Observational Assessment of Preschool Disruptive Behavior, Part II: Validity of the Disruptive Behavior Diagnostic Observation Schedule (DB -DOS). *Journal of the American Academy of Child and Adolescent Psychiatry, 47*, 632-640.

Wakschlag, L.S., Henry, D.B., Tolan, P.H., Carter, A.S., Burns J.L., & Briggs-Gowan, M.J. (2012). Putting theory to the test: modeling a multidimensional, developmentally-based approach to preschool disruptive behavior. *Journal of the American Academy of Child and Adolescent Psychiatry, 51,* 593-604.

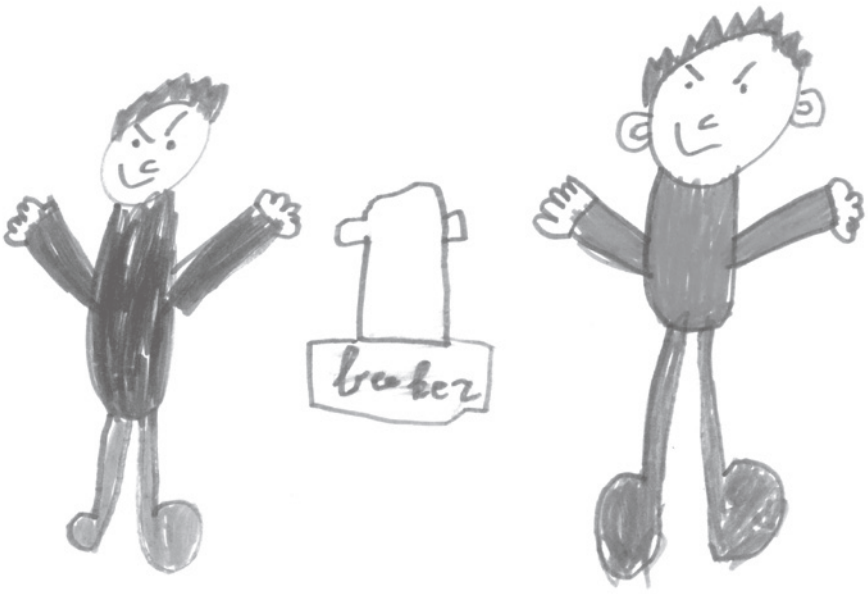
Wakschlag, L.S., Leventhal, B.L., Briggs-Gowan, M.J., Danis, B., Keenan, K., Hill, C., …Carter, A.S. (2005). Defining the "disruptive" in preschool behavior: what diagnostic observation can teach us. *Clinical Child and Family Psychology Review, 8*, 183-201.

Wilens, T.E., Biederman, J., Brown, S., Monuteaux, M., Prince, J., & Spencer, T.J. (2002). Patterns of psychopathology and dysfunction in clinically referred preschoolers. *Journal of Developmental & Behavioral Pediatrics, 23*, S31-6.

CHAPTER 3

Clinical usefulness of the Kiddie-Disruptive Behavior Disorder Schedule in the diagnosis of DBD and ADHD in preschool children

Tessa L. Bunte, M.D., Kim Schoemaker, M.S., David J. Hessen, Ph.D., Peter G.M. van der Heijden, Ph.D., Walter Matthys, M.D., Ph.D.



Bunte, T.L., Schoemaker, K., Hessen, D.J., van der Heijden, P.G.M., Matthys, W. (2013). Clinical usefulness of the Kiddie-Disruptive Behavior Disorder Schedule in the diagnosis of DBD and ADHD in preschool children. *Journal of Abnormal Child Psychology, 41*, 681-690.

ABSTRACT

The aim of the present study was to investigate the clinical usefulness of a semi-structured diagnostic parent interview, i.e., the Kiddie-Disruptive Behavior Disorder Schedule (K-DBDS), in preschool children. For Oppositional Defiant Disorder (ODD), to define symptoms two coding methods were compared, i.e., one based on the threshold "often" and the other based on the frequency of behaviors in combination with the presence of clinical concern. For Attention-Deficit/Hyperactivity Disorder (ADHD), to define symptoms, two coding methods were compared, i.e., one with and one without consideration of pervasiveness across contexts. Participants were referred preschool children with externalizing behavioral problems (N=193; 83% male) and typically developing (TD) children (N=58; 71% male). The referred children were given a diagnosis of either ODD/CD (N=39), or ADHD (N=58) or comorbid ODD/CD+ADHD (N=57) or no diagnosis (N=39) based on best-estimate diagnosis. Receiver Operating Characteristic curve analyses showed that a cutoff score of four ODD symptoms using "often" as the threshold for frequency of behaviors led to a sensitivity of 87% and a specificity of 93%; the coding method which included the frequency of behaviors yielded a sensitivity of 56% and a specificity of 100%. For ADHD, a clinical cutoff score of five symptoms without the pervasiveness criterion yielded a sensitivity of 83% and a specificity of 98%; when the pervasiveness criterion was included sensitivity was 77% and specificity 98%. In the clinical assessment of ODD and ADHD in preschool children, the K-DBDS may be used with ODD symptom definition based on the threshold "often" and ADHD pervasiveness across contexts not included.

Keywords: Preschoolers, ODD, ADHD, Structured interview

INTRODUCTION

Epidemiological studies have demonstrated that prevalence rates of oppositional defiant disorder (ODD), conduct disorder (CD) and attention-deficit/hyperactivity disorder (ADHD) in preschool children are similar to those later in childhood (Egger & Angold, 2006). Moreover, it has been shown that the symptoms and associated impairment of preschool children diagnosed with ODD, CD and ADHD are likely to persist into school age (Keenan et al., 2011; Lahey et al., 2004). Thus, in view of appropriate treatment of these disorders, accurate identification of preschool children with ODD, CD and ADHD is essential. However, clinicians have long been reluctant to diagnose young children with these disorders as the manifestations of most of the symptoms (e.g., losing temper, hyperactivity) are not atypical per se when observed in a preschooler (Keenan et al., 2007). Thus, a point of concern is the possibility of over-identification of preschool children and diagnosing them wrongly as having ODD, CD or ADHD as most preschoolers exhibit some of the behaviors that fall under the label of disruptive behavior.

The assessment of preschool children necessitates a combination of information about the child's behavior from parents and teachers/caregivers, and clinical observation. Parent reports are particularly important as they are based on children's behaviors across a long period of time. Teacher/caregiver reports are essential as well for they are based on information of children's behavior in another setting than the home. It should be noted that mothers report more symptoms and diagnoses of ODD and CD than teachers, and that mother-teacher agreement on both ODD and CD symptoms and diagnoses is low, due to different contexts and perceptions (Strickland, Hopkins, & Keenan, 2012). Parent and teacher reports may be obtained using standardized questionnaires such as the ASEBA (Achenbach & Rescorla, 2000). Structured diagnostic interviews with parents such as the Kiddie-Disruptive Behavior Disorder Schedule (K-DBDS; Keenan et al., 2007) and the Preschool Age Psychiatric Assessment (PAPA; Egger et al., 2006), however, have the advantage of eliciting information necessary to make clinical diagnoses, based on DSM-IV-TR criteria (American Psychiatric Association, 2000).

The K-DBDS has been shown to have satisfactory interrater and test-retest reliability with regard to ODD and CD symptoms (Keenan et al., 2007). Validity was demonstrated via associations with impairment and the differentiation between referred and nonreferred children (Keenan et al., 2007). Furthermore, the predictive validity of ODD and CD diagnosis has been investigated (Keenan et al., 2011). It was shown that 80% of preschoolers diagnosed with ODD and approximately 60% of preschoolers diagnosed with CD met criteria for the same disorder at 3 year follow-up. Thus, most DSM-IV-TR ODD and CD symptoms can be reliably and validly assessed in preschool children using the K-DBDS. In

contrast, the reliability and validity of ADHD symptoms have not been investigated thus far using the K-DBDS.

Further research is needed regarding the clinical use of the K-DBDS, i.e., to identify individual clinical cases. Two issues with regard to the operationalization of symptoms are crucial here: 1. whether the definition of ODD symptoms should be based either on the threshold "often" or on a more specific frequency threshold; 2. whether, in the definition of ADHD symptoms, pervasiveness across contexts should be included or not.

With regard to ODD, the DSM-IV-TR uses the word "often", but it is difficult to know what the frequency threshold of the word "often" is in the DSM-IV-TR (American Psychiatric Association, 2000) and what it means for the preschool age. The frequency of the behaviors indeed depends on the age of children. For example, for the ODD symptom "often loses temper", for children older than six years of age a frequency at the 90th percentile is two times a week, but for children under the age of six the 90th percentile is represented by a frequency of two to three times a day (Egger & Angold, 2006). The ADHD and Disruptive Behavior Disorders Work Group of the Diagnostic and Statistical Manual for Mental Disorders, Fifth Edition (DSM-5) is considering adopting a frequency threshold for symptoms of ODD (www.dsm5.org/Proposed Revisions). The work group questioned whether more objective frequency criteria can be used to define symptom thresholds instead of using only the criteria "often" for ODD. Keenan (2012), however, argues to keep using the DSM-5 symptom threshold "often" for ODD for preschoolers instead of a frequency threshold because the latter could result in a sizeable group of children who would no longer meet diagnostic criteria, despite demonstrating significant morbidity concurrently and prospectively.

The K-DBDS offers an opportunity to examine whether a frequency threshold should be included or not in the diagnosis of ODD. Indeed, to determine if a behavior is an ODD symptom according to the K-DBDS, two coding methods may be used (K-DBDS Preschool Version Instruction Manual; Keenan, Wakschlag, & Danis, 2001), here called the Qualitative Coding Method and the Specific Coding Method. The Qualitative Coding Method adheres closely to the DSM-IV, i.e., a behavior is coded as a symptom based on the threshold "often". According to the Specific Coding Method, however, a behavior is coded as a symptom not only when it often occurs but also based on follow up probes on the frequency of occurrence, i.e., either many times per day or 1-2 times a day, the latter in combination with a qualitative indicator (severity, duration, support and pervasiveness), thus specifying clinical concern

With regard to ADHD, the issue is whether, in the definition of symptoms, pervasiveness across contexts should be included or not. The DSM-IV-TR criteria for ADHD require that some impairment from the symptoms is present in two or more settings (e.g., at school/work and at home) (American Psychiatric Association, 2000). It is, however, unclear why

children who meet all diagnostic criteria for ADHD but display impairment in only one setting (e.g., seriously impaired academic functioning) should not be given the diagnosis of ADHD. Although functional impairment is a diagnostic criterion for all mental disorders, it has been questioned why impairment in multiple settings should be required only for ADHD (Lahey et al., 2004).

The K-DBDS may be used to examine whether, in the definition of ADHD symptoms, pervasiveness across contexts should be included. Indeed, according to the Qualitative Coding Method a behavior is coded as a symptom based on a threshold for frequency of "a lot", i.e., in a way that is identical to that of the DSM-IV-TR, but without the criterion of pervasiveness across contexts. On the contrary, according to the Specific Coding Method a behavior is coded as a symptom only when additional pervasiveness probes are endorsed, which is in accordance with the DSM-IV-TR. One may question whether the coding method without the pervasiveness question (Qualitative Coding Method) results in the generation of many false positive diagnoses (low specificity) because the symptoms with a frequency threshold "often" or "a lot" may also be a characteristic of typically developing preschool children. On the other hand, the coding method with the pervasiveness question across contexts (Specific Coding Method) could be too stringent; as a result, clinical children might be missed (low sensitivity).

The K-DBDS has also been developed to assess CD symptoms. Some of the CD symptoms are adapted with more developmentally-appropriate language, whereas four CD symptoms (breaking into a house, car or building; running away from home overnight; often staying out late; truancy) are not assessed because of a lack of face validity in preschool children. The two coding methods are in line with those for ODD, although there are only a few CD symptoms (physical aggression; bullies, threatens or intimidates others; lies) with the word "often" in the symptom description. For the other CD symptoms, according to the Qualitative Coding Method the occurrence of a behavior is sufficient to be considered a symptom. For the Specific Coding Method, the frequency thresholds and qualitative indicators vary by symptom.

Another issue in the diagnosis of ODD, CD and ADHD is the total number of symptoms required for a diagnosis. The DSM-5 work group proposed to keep the same thresholds for ODD, CD and ADHD as used with the DSM-IV-TR. For ODD, based on the DSM-IV field trial a symptom threshold of four symptoms was chosen by the DSM-IV Child Disorder Workgroup (Lahey, Applegate, Barkley et al., 1994). However, Angold & Costello (1996) have shown that children with two or three symptoms (plus impairment) are similar to those with four or more symptoms in terms of outcome 1 year later. For ADHD, although results of the field trial suggested a threshold of five hyperactivity-impulsivity symptoms, the DSM-IV Child Disorder Workgroup decided on a six out of nine symptom threshold for the ADHD symptoms of hyperactivity-impulsivity to reduce false positives; with regard

to inattentive symptoms a threshold of six symptoms seemed to be appropriate (Lahey, Applegate, Mc Burnett et al., 1994).

ADHD subtypes (predominantly hyperactive-impulsive, predominantly inattentive, and combined subtype), however, may not be the best descriptors of the disorder in the preschool age group as argued by Lahey et al., (2005) and Lahey & Willcutt (2010). Indeed, the subtypes seem too unstable for use in the clinical assessment of individual children. Preschool children rarely remain in the hyperactive-impulsive subtype over time; rather, they sometimes desist from ADHD but mostly shift to the combined subtype in later years (Lahey et al., 2005). This may be related, in part, to a decline in hyperactive-impulsive symptoms and an increase in inattention symptoms with age, resulting in a later diagnosis of combined subtype for the majority of those initially classified as hyperactive-impulsive subtype (Hardy et al., 2007). Consistent with these findings, the base rate of inattentive symptoms in clinically referred preschoolers has been found to be lower and more variable than the base rate of hyperactive-impulsive symptoms (Murray et al., 2007).

Although the main aim of the study was to examine the clinical validity of the K-DBDS, we first studied the reliability and convergent/divergent validity of the K-DBDS in a sample of referred preschool children with behavioral problems. We then examined clinical validity of K-DBDS symptom scores on a diagnostic group level by comparing a typically developing group (TD) with the following groups; ODD and CD (also called disruptive behavior disorder; DBD), ADHD, Comorbid DBD+ADHD, DBD total (DBD and DBD+ADHD), and ADHD total (ADHD and DBD+ADHD). Finally, to study clinical validity on an individual level we examined which symptom threshold ("often" or a more specific frequency) can be used in the diagnosis of ODD/CD and whether pervasiveness across context is necessary in the diagnosis of ADHD. We also investigated the total number of symptoms required for a diagnosis, by using Receiver Operating Characteristic (ROC) curve analysis, in the diagnosis of ODD, CD and ADHD on an individual level.

METHODS

Participants

Participants were 251 children aged 42-66 months (mean=55 months, SD=7.8); 80% were male, 86% White, 12% Turkish/Moroccans, 2% African, 0.5% Asian. In terms of educational level of the parents: elementary level accounted for 6%, secondary level 25%, intermediate vocational level 34%, higher vocational level 16% and university 19%. General practitioners, pediatricians and well-baby clinics in the province of Utrecht were invited to refer children aged 3.5-5.5 with severe behavior problems to the Outpatient Clinic for Preschool Children with Behavioral Problems, Department of Psychiatry, University Medical

Center Utrecht, the Netherlands. Children were included if they scored at or above the 90th percentile either on the Aggressive Behavior scale or on the Attention Problems scale of the Child Behavior Checklist/1.5-5 (CBCL/1.5-5) or Child-Teacher Report Form/1.5-5 (C-TRF/1.5-5) (Achenbach & Rescorla, 2000). In the study, 193 referred children were included, of whom 154 were diagnosed with ODD, CD or ADHD.

The typically developing (TD) group (N=58) was recruited from regular elementary schools and daycare centers. Children with a score in the clinical range either on the Attention Problems scale or on the Aggressive Behavior scale of the CBCL/1.5-5 or C-TRF/1.5-5 were excluded. All the children (referred and TD) with an IQ below 70, estimated with the average score on the Raven Coloured Progressive Matrices (Raven, Court, & Raven, 1998) and the Peabody Picture Vocabulary Test-III-NL (Dunn & Dunn, 2005; Schlichting, 2005), were excluded. None of the participants were on medication.

In order to study the reliability and convergent/divergent validity, all children referred with behavioral problems (N=193) and all TD children (N=58) were included. To examine interrater reliability, 27% (N=68) of the digitally recorded K-DBDS were randomly selected and coded by a second, different research assistant. Of those selected, 79% (N=54) were from clinically referred children, 87% were boys. To examine test-retest reliability 14% (N=35) of the original group was re-assessed after 1 week (mean=7.6 days, SD=1.7, range 6-16 days), 74% (N=26) of these preschoolers were referred for behavior problems and 26% (N=9) were TD children, 81% were boys.

For the purpose of the clinical validity study, each child of the referred group was given a diagnosis of either DBD (ODD, CD or DBD-NOS [i.e. defined as either three ODD symptoms and at least one CD symptom or at least one ODD symptom and two CD symptoms]), ADHD or comorbid (DBD and ADHD), or no diagnosis. As there is no gold standard by which to diagnose a psychiatric disorder, the procedure of best-estimate diagnosis was used, i.e., an independent child psychiatrist and clinical child psychologist considered the information about the child, discussed all the findings and reached consensus (Lord et al., 2006). Both clinicians were unfamiliar with the child. The information about the child consisted of:

1. Interview with the parents about the reasons for referral and the development of the child; 2. The scores of the Attention Problem scale and the Aggressive Behavior Scale of the CBCL completed by parents (CBCL/1.5-5) and the C-TRF completed by teachers or daycare caregivers (C-TRF/1.5-5); 3. The scores on the Children's Global Assessment Scale (CGAS; Schaffer et al., 1983): i.e., a measure of the impairment of the functioning of the child's behavior, filled out by the parents as well as the teacher/caregiver; 4. The scores on the Raven Coloured Progressive Matrices (Raven et al., 1998) and the Peabody Picture Vocabulary Test III-NL (Dunn & Dunn 2005; Schlichting, 2005); 5. The results of the executive function tasks; Three inhibition tasks (GoNoGo task, Snack Delay and Shape School),

and two working memory tasks (Delayed Alternation and Nine boxes) (Schoemaker et al., 2012); 6. The videotaped observations of the child's behavior during the psychological tasks.

In order to reach a clinical consensus diagnosis, the DSM-IV-TR criteria of ODD, CD or DBD-NOS and ADHD were strictly applied. On the basis of the consensus best-estimate diagnosis, children were positioned in the DBD, ADHD or comorbid (DBD and ADHD) or TD group. In this way, the children were divided into four groups: TD group (N=58), DBD group (N=39), ADHD group (N=58), ADHD+DBD group (N=57). The characteristics of the four groups are displayed in Table 1. In the clinical validity study, there were group differences in estimated IQ (p<0.001), with the TD group significantly outperforming the three clinical groups, who performed similarly to each other. The DBD group and the comorbid DBD+ADHD group significantly differed in gender from the TD group (p<0.05) and the DBD group significantly differed in age from the TD group (p<0.01).

Table 1 *Means (and SD) for the Demographics and Control Variables in the Four Groups*

	TD (N=58)	DBD (N=39)	ADHD (N=58)	DBD+ADHD (N=57)
	Mean (SD)	Mean (SD)	Mean (SD)	Mean (SD)
Age (months)	56.0 (7.3)	51.5 (8.2)**	54.5 (7.4)	54.0 (7.9)
% Boys	71	90*	76	86*
IQ[a]	112.4 (10.4)	101.8(10.2)***	101.6(11.5)***	101.0 (12.3) ***
CBCL aggression	50.5 (1.3)	74.9 (10.4)***	63.3 (10.6)***	76.3 (11.7)***
CBCL attention	50.8 (2.1)	64.1(8.3)***	68.6 (6.5)***	69.9 (6.3)***
TRF aggression	52.0 (3.1)	62.2 (10.2)***	62.1 (9.1)***	69.0 (11.2)***
TRF attention	52.0 (3.8)	58.9 (8.3)***	70.4 (10.7)***	70.6 (11.6)***

Note. TD=typically developing; DBD=disruptive behavior disorder; ADHD=attention-deficit/hyperactivity disorder; IQ[a]=estimate based on mean Raven and Peabody; CBCL=Child Behavior Checklist; TRF=Teacher Report Form; *p<.05. **p <.01. ***p<.001. (compared with the TD Group).

Procedure

Children were assessed during one morning session. First, general intellectual functioning was assessed followed by the executive functions tasks. After a break, the DB-DOS and K-DBDS were administered. The tasks were administered in a quiet and non-distractible room with a one-way mirror. Parents received a small financial compensation for participating, children received two small gifts. Written informed consent from the parents was

gathered before participating. The Medical Ethical Review Committee of the University Medical Center Utrecht, the Netherlands approved this study.

Measures

K-DBDS. The K-DBDS is a semi-structured clinical interview assessing DSM-IV (American Psychiatric Association, 2000) ODD, CD and ADHD symptoms in preschool children (Keenan et al., 2007). It is based on the Schedule for Affective Disorders and Schizophrenia for School-Age Children (Orvaschel & Puig-Antich, 1995). Administrating the K-DBDS takes approximately 45 minutes, coding takes approximately 10 minutes. In our study, child psychiatrists and clinical child psychologists administrated the K-DBDS. To determine if a behavior is a symptom according to the K-DBDS, two coding methods have been used (K-DBDS Preschool Version Instruction Manual; Keenan, Wakschlag, & Danis, 2001). For ODD: 1. The Qualitative Coding Method: the threshold "often" is used. 2. The Specific Coding Method: the threshold "often" is used and follow-up probes on the frequency of occurrence, i.e., either many times per day or 1-2 times a day and a qualitative indicator (severity, duration, support and pervasiveness). With regard to CD, the two coding methods are in line with those for ODD (for differences see Introduction). For ADHD: 1. The Qualitative Coding Method: the threshold "a lot" is used. 2. The Specific Coding Method: the threshold "a lot" is used in combination with the pervasiveness threshold, i.e., the behavior must be endorsed in more than one context.

CBCL/C-TRF. Disruptive behavior symptoms and attention problems were assessed using the scores of the Attention Problem scale and the Aggressive Behavior Scale of the CBCL/1.5-5 (Achenbach & Rescorla, 2000) completed by parents and the C-TRF/1.5-5 (Achenbach & Rescorla, 2000) completed by teachers or daycare caregivers.

Children's Global Assessment Scale (CGAS; Schaffer et al., 1983). To provide a global assessment of functioning, parents and teachers/caregivers completed the non-clinician version of the CGAS is scored from 0-100 with lower scores indicating greater impairment.

Impact on family scale (IFS; Sheeber & Johnson, 1992). Completed by the mother to assess the social, financial and personal burden resulting from child behavior problems. In the present study, the Cronbach alpha for this scale was 0.95.

Social Skills Rating Scale (SSRS; Gresham & Elliot, 1990). Parents' reports on the summary score of the IFS were used to evaluate children's social skills. This scale was completed during the parent module of the DB-DOS. In the present study the Cronbach alpha for this scale was 0.91.

Early Child Inventory (ECI; Gadow & Sprafkin, 1996). Parents and teachers/caregivers reported on child behavior symptoms with the ECI, a DSM based checklist. The pervasive developmental scale of the ECI was used in view of examining divergent validity of DBD

and ADHD scores. In the present study the Cronbach alpha for this scale was 0.77 (parent) and 0.82 (teacher).

Statistical analyses

For all scales used, the percentage of missing data was at most 4.1%, with the exception of the ECI teacher questionnaire (4.8%) and the CGAS teacher questionnaire (6.8%). For the reliability analyses the internal consistency was examined with Cronbach alpha. Interrater agreement and test-retest reliability were assessed using intraclass correlation coefficients. Convergent and divergent validity were examined by computing the correlations of K-DBDS total symptom scores and multi-method assessment of child functioning as evaluated by parents' (CGAS, CBCL/1.5-5, SSRS, ECI), teacher/caregivers' reports (CGAS, C-TRF/1.5-5, ECI).

To assess clinical validity, examined on a diagnostic group level, analyses of (co)variance (AN(C)OVA) were conducted to compare the DBD and ADHD groups (with and without comorbid children) and the comorbid group separately with the TD group while controlling for age and gender for the DBD group/DBD total group and controlling for gender for the comorbid DBD+ADHD group. The analyses were carried out with and without controlling for IQ, using Bonferroni corrections. Statistical significance was based on two-sided tests. The two coding methods were compared as well. Patients groups were not compared with each other as this was not an aim of the study.

To operationalize a symptom in the diagnosis of ODD, CD and ADHD on an individual level in the preschool period, the two coding methods were compared. An appropriate cutoff score of the number of symptoms for ODD, CD and ADHD, using the two coding methods, was required to divide clinical from nonclinical cases. For this Receiver Operating Characteristic (ROC) curve analysis was used. Receiver Operating Characteristic (ROC) curves are a useful way to interpret sensitivity and specificity levels as a function of varying cutoff scores of a quantitative scale. An overall indication of the diagnostic accuracy of cutoffs is the area under the ROC-curve (AUC) (Krzanowski & Hand, 2009).

RESULTS

Reliability

Reliability of the ODD, CD and ADHD domain and symptoms. Three CD symptoms (sexually coercive, sets fire and steals with confrontation) were rated as present for fewer than 3 children. Because estimates of reliability can be unduly biased by differences between a single pair of raters for items with a very low occurrence, these items were excluded from reliability analyses (Keenan et al., 2007).

Internal consistency. The ODD domain exhibited moderate to good internal consistency for the two different scoring methods (Table 2). The CD domain (without the three above mentioned CD items) exhibited moderate internal consistency for the CD Qualitative Coding Method and a low internal consistency for the CD Specific Coding Method. As the Cronbach's alpha for the latter method are clearly too low, further analyses were only conducted using the Qualitative Coding Method. The ADHD domain exhibited good internal consistency for the two different scoring methods.

Interrater reliability. Results showed excellent interrater reliability for the total number of symptoms for ODD, CD and ADHD (Table 2). The symptom reliability showed good to excellent interrater reliability for ODD symptoms (Table is available from first author) with for the Qualitative Coding Method kappas ranging from 0.78 to 1.0, and the Specific Coding Method kappas ranging from 0.90 to 1.0. The CD items: "physically cruelty to people", "forced sexual activity", "stealing with confrontation" and "fire setting" had base rates <3 in both samples and were excluded from reliability analyses. Interrater reliability for CD symptoms for the Qualitative Coding Method had kappas ranging from 0.65 to 1.0. The symptom reliability for ADHD showed excellent reliability with the Qualitative Coding Method, with kappas ranging from 0.94 to 1.0, and for the Specific Coding Method kappas ranging from 0.90 to 1.0.

Test-retest reliability. Results of correlation coefficients (ICC) showed good (ODD domain), moderate (CD domain), to excellent (ADHD domain) reliability of the number of symptoms (Table 2). The symptom test-retest reliability showed moderate test-retest reliability for ODD symptoms (Table available from first author) with the Qualitative Coding Method

Table 2 *Reliability of DSM-IV ODD, CD and ADHD Domains*

Domain	Cronbach alpha	Interrater ICC	Test-retest ICC
	N=251	N=68 (14 TD) *(CI 95%)*	N=35 (9 TD) *(CI 95%)*
ODD (QCM)	.75	.98 (.97-.99)	.92 (.83-.96)
ODD (SCM)	.75	1.0 (.99-1.0)	.86 (.73-.93)
CD (QCM)	.67	.96 (.94-.98)	.70 (.40-.85)
CD (SCM)	.54	.99 (.99-1.0)	.64 (.29-.82)
ADHD (QCM)	.88	1.0 (.99-1.0)	.97 (.94-.98)
ADHD (SCM)	.88	1.0 (.99-1.0)	.97 (.94-.98)

Note. TD=typically developing; ODD (QCM)=Oppositional Defiant Disorder (Qualitative Coding Method); ODD (SCM)=Specific Coding Method; CD=Conduct Disorder; ADHD=attention-deficit/hyperactivity disorder; ICC=Intraclass Correlation Coefficient. Three CD symptoms (sexually coercive, sets fire and steals with confrontation) were rated as present fewer than 3 times; Because estimates of reliability can be unduly biased by differences between a single pair of raters for items with a very low occurrence, these items were excluded from reliability analyses.

kappas ranging from 0.30 to 0.75, and for the Specific Coding Method kappas ranging from 0.24 to 0.77. For the CD test-retest symptom score, many items were rated as present fewer than 3 times in both tests and these items were excluded from reliability analyses. The symptom reliability showed moderate test-retest reliability for CD symptoms with the Qualitative Coding Method kappas ranging from 0.37 to 0.82. The symptom test-retest reliability for ADHD showed moderate reliability for ADHD symptoms with the Qualitative Coding Method and the Specific Coding Method kappas ranging from 0.30 to 0.94.

Validity

Convergent validity. Significant correlations were established between K-DBDS total symptom scores and parent and teacher/caregiver questionnaires of symptoms and impairment (CGAS, CBCL/1.5-5, C-TRF/1.5-5, IFS) (Table 3).

Divergent clinical validity. Small but significant correlations were found between the majority of K-DBDS scores and parent reports of social skills on the SSRS and the correlation between the ECI-PDD part and the K-DBDS symptom scores was weak.

Table 3 *Associations of the Kiddie-Disruptive Behavior Disorder Schedule and Parent and Teacher Assessments of Child Disruptive Behavior and Functioning*

	ODD	ODD	CD	ADHD	ADHD
	QCM	SCM	QCM	QCM	SCM
Parent					
CGAS	-.64***	-.57***	-.50***	-.69***	-.68***
CBCLagg	.68***	.65**	.61***		
CBCLatt				.75***	.76***
IFS (M)	.65***	.62***	.52***	.64***	.62***
SSRS	-.50***	-.47***	-.38***	-.48***	-.46***
ECI	.44***	.44***	.40***	.39***	.39***
Teacher					
CGAS	-.27***	-.21**	-.35***	-.45***	-.47***
TRF agg	.26***	.21**	.35***		
TRF att				.41***	.42***
ECI	.11*	.07	.20**	.22**	.22**

Note. ODD=Oppositional Defiant Disorder; CD=Conduct disorder; QCM=Qualitative Coding Method; SCM=Specific Coding Method; Three CD symptoms (sexually coercive, sets fire and steals with confrontation) were rated as present fewer than 3 times; ADHD=attention-deficit/hyperactivity disorder; CGAS=Children's Global Assessment Scale; CBCLagg=Child Behavior Checklist Aggression Scale; CBCLatt=Child Behavior Checklist Attention Scale; IFS=Impact on Family Scale; SSRS=Social Skills Rating Scale; IFS (M)=Impact on Family Scale (Mother); ECI=Early Child Inventory; *p <.05. **p <.01. ***p <.001.

Clinical validity on a group level. Table 4 illustrates the means and standard deviations of the K-DBDS scores based on the different coding methods in the six different groups: TD, DBD, ADHD, Comorbid (DBD+ADHD), DBD total (DBD and DBD+ADHD), and ADHD total (ADHD and DBD+ADHD). Analyses of (co)variance were used to examine group differences in K-DBDS total symptom scores between the TD group and the clinical groups. All clinical groups differed significantly (p<.001) from the TD group on all K-DBDS total symptom scores based on the different coding methods using covariate analyses for gender for the DBD, the comorbid DBD+ADHD and the total DBD group, and for age for the DBD and the DBD total group. The results are similar for the analyses when controlling for IQ.

Clinical validity on an individual level. To study the clinical validity of the K-DBDS on an individual level with respect to ODD, a cutoff score of the total number of ODD symptoms for the two coding methods was examined. First we determined the best cutoff score for the total number of ODD symptoms for the Qualitative Coding Method (Table 5). Therefore, we used the total ODD group (ODD and ODD+ADHD based on consensus

Table 4 *Means (SD) and F-value of the K-DBDS in the Six Groups*

	TD	DBD	ADHD	DBD+ ADHD	DBD only and DBD+ ADHD	ADHD only and DBD+ ADHD
	(N=58)	(N=39)	(N=58)	(N=57)	(N=96)	(N=115)
	Mean(SD)	Mean(SD)	Mean(SD)	Mean(SD)	Mean(SD)	Mean(SD)
	F-value	F-value	F-value	F-value	F-value	F-value
ODD (QCM)	1.2 (1.3)	5.6 (1.5)	3.5 (1.8)	5.5 (1.6)	5.6 (1.6)	4.5 (2.0)
		233.7***	62.3***	252.2***	313.6***	131.3***
ODD (SCM)	.3 (.5)	3.6 (1.9)	1.6 (1.5)	3.8 (1.8)	3.7 (1.9)	2.7 (2.0)
		137.3***	36.2***	179.7***	166.2***	76.8***
CD (QCM)	.4 (.7)	2.8 (1.5)	1.5 (1.2)	3.3 (2.0)	3.1 (1.8)	2.4 (1.8)
		58.3***	40.1***	101.3***	83.1***	62.9***
ADHD (QCM)	1.7 (2.0)	7.1 (4.0)	10.5 (2.9)	11.1 (3.3)	9.4 (4.1)	10.8 (3.1)
		63.7***	358.8***	320.2***	170.2***	404.4***
ADHD (SCM)	1.5 (1.9)	6.7 (3.9)	10.2 (2.9)	10.5 (3.4)	8.9 (4.0)	10.3 (3.2)
		63.6***	362.9***	295.0***	163.6***	386.2***

Note. SD=Standard Deviation; K-DBDS=Kiddie-Disruptive Behavior Disorder Schedule; TD=typically developing; DBD=disruptive behavior disorder; ADHD=attention-deficit/hyperactivity disorder; ODD=Oppositional Defiant Disorder; CD=Conduct Disorder; QCM=Qualitative Coding Method; SCM=Specific Coding Method; Three CD symptoms (sexually coercive, sets fire and steals with confrontation) were rated as present fewer than 3 times.
***p <.001.

best estimate diagnosis) with the TD group for the ROC curve analysis. Using a cutoff score of four ODD symptoms yielded the best balance between a sensitivity of 87% and a specificity of 93%; a sensitivity of 87% means that nearly 9 out of 10 children with ODD will obtain the diagnosis according to consensus best-estimate diagnosis and a specificity of 93% means that 9 out of 10 TD children will not be diagnosed with ODD whereas 7% (nearly 1 out of 10) will be diagnosed with ODD while the diagnosis is not given for the Qualitative Coding Method; the ROC analyses yielded an AUC of .97 (95% Confidence Interval (C.I.)=[.95-1.0]).

For the cutoff score of four ODD symptoms, the Specific Coding Method yielded a sensitivity of 56% and a specificity of 100%. Using a cutoff score of two ODD symptoms yielded the best balance between sensitivity (83%) and specificity (98%) for the Specific Coding Method; the ROC analyses yielded an estimated AUC of .96 (95% C.I.=[.93-.96]). With

Table 5 *ROC Curve*

	K-DBDS score ≥	Sensitivity	Specificity
ODD N=92			
ODD (QCM)	2	99	62
ODD (QCM)	3	98	86
ODD (QCM)	**4**	**87**	**93**
ODD (QCM)	5	77	98
ODD (SCM)	1	97	71
ODD (SCM)	**2**	**83**	**98**
ODD (SCM)	3	71	100
ODD (SCM)	4	56	100
ADHD HI N=110			
ADHD HI (QCM)	4	93	91
ADHD HI (QCM)	**5**	**83**	**98**
ADHD HI (QCM)	6	64	100
ADHD HI (SCM)	**4**	**90**	**95**
ADHD HI (SCM)	5	77	98
ADHD HI (SCM)	6	62	100

Note. ROC=Receiver Operating Characteristic; K-DBDS=Kiddie-Disruptive Behavior Disorder Schedule; ODD=Oppositional Defiant Disorder; ADHD HI=attention-deficit/hyperactivity disorder, predominantly hyperactive-impulsivetype; QCM=Qualitative Coding Method; SCM=Specific Coding Method; Bolded numbers indicate maximized sensitivities and specificities depending on criteria used in selecting cutoff scores.

regard to CD, comparison of the two coding methods was not possible as according to consensus best estimate diagnosis only one child was diagnosed with CD.

With regard to ADHD, according to consensus best estimate diagnosis, only six children were diagnosed with the predominantly inattentive type of ADHD. As a consequence, it was not possible to examine the clinical validity on an individual basis for the inattentive type of ADHD type. Neither was it possible to differentiate between the ADHD predominantly hyperactive-impulsive type and the ADHD combined subtype. Thus, comparison of the two coding methods was only possible for the total ADHD predominantly hyperactive-impulsive type group (ADHD predominantly hyperactive-impulsive type and DBD+ADHD predominantly hyperactive-impulsive type; N=110). If a cutoff score of six symptoms of the total number of ADHD hyperactive-impulsive symptoms was used, 64% (Qualitative Coding Method) and 62% (Specific Coding Method) of the children in the ADHD predominantly hyperactive-impulsive type total group (ADHD predominantly hyperactive-impulsive type and DBD+ADHD predominantly hyperactive-impulsive type) were given an ADHD diagnosis. The specificity was 100% for both coding systems, see Table 5. The ROC analyses yielded an estimated AUC for both coding methods of .98 (95% C.I.=[0.96-1.0]).

To optimize the clinical validity on an individual level, a cutoff score for each coding method of the total number of ADHD hyperactive-impulsive symptoms was examined. We used the total ADHD predominantly hyperactive-impulsive type group (ADHD predominantly hyperactive-impulsive type and DBD+ADHD predominantly hyperactive-impulsive type) with the TD group for the ROC curve analysis. Using a cutoff score of five ADHD hyperactive-impulsive symptoms, yielded a sensitivity of 83% and a specificity of 98%, for the Qualitative Coding Method, which is satisfactory (Table 5). Using a cutoff score of five ADHD hyperactive-impulsive symptoms for the Specific Coding Method yielded a slightly worse sensitivity (77%) and equal specificity (98%).

DISCUSSION

The results of reliability (internal consistency, interrater reliability, test retest reliability) and validity (convergent and divergent) of the K-DBDS in a European sample using different coding methods are similar to the findings of the Keenan et al., study (2007) in a U.S. sample. With respect to the clinical validity of the K-DBDS on a group level, all clinical groups differed from the TD group on ODD/CD and ADHD total symptom scores. The current study not only replicated but also extended the US study by examining clinical validity on an individual level. Hence, to operationalize ODD and ADHD symptoms in the preschool period two coding methods were compared.

With regard to ODD, using "often" as threshold for frequency of behaviors (Qualitative Coding Method) and a cutoff score of four ODD symptoms leads to a sensitivity of 87% and specificity of 93%. We compared this coding method with a more specific frequency threshold (Specific Coding Method) using a cutoff score of four ODD symptoms; the Specific Coding Method yielded a sensitivity and specificity of 56% and 100%. Since for clinical use a high sensitivity and an acceptable specificity (>70%) are required (Kim & Lord, 2012), sensitivity of the Specific Coding Method is clearly less satisfactory than the sensitivity of the Qualitative Coding Method. Thus, using other criteria than the qualification "often" to define a symptom as suggested by the DSM-5 ADHD and Disruptive Behavior Disorders Work Group will result in a sizeable group of children who will no longer be diagnosed with ODD, which is in line with the results of the study by Keenan (2012). On the other hand, the specificity of the Qualitative Coding Method is 93% which means that 7% of children without a disorder would be diagnosed with ODD. Clearly, diagnosing a child with a psychiatric disorder while there is no disorder is inappropriate. We should, however, keep in mind that in everyday clinical practice the decision of diagnosing a child with ODD is not based on the results of a single measure such as the K-DBDS but on the combination of results from multiple measures such as standardized parent and teacher/caregiver rating scales and an observation of the child (Le Couteur & Gardner, 2008; Matthys & Lochman, 2010).

Using ROC curve analysis we explored whether a better sensitivity/specificity balance could be obtained for the Specific Coding Method for ODD. Surprisingly, the Specific Coding Method with a cutoff score of two symptoms for ODD yielded a sensitivity and specificity of 83% and 98%. Thus, results of the Qualitative Coding Method for ODD are similar to those of the Specific Coding Method. However, the Specific Coding Method has two symptoms as a cutoff score compared to the cutoff score of four ODD symptoms according to DSM-IV-TR criteria (American Psychiatric Association, 2000). Apparently, if the threshold is "often" and the follow up probes specify (severe) clinical concern, a lower total symptom threshold is generated. It may be that the inclusion of clinical concern refers not so much to the presence of symptoms as to the severity of symptoms. As a result, preschool children with only two severe ODD symptoms will be diagnosed with ODD. We, however, do not suggest that the frequency threshold of four ODD symptoms should be changed as the predictive validity of ODD using this criterion has been demonstrated (Keenan et al., 2011). Thus, a cutoff score of four ODD symptoms using "often" as threshold for frequency of behavior seems satisfactory.

For ADHD in preschool children, ADHD subtypes (predominantly hyperactive/impulsive, predominantly inattentive, and combined subtype) may not be the best descriptors of the disorder (Lahey et al., 2005; Lahey & Willcutt, 2010). In the present study, using the procedure of best-estimate diagnosis we felt uncertain to about diagnosis of the inattentive subtype. This is consistent with the study by Murray et al., (2007) in which the base rate of inattentive symptoms in clinically referred preschoolers was lower and more variable than hyperactive-

impulsive symptoms. As a consequence it was not possible to differentiate between the ADHD predominantly hyperactive-impulsive type and the ADHD combined type. Thus, results with regard to ADHD were based on the presence of the numbers of hyperactive-impulsive symptoms only. Using a cutoff score of six ADHD hyperactive-impulsive symptoms, leads to a moderate sensitivity of 64% for the Qualitative Coding Method and of 62% for the Specific Coding Method and a specificity of 100% for both. This sensitivity is clearly unsatisfactory for a clinical assessment tool. We explored, by using ROC curve analysis, whether a better sensitivity could be obtained and a satisfactory specificity, using a lower cutoff score of 5 symptoms instead of 6 symptoms for ADHD (see DSM-IV field trial; Lahey et al., 1994b). A cutoff score of five ADHD hyperactive-impulsive symptoms yielded a good sensitivity and specificity (83% and 98%) for the Qualitative Coding Method, whereas for the Specific Coding Method the sensitivity and specificity balance was slightly worse (77% and 98%). In sum, the present study shows that ADHD in preschool children may be diagnosed based on a symptom threshold of "a lot" for five ADHD symptoms, while the pervasiveness threshold that the behavior must be endorsed in more than one context does not need to be included.

A number of limitations should be noted. First, research questions with regard to CD and ADHD subtypes could not be answered as only one child with CD and only six children with ADHD Inattentive type were diagnosed. Second, assessment of comorbidity with anxiety and mood disorders was not included in the study. Third, the test-retest reliability was only moderate for ODD in line with the Keenan et al. study (2007) and could not be validly assessed for CD because the low frequency of scores on CD symptoms. Fourth, in the course of studying divergent clinical validity, small but significant minor correlations were found between the majority of K-DBDS scores and parent reports of social skills on the SSRS. One may question whether the SSRS is an appropriate measure to examine divergent clinical validity as social skills are less well developed in children with DBD and ADHD (Ronk, Hund, & Landau, 2011). Therefore we added the ECI-PDD part in view of studying divergent validity. As expected, the correlation between the ECI-PDD part and the K-DBDS symptom scores was weak.

Clinical implications

Results of the present study suggest that for the diagnosis of ODD in preschool children it is appropriate to code the K-DBDS using the Qualitative Coding Method. According to this method, the behavior is coded in a way that is identical to that of the DSM-IV-TR, i.e., the threshold "often" is used, with a cutoff score of four ODD symptoms. With a sensitivity of 87% and a specificity of 93% it is important to keep in mind that diagnosing a child with ODD is not only based on information from the parent but also from teacher/caregiver, and from clinical observation. The Qualitative Coding Method (without the pervasiveness across contexts probes) may also be used for the diagnosis of ADHD, at least for the

hyperactive-impulsive type, using five ADHD hyperactive-impulsive symptoms as a cutoff score. With a sensitivity of 83% and a specificity of 98% there is minimal risk for false positives. Clearly, diagnosing a young child with ADHD and starting pharmacological or psychological treatment in the absence of the disorder is inappropriate. When a diagnosis is missed in a young child and symptoms persist, a clinical reassessment is needed to give further clarification. Using the Qualitative Coding Method for ODD and ADHD makes the K-DBDS less complicated to administer and less time consuming, which is an improvement for clinical practice.

Acknowledgments

Disclosure: Ms. Bunte, Schoemaker, and Dr. Hessen, van der Heijden, and Matthys report no biomedical financial interests or potential conflicts of interest.

The authors are grateful to the parents and children who participated in this study.

REFERENCES

Achenbach, T.M. & Rescorla, L.A. (2000). *Manual for the ASEBA preschool forms and Profiles.* Burlington: Research Center for Children, Youth & Families VT: University of Vermont.

American Psychiatric Association. (2000). *Diagnostic and statistical manual of mental disorders (4th Edition, text-revised)*. Washington, DC: Author.

Angold, A. & Costello, E.J. (1996). Toward establishing an empirical basis for the diagnosis of oppositional defiant disorder. *Journal of the American Academy of Child and Adolescence Psychiatry, 35*, 1205-12.

Dunn, L.M. & Dunn, L.M. (2005). *Peabody picture vocabulary test (III-NL)* [Dutch version by Schlichting]. Amsterdam: Harcourt Assessment.

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.

Egger, H.L., Erkanli, A., Keeler, G., Potts, E., Walter, B.K. & Angold, A. (2006). Test-retest reliability of the Preschool Age Psychiatric Assessment (PAPA). *Journal of the American Academy of Child and Adolescence Psychiatry, 45*, 538-549.

Gadow, K. & Sprafkin, J. (1996). *Early childhood inventories manual.* Stony Brook, NY: Checkmate Plus.

Gresham, F. & Elliot, S. (1990). *Social Skills Rating System*. Circle Pines, MN: American Guidance Service.

Hardy, K.K., Kollins, S.H., Murray, D.W., Riddle, M.A., Greenhill, L., Cunningham, C., ….Chuang, S.Z. (2007). Factor structure of parent- and teacher-rated attention-deficit/hyperactivity disorder symptoms in the Preschoolers with Attention -Deficit/Hyperactivity Disorder Treatment Study (PATS). *Journal of Child and Adolescence Psychopharmacology, 17*, 621-34.

Keenan, K. (2012). Mind the gap: Assessing impairment among children affected by proposed revisions to the diagnostic criteria for oppositional defiant disorder. *Journal of Abnormal Psychology, 121*, 352-9.

Keenan, K., Boeldt, D., Chen, D., Coyne, C., Donald, R., Duax J,…Humphries, M. (2011). Predictive validity of DSM-IV oppositional defiant and conduct disorders in clinically referred preschoolers. *Journal of Child Psychology and Psychiatry, 52*, 47-55.

Keenan, K., Wakschlag, L.S., & Danis, B. (2001). Kiddie-Disruptive Behavior Disorder Schedule (version 1.1).

Keenan, K., Wakschlag, L.S., Danis, B., Hill, C., Humphries, M., Duax, J. & Donald, R. (2007). Further evidence of the reliability and validity of DSM-IV ODD and CD in preschool children. *Journal of the American Academy of Child and Adolescent Psychiatry, 26*, 457-468.

Kim, S.H. & Lord, C. (2012). New Autism Diagnostic Interview-revised algorithms for toddlers and young pre-schoolers from 12-47 months of age. *Journal of Autism and Developmental Disorders, 42*, 82-93.

Krzanowski, W.J. & Hand, D.J. (2009). *ROC Curves for Continuous Data.* London: Chapman and Hall. .

Lahey, B.B., Applegate, B., Barkley, R.A., Garfinkel, B., McBurnett, K., Kerdyk, L., … Newcorn, J. (1994). DSM-IV field trials for oppositional defiant disorder and conduct disorder in children and adolescents for oppositional defiant disorder and conduct disorder. *American Journal of Psychiatry, 151*, 1163-1171.

Lahey, B.B., Applegate, B., McBurnett, K., Biederman, J., Greenhill, L., Hynd, G.W.,…Shaffer, D. (1994). DSM-IV field trials for attention deficit hyperactivity disorder in children and adolescents. *American Journal of Psychiatry, 151*, 1673-1685.

Lahey, B.B, Pelham, W.E., Loney, J., Kipp, H., Ehrhardt, A., Lee, S.S., …Massetti, G. (2004). Three-year predictive validity of DSM-IV attention deficit hyperactivity disorder in children diagnosed at 4-6 years of age. *American Journal of Psychiatry, 161*, 2014-20.

Lahey, B.B., Pelham, W.E., Loney, J., Kipp, H., Lee, S.S. & Wilcutt, E. (2005). Instability of the DSM-IV subtypes of ADHD from preschool through elementary school. *Archives of General Psychiatry, 62*, 896-902.

Lahey, B.B., & Willcutt, E.G. (2010). Predictive validity of a continuous alternative to nominal subtypes of attention-deficit/hyperactivity disorder for DSM-V. *Journal of Clinical Child and Adolescence Psychology, 39*, 761-75.

Le Couteur, A., & Gardner, F. (2008). Use of structured interviews and observational methods in clinical settings. In M. Rutter, D. Bishop, D. Pine, S. Scott, J. Stevenson, E. Taylor & A. Thapar (Eds.), Rutter's child and adolescent psychiatry, fifth edition (pp. 271-288). Oxford: Blackwell.

Lord, C., Risi, S., DiLavore, S., Shulman, C., Thurm, A. & Pickles, A. (2006). Autism from 2 to 9 years of age. *Archives of General Psychiatry, 63*, 694-701.

Matthys, W. & Lochman, J.E. (2010). *Oppositional defiant disorder and conduct disorder in childhood*. Chichester: Wiley-Blackwell.

Murray, D.W., Kollins, S.H., Hardy, K.K., Abikoff, H.B., Swanson, J.M., Cunningham, C., … Chuang, S.Z. (2007). Parent versus teacher ratings of attention -deficit/hyperactivity disorder symptoms in the Preschoolers with Attention -Deficit/Hyperactivity Disorder Treatment Study (PATS). *Journal of Child and Adolescence Psychopharmacology, 17*, 605-20.

Orvaschel, H. & Puig-Antich, J. (1995). Schedule for Affective Disorders and Schizophrenia for School-Age Children-Epidemiologic 5th Version. Ft. Lauderdale, FL: Nova University.

Raven, J.C., Court, J.H., & Raven, J. (1998). *Raven coloured progressive matrices*. Oxford: Oxford Psychologist Press.

Ronk, M.J., Hund, A.M. & Landau, S. (2011). Assessment of social competence of boys with attention deficit/ hyperactivity disorder: problematic peer entry, host responses, and evaluations. *Journal of Abnormal Child Psychology*, *39*, 829-40.

Schaffer, D., Goud, M.S., Brasic, J., Ambrosini, P., Fisher, P., Bird, H. & Aluwahlia, S. (1983). A children's global assessment scale (CGAS). *Archives of General Psychiatry*, *40*, 1228-1231.

Schlichting, L. (2005). Peabody picture vocabulary test III-NL. Amsterdam: Hartcourt Assessment.

Schoemaker, K., Bunte, T., Wiebe, S.A., Espy, K.A., Dekovic, M., & Matthys, W. (2012). Executive function deficits in preschool children with ADHD and DBD. *Journal of Child Psychology and Psychiatry*, *53*, 111-119.

Sheeber, L., & Johnson, J. (1992). Applicability of the Impact on Family Scale for assessing families with behaviorally difficult children. *Journal of Child Psychology and Psychiatry, 17*, 155-159.

Strickland, J., Hopkins, J., & Keenan, K. (2012). Mother-Teacher agreement on preschoolers' symptoms of ODD and CD: Does Context matter? *Journal of abnormal Child Psychology*, doi 10.1007/s10802-012-9622-y.

APPENDIX

Table Supplement *Reliability of DSM-IV ODD, CD and ADHD symptoms*

Domain	Interrater (QCM) (N=68) kappa	Interrater (SCM) (N=68) kappa	Test-retest (QCM) (N=35) kappa	Test-retest (SCM) (N=35) kappa
ODD				
1.	0.88	1.0	0.57	0.77
2.	0.96	0.90	0.52	0.44
3.	0.82	0.96	0.77	0.72
4.	1.0	1.0	0.54	0.24
5.	0.88	0.91	0.46	0.30
6.	0.97	1.0	0.58	0.57
7.	0.78	0.95	0.75	0.31
8	0.90	1.0	0.30	0.31
CD				
12.	0.74	1.0	0.54	0.74
13.	1.0	#	#	#
14.	0.94	0.95	0.50	0.62
15.	1.0	#	#	#
16.	0.93	0.88	0.58	0.44
18.	0.65	#	0.37	#
19.	0.92	#	0.82	#
21.	0.95	#	0.58	#
ADHD				
25.	1.0	0.97	0.49	0.42
26.	0.94	0.97	0.71	0.72
27.	0.97	0.97	0.54	0.59
28	0.97	0.93	0.74	0.66
29	1.0	1.0	#	#
30	1.0	0.93	0.92	0.75
31	0.97	0.97	0.82	0.88
32.	0.96	0.90	0.91	0.72
33.	1.0	1.0	0.30	0.30
34.	1.0	1.0	0.72	0.67
35.	1.0	0.97	0.87	0.88

Table Supplement *Continued*

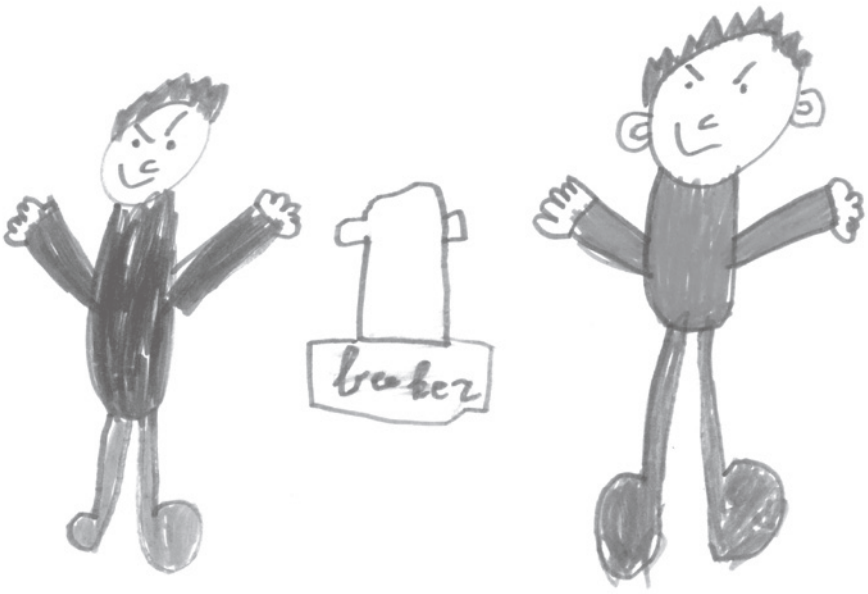
Domain	Interrater (QCM) (N=68) kappa	Interrater (SCM) (N=68) kappa	Test-retest (QCM) (N=35) kappa	Test-retest (SCM) (N=35) kappa
36.	1.0	0.96	0.78	0.86
37.	1.0	1.0	0.72	0.78
38.	0.97	0.97	0.83	0.83
39.	1.0	1.0	0.94	0.94
40.	1.0	1.0	0.80	0.80
41.	0.94	0.94	0.60	0.60
42.	0.94	0.97	0.60	0.66

Note. #=less than 3 times in both sessions; ODD=Oppositional Defiant Disorder; CD=Conduct disorder; ADHD=attention-deficit/hyperactivity disorder; Three CD symptoms (sexually coercive, sets fire and steals with confrontation) were rated as present fewer than 3 times; QCM=Qualitative Coding Method; SCM=Specific Coding Method;
ODD: 1. Often defiant; 2. Often argues with adults; 3. Often loses temper; 4. Often easily annoyed; 5. Often angry or resentful; 6. Often annoys others; 7. Often blaming others; 8. Often spiteful or vindictive
CD: 12. Physical aggression; 13. Often bullies, threatens or intimidates; 14. Uses object to harm; 15. Physically cruelty to people; 16. Physically cruelty to animals 17. Forced sexual activity; 18. Often lies; 19. Stealing without confrontation 20. ; Steals with confrontation; 21. Destroys property; 22. Fire setting
ADHD: 25. Careless/inattention; 26. Difficulty sustaining attention; 27. Doesn't listen; 28. Difficulty following instructions; 29. Difficulty organizing tasks; 30. Difficulty with sustained mental effort; 31. Easily distracted; 32. Often loses things; 33. Often forgetful
34. Fidgets/Squirms; 35. Difficulty remaining seated; 36. Runs/climbs excessively/inappropriately; 37. Difficulty playing quietly; 38. Talks excessively; 39.Often on the go; 40. Blurts out answers; 41. Difficulty waiting turns; 42. Interrupts/intrudes
ADHD: 25. Careless/inattention; 26. Difficulty sustaining attention; 27. Doesn't listen; 28. Difficulty following instructions; 29. Difficulty organizing tasks; 30. Difficulty with sustained mental effort; 31. Easily distracted; 32. Often loses things; 33. Often forgetful
34. Fidgets/Squirms; 35. Difficulty remaining seated; 36. Runs/climbs excessively/inappropriately; 37. Difficulty playing quietly; 38. Talks excessively; 39.Often on the go; 40. Blurts out answers; 41. Difficulty waiting turns; 42. Interrupts/intrudes

CHAPTER 4

Stability and change of ODD, CD and ADHD diagnosis in referred preschool children

Tessa L. Bunte, M.D., Kim Schoemaker, M.S., David J. Hessen, Ph.D., Peter G.M. van der Heijden, Ph.D., Walter Matthys, M.D., Ph.D.



Bunte, T.L., Schoemaker, K., Hessen, D.J., van der Heijden, P.G.M., Matthys, W. (2014). Stability and change of ODD, CD and ADHD diagnosis in referred preschool children. Accepted for publication in the *Journal of Abnormal Child Psychology*.

ABSTRACT

Longitudinal studies have shown that preschool children's diagnosis of Oppositional Defiant Disorder (ODD), Conduct Disorder (CD) and Attention-Deficit/Hyperactivity Disorder (ADHD) are likely to persist into school age. However, limited attention has been paid to instability of diagnosis. Moreover, studies did not examine whether children who were not diagnosed at baseline reached the symptom threshold at later assessments. The aim of the present study, therefore, was to investigate both stability and change of ODD, CD and ADHD diagnosis in children aged 3.5-5.5 years. For diagnosing these disorders, a semi-structured diagnostic parent interview, i.e., the Kiddie-Disruptive Behavior Disorder Schedule (K-DBDS), was used at the first assessment and at follow-up assessments (9 and 18 months). Five diagnostic stability groups (chronic, partial remission, full remission, new onset, no diagnosis) were compared with regard to impairment and number of symptoms. Participants were referred preschool children with externalizing behavioral problems (N=193; 83% male) and typically developing (TD) children (N=58; 71% male). Follow-up assessments allowed to distinguish children belonging to the chronic group of ODD, CD or ADHD from those belonging to one of the remission groups. In addition, there was a substantial number of children with a new onset diagnosis. In conclusion, as a complement to studies showing stability of ODD, CD and ADHD diagnosis into school age, present findings point to changes of diagnosis in the preschool and early school period. Diagnostic reassessments therefore are needed in this age group. In these clinical evaluations attention should be paid both to impairment and number of symptoms.

Keywords: Preschoolers, ODD, CD, ADHD, Assessment, Structured Interview

INTRODUCTION

There is an increasing number of young children referred to mental health clinics (Howell & Teich, 2008; Ringel & Sturm, 2001). Among these children attention deficit and disruptive behavior problems are the most common reason for referral (Wilens et al., 2002**).** Although trajectories of chronic conduct problems and hyperactivity/attention problems can already be identified in the toddler and preschool years (Shaw, Lacourse & Nagin, 2005; Galéra et al., 2011), clinicians have long been reluctant to diagnose young children with Oppositional Defiant Disorder (ODD), Conduct Disorder (CD) and Attention-Deficit/Hyperactivity Disorder (ADHD) (American Psychiatric Association, 2013). Yet, epidemiological studies have demonstrated that the prevalence rates of ODD, CD and ADHD in preschool children are similar to those later in childhood (Egger & Angold, 2006).

Evidence for the stability of ODD, CD and ADHD diagnoses is critical for refuting the claim that these disorders pathologize normative behavior in young children (Keenan et al., 2011). Three studies have shown evidence for the stability of these disorders from preschool to elementary school age, but significant change as well. With regard to ODD and CD, in a study on the predictive validity of ODD and CD in 3-5-year old children Keenan et al., (2011) showed that 73.0%, 66.3% and 51.7% diagnosed with ODD at baseline met the criteria at 12-, 24-, and 36 month follow-up. Of the children who met criteria for CD at baseline, 48.6%, 33.3%, and 26.0% met criteria at 12-, 24-, and 36 month follow-up. With regard to ADHD, Lahey et al., (2004) examined stability of diagnosis in a three year predictive validity study in children diagnosed at 3.8-7 years of age with ADHD. This study found that 79% of the children who met criteria for ADHD at study entry received the diagnosis on at least two of three assessments. Likewise, in a follow-up of the Preschool ADHD Treatment Study (baseline mean age: 4.4 years), the rate of diagnosis irrespective of medications status at 3 year follow-up was 76% and at 6 year follow-up 77.2% (Riddle et al., 2013).

Results of these studies support the predictive validity of ODD, CD and ADHD diagnosis in preschool children, but they also point to instability of diagnosis from preschool to school age. Moreover, the studies did not examine whether children who were not diagnosed at baseline reached the symptom threshold at later assessments possibly due to increasing environmental expectations over the preschool and early school period. The aim of the present study, therefore, was to investigate both stability and change of ODD, CD and ADHD diagnosis in preschool children within a smaller time frame than in previous studies. To examine the stability and change in ODD, CD and ADHD diagnosis in preschoolers, an important issue is how to differentiate clinical from normative transient disruptive behaviors within the preschool age, as most preschoolers exhibit at least some of the behaviors that fall under the rubric of disruptive behavior. This makes clinical assessment

of preschool behavioral problems particularly subtle and complex (Wakschlag et al., 2005). Recently, however, assessment procedures have been developed, including (semi-) structured interviews with the parents such as the Kiddie Disruptive Behavior Disorder Schedule (K-DBDS; Keenan et al., 2007, Bunte et al., 2013b) and the Preschool Age Psychiatric Assessment (PAPA; Egger et al., 2006), and standardized observational assessments such as the Disruptive Behavior Diagnostic Observation Schedule (DB-DOS; Wakschlag et al., 2008ab, Bunte et al., 2013a).

The K-DBDS has been shown to have satisfactory interrater and test-retest reliability with regard to ODD/CD symptoms (Keenan et al., 2007) and ADHD symptoms (Bunte et al., 2013b) at the preschool age. Furthermore, Bunte et al. (2013b) have investigated the clinical usefulness of the K-DBDS in a European sample of 193 referred preschool children with externalizing behavior problems and 58 typically developing children. Two issues with regard to the operationalization of ODD and ADHD symptoms were considered using different coding methods: 1. whether the definition of ODD symptoms should be based either on the threshold "often" or on a more specific frequency threshold (e.g., two or three times a day; Egger & Angold, 2006); 2. whether in the definition of ADHD symptoms pervasiveness across contexts should be included or not. It was shown that ODD in preschool children may be diagnosed using a cutoff score of four ODD symptoms with "often" as a threshold for frequency of behavior. For ADHD, a threshold of five hyperactive-impulsive symptoms was satisfactory without inclusion of pervasiveness across contexts (Bunte et al., 2013b).

A limitation of the Bunte et al. study was that the operationalization of the ADHD predominantly inattentive type could not be examined due to low numbers of children diagnosed with this subtype. Indeed, according to consensus best estimate diagnosis only six children were diagnosed with the predominantly inattentive type of ADHD. ADHD subtypes (predominantly hyperactive/impulsive, predominantly inattentive, and combined subtype), however, may not be the best descriptors of the disorder in the preschool age group. Indeed, the subtypes seem too unstable for use in the clinical assessment of individual children. Preschool children rarely remain in the ADHD predominantly hyperactive-impulsivesubtype (ADHD HI) over time; rather, they sometimes desist from ADHD but mostly shift to the combined subtype in later years (Lahey et al., 2005). This may be related, in part, to a decline in hyperactive-impulsivesymptoms and an increase in inattention symptoms with age, resulting in a later diagnosis of combined subtype for the majority of those initially classified as hyperactive-impulsivesubtype (Hardy et al., 2007). Consistent with these findings, the base rate of inattentive symptoms in clinically referred preschoolers has been found to be lower and more variable than the base rate of hyperactive-impulsivesymptoms (Murray et al., 2007). Finally, in a sample of 1155 preschool children it was shown that ADHD symptomatology was best represented by a

single latent factor that exhibited partial measurement invariance from 3-5 years of age (Willoughby, Pek & Greenberg, 2012).

The present study extends the Bunte et al. (2013b) study by reassessing the 3.5–5.5 year old children 9 months and 18 months after the first assessment. The first aim of the study was to examine stability and change of ODD, CD and ADHD HI diagnosis in terms of percentages of children being diagnosed again at 9 month and 18 month follow-up. We also anticipated that referred preschool children with externalizing behavior problems not diagnosed at the first assessment may be diagnosed at 9 month and/or 18 month follow-up. The second aim was to examine the predictive validity of an ODD, CD and ADHD HI diagnosis. The third aim was to examine impairment and number of symptoms in groups that differed in stability of diagnosis (chronic, partial remission, full remission, new onset, no diagnosis).

METHOD

Participants

Participants were 251 children aged 42-66 months (mean=55 months, SD=7.8); 80% were male, 86% White, 12% Turkish/Moroccans, 2% African, 0.5% Asian. In terms of educational level of the parents: primary accounted for 6%, secondary 25%, intermediate vocational 34%, higher vocational 16% and university 19%. General practitioners, pediatricians and well-baby clinics in the province of Utrecht were invited to refer children aged 3.5-5.5 with externalizing behavior problems to the Outpatient Clinic for Preschool Children with Behavioral Problems, Department of Psychiatry, University Medical Center, Utrecht, the Netherlands. Children were included if they scored at or above the 90-th percentile either on the Aggressive Behavior scale or on the Attention Problems scale of the Child Behavior Checklist (CBCL/1.5-5) or the Child-Teacher Report Form (C-TRF/1.5-5) (Achenbach & Rescorla, 2000). The typically developing (TD) group (N=58) was recruited from regular elementary schools and daycare centers. Children with a score in the clinical range either on the Attention Problems scale or on the Aggressive Behavior scale of the CBCL/1.5-5 or C-TRF/1.5-5 were excluded. All the children (referred and TD) with an IQ below 70, estimated with the average score on the Raven Coloured Progressive Matrices (Raven, Court, & Raven, 1998) and the Peabody Picture Vocabulary Test-III-NL (Dunn & Dunn, 2005; Schlichting, 2005), and children diagnosed with Autistic Disorder or Pervasive Developmental Disorder-Not Otherwise Specified (PDD-NOS), based on best estimate diagnosis, were excluded. In the study 193 of the 208 referred children were included (15 children were excluded because of an IQ below 70, Autistic Disorder/PDD-NOS diagnosis and in one case refusal of the parents to use the data of their child for the study) of whom

154 were diagnosed with ODD, CD or ADHD. None of the participants were on medication. Children were followed for a period of 9 months (Mean=8.9 months, range=7-12 months, SD=0.88 months) and 18 months (Mean=18.1 months, range=15-24 months, SD=1.2 months). In view of examining the stability of ODD, CD and ADHD diagnosis, a diagnostic evaluation was conducted for each child at the first assessment (T1) and the follow-up assessments 9 months (T2) and 18 months later (T3) based on the K-DBDS. A child was placed in the ODD, CD or ADHD HI group, and if there was comorbidity, also in the comorbidity group (ODD+ADHD or CD+ADHD). To examine stability and change of ODD, CD and ADHD HI diagnosis, mutually exclusive diagnostic stability groups were created. Children were categorized as: 1. no diagnosis, i.e., no diagnosis at T1, T2 and T3; 2. chronic diagnosis, i.e., meeting diagnostic criteria at T1, T2 and T3; 3. full remission, i.e., a diagnosis at T1 and no diagnosis at T2 and T3; 4. partial remission, i.e., a diagnosis at T1 and no diagnosis at either T2 or T3; 5. new onset, i.e., no diagnosis at T1 but a diagnosis either at T2 or T3. Several children received treatment after the first assessment: individual parent /counseling at home or at the outpatient clinic, the Incredible Years parent program (Webster-Stratton & Reid, 2010), and/or pharmacotherapy (methylphenidate in most cases, but atomoxetine and risperidone were prescribed as well). Between T1 and T2 128 (66%) and between T1 and T3 137 (71%) of the referred preschool children have received treatment.

Procedure

At each assessment children were evaluated in a single morning session. At T1, two measures of intellectual functioning were administered followed by executive functions (EF) tasks (Schoemaker et al., 2012). After a break the child was observed using the DB-DOS (Wakschlag et al., 2008ab; Bunte et al, 2013a) and the K-DBDS was administered. At T2, EF tasks and the K-DBDS were administered. At T3, the assessment was similar to T1 with the exception of the assessment of intellectual functioning. The K-DBDS at T1 and T3 was administrated by the same child psychiatrist/clinical child psychologist while the K-DBDS at T2 was administered by trained research assistants; the research assistants were clinical child psychologists, who were trained by the child psychiatrists involved in the study. Retention was high: 96% at T2 and 95% at T3. There were no significant differences between those participants who were and were not seen at T2 and T3 in terms of age, sex, referral status or diagnosis at T1.

Parents received a small financial compensation for participating, children received two small gifts. Written informed consent from the parents was gathered before participating. The Medical Ethical Review Committee of the University Medical Center Utrecht, approved this study.

Measures

K-DBDS; the K-DBDS is a semi-structured clinical interview assessing DSM-IV (American Psychiatric Association, 2000) ODD, CD and ADHD symptoms in preschool children (Keenan et al., 2007). It is based on the Schedule for Affective Disorders and Schizophrenia for School-Age Children (Orvaschel & Puig-Antich, 1995). Administrating the K-DBDS takes approximately 45 minutes, coding takes approximately ten minutes. To determine if the behavior is a symptom according to the K-DBDS, the Qualitative Coding Method (QCM) was used (K-DBDS Preschool Version Instruction Manual and Bunte et al., 2013b). For ODD: 1. the threshold "often" is used for frequency of behaviors. For CD: the coding methods is in line with those for ODD, although there are only a few CD symptoms with the word "often" in the symptom description. For the other CD symptoms, according to the QCM the occurrence of behavior is sufficient to be considered a symptom. For ADHD HI: according to the QCM the threshold "often" is used for frequency of behaviors; the threshold of six symptoms is used, and the pervasiveness threshold that the behavior must be endorsed in more than one context is not included.

CBCL/C-TRF. Disruptive behavior symptoms and attention problems were assessed using the scores of the Attention Problem scale and the Aggressive Behavior Scale of the Child Behavior Checklist (Achenbach & Rescorla, 2000) (CBCL/1.5-5) completed by parents and the Child-Teacher Report Form (Achenbach & Rescorla, 2000) (C-TRF/1.5-5) completed by teachers or daycare caregivers;

Children's Global Assessment Scale (CGAS; Schaffer et al., 1983). To provide a global assessment of functioning, parents completed the non-clinician version of the CGAS. The CGAS is scored from 0-100 with lower scores indicating greater impairment.

Statistical analyses

For all scales used, the percentage of missing data for the three different assessments was: 1.6-2.7% for the K-DBDS (this percentage represented the incomplete questionnaires), 0.8-1.6% for the CGAS parent questionnaire. Children were excluded from the analyses if there had missing data in these analyses.

First, in order to examine the stability and change of ODD, CD and ADHD HI diagnosis in terms of percentages of children being again diagnosed at T2 and T3, we used descriptive statistics (frequencies) for the rates at T1, T2 and T3. Second, to study the predictive validity of an ODD, CD, ADHD HI, or comorbid diagnosis (ODD+ADHD or CD+ADHD), logistic regression analyses were used with diagnosis at T1, and sex and treatment entered simultaneously to estimate conditional odds ratios. We controlled for treatment as some children were treated and others not. Information was collected on: Type of treatment: individual parent counseling at home (yes/no) /at the clinic (yes/no), the Incredible Years parent program (yes/no), and pharmacotherapy (yes/no); however, the total numbers in

each of the treatment separately were small and therefore we constructed the variable "Treatment", as a dichotomous variable (yes/no) where all types of treatment are taken together. Third, general linear model analyses (mixed repeated measures ANOVA) were carried out, namely for (i) impairment over time and (ii) number of symptoms over time, in groups that differed in stability of diagnosis (chronic, partial remission, full remission, new onset, no diagnosis); post hoc comparisons were corrected using Bonferonni. Fourth, we conducted ANOVA analyses, at T1, T2 and T3 with the chronic group compared to the full and partial remission groups. To examine whether the diagnosis groups differed in sex and treatment between T1 and T2, and between T1 and T3, Pearson Chi-Square tests were carried out.

RESULTS

Rates of ODD, CD and ADHD HI at T1. Among the 251 preschool children 142 met criteria for ODD with a threshold of four symptoms (96.5% of them were from the clinically referred group) (45.1% also met CD and 52.1% of them also met ADHD HI criteria) and 79 preschoolers met criteria for CD with a threshold of three symptoms (98.7% clinically referred) (81.0% of them also met ODD and 59.5% of them also met ADHD HI criteria), see also Figure 1. Among the 251 preschool children 94 children met criteria of ADHD HI with a threshold of six symptoms (98.9% clinically referred) (50.0% of them also met CD and 43.6% of them also met ODD criteria), see also Figure 1.

Stability and change of diagnosis at T2 and T3. Figure 1 shows numbers of children with diagnosis at T1, T2 and T3 as well as numbers of children with symptoms beneath diagnostic thresholds. For example, for ODD diagnosis at T1 142 children met ODD criteria. Of these children diagnosed at T1, 59.2% of the children diagnosed at T1 with ODD, still met ODD criteria at T2 and 62.0% of the children diagnosed at T1 with ODD still met ODD criteria at T3. Very few children diagnosed at T1 were symptom free at T2 and T3. Six children (4.2%), who were in remission of ODD at T2 or T3, developed a new onset of CD on T2 or T3. Five children (3.5%), who were in remission of CD at T2 or T3, developed a new onset of ODD on T2 or T3. These were all referred children.
Children were categorized as having either 1. no diagnosis , i.e., no diagnosis at T1, T2 and T3; 2. chronic diagnosis, i.e., meeting diagnostic criteria at T1, T2 and T3; 3. full remission, i.e., a diagnosis at T1 and no diagnosis at T2 and T3; 4. partial remission, i.e., a diagnosis at T1 and no diagnosis at either T2 or T3; 5. new onset, i.e., no diagnosis at T1 but a diagnosis either at T2 or T3. For ODD, CD, ADHD HI, ODD+ADHD HI and CD+ADHD HI these amounts and percentages are shown in Table 1. The "No diagnosis" group is not the same as the

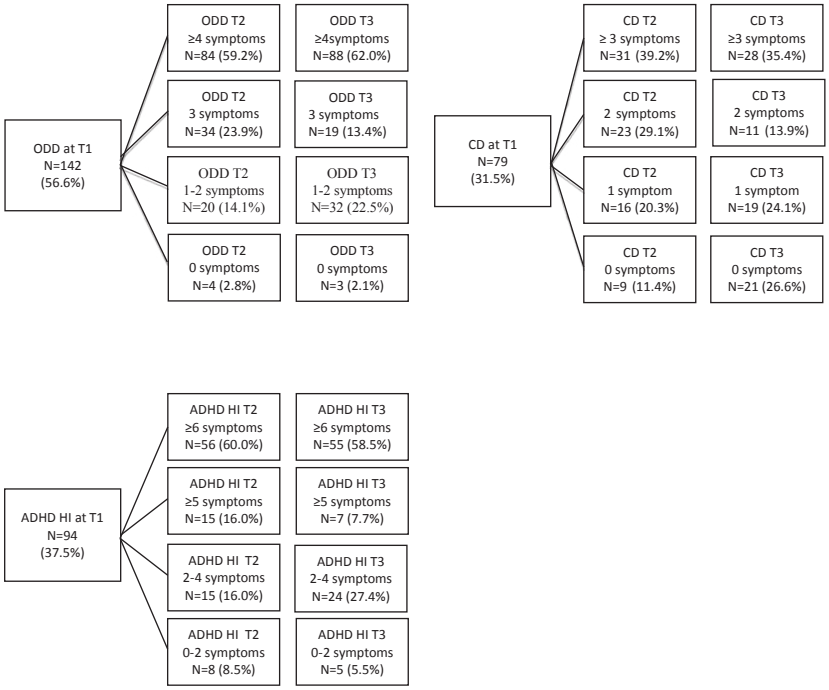


Figure 1. Outcome for preschoolers meeting criteria for DSM IV ODD, CD and ADHD at ages 3.5-5.5 years. T1=Baseline; T2=9 months follow-up; T3=18 months follow-up; ODD=Oppositional Defiant Disorder; CD=Conduct Disorder; ADHD HI=attention-deficit/hyperactivity disorder predominantly hyperactive-impulsive type.

Table 1 *Number (N) and percentages of the five different stability groups for ODD, CD, ADHD HI, ODD+ADHD HI and CD+ADHD HI*

	No diagnosis	Chronic	Full remission	Partial remission	New onset	Total
ODD	79 (35%)	64 (28%)	26 (11%)	36 (16%)	23 (10%)	228 (100%)
CD	131 (57%)	17 (7%)	31 (14%)	24 (10%)	27 (12%)	230 (100%)
ADHD HI	117 (50%)	42 (18%)	24 (10%)	26 (11%)	26 (11%)	235 (100%)
ODD+ADHD HI	141 (62%)	26 (11%)	20 (9%)	23 (10%)	18 (8%)	228 (100%)
CD+ADHD HI	166 (72%)	8 (3%)	21 (9%)	14 (6%)	21 (9%)	230 (100%)

Note: ODD=Oppositional Defiant Disorder; CD=Conduct Disorder; ADHD HI=attention-deficit/hyperactivity disorder predominantly hyperactivity-impulsivity type.

TD group; for example, the "no ODD diagnosis group" consisted of TD children but also children with an ADHD diagnosis.
We examined how many preschool children without a diagnosis at T1 were given a diagnosis at T2 or T3. There were 23 (9.2%) new cases of ODD (83% referred): 60.1% had three symptoms at T1, 8.6% had two symptoms at T1, 21.7% had one symptom at T1, and 8.6% had no symptoms at T1. There were 27 (10.8%) new cases of CD (100% referred): half of them had two symptoms at T1, nearly half of them had one symptom at T1, and 2.7% had no CD symptoms at T1. There were 26 (10.4%) (93% referred children) new cases of ADHD HI using the six symptom threshold: 51.8% had five symptoms, 35.1% had four symptoms at T1, 10,8% had three symptoms at T1, 2.7% had two symptoms at T1, and no new cases with one or no symptom at T1.

Predictive validity of preschool diagnosis. The likelihood of meeting criteria for a diagnosis of ODD, CD and ADHD HI at a subsequent assessment was tested using logistic regression with diagnosis at T1, and sex and treatment entered simultaneously. Sex had no significant effect; given the small percentage of girls (17%), these findings cannot be applied for girls and therefore was dropped from the models. The odds ratio (95% CI) of meeting criteria for ODD, given a diagnosis at T1, were for T2 and T3 9.92 (4.79-20.59) and 6.06 (3.17-11.60), respectively, controlled for treatment (see Table 2; all odds ratios are conditional on the presence of the other variables in the model). The odds ratio (95% CI) of the covariate treatment was 3.16 (1.62-6.16) and 4.04 (2.10-7.76) for T2 and T3 respectively. Conditional odds ratios for CD, ADHD HI, ODD+ ADHD and CD+ADHD are shown in Table 2. We also conducted logistic regression analyses without controlling for treatment, and this resulted in higher odds ratios for ODD, CD, ADHD HI and ODD+ADHD HI and CD+ADHD HI. For example, the odds ratio (95% CI) of ODD at T3 controlling for treatment was 6.06 (3.17-11.60) while it was 8.95 (4.84-16.56) not controlling for treatment. Likewise, the odds ratio of CD at T3 was 3.94 (1.95-7.98) controlling for treatment while it was 5.02 (2.54-9.89) not controlling for treatment.

Impairment as function of diagnostic stability of ODD, CD, ADHD, ODD+ADHD or CD+ADHD.

Figure 2 shows the results of average ratings of impairment over time for the coding methods for ODD, CD, ADHD HI, ODD+ADHD HI and CD+ADHD HI. General linear model analyses (mixed repeated measures ANOVA) with post-hoc comparisons using Bonferroni correction were conducted to test, first, whether the five groups differed from each other over time (see Table 3), and second, using ANOVA, whether the chronic group differed from the full- and partial remission group at T1, T2 and T3 (see Table 4).

Table 2 *Predictive Validity of Preschool ODD, CD, ADHD, ODD+ADHD and CD+ADHD across 18 months*

	B	S.E.	Sign.	Lower OR	95% CI for OR	Upper OR
ODD T2						
ODD T1	2.30	.37	.000	4.79	9.92	20.54
Treatment	1.15	.34	.001	1.62	3.16	6.16
ODD T3						
ODD T1	1.80	.33	.000	3.17	6.06	11.60
Treatment	1.40	.33	.000	2.10	4.04	7.76
CD T2						
CD T1	1.88	.37	.000	3.16	6.47	13.22
Treatment	.70	.39	.069	.95	2.02	4.29
CD T3						
CD T1	1.37	.36	.000	1.95	3.94	7.98
Treatment	1.61	.47	.001	1.98	5.00	12.64
ADHD HI T2						
ADHD HI T1	2.55	.37	.000	6.20	12.79	26.38
Treatment	.50	.37	.17	.80	1.65	3.37
ADHD HI T3						
ADHD HI T1	1.71	.33	.000	2.92	5.47	10.52
Treatment	1.46	.38	.000	2.07	4.32	9.02
ODD+ADHD T2						
ODD+ADHD T1	2.12	.41	.000	3.72	8.35	18.20
Treatment	.93	.46	.043	.93	2.53	6.22
ODD+ADHD T3						
ODD+ADHD T1	2.00	.36	.000	3.62	7.36	14.93
Treatment	1.29	.44	.004	1.52	3.62	8.60
CD+ADHD T2						
CD+ADHD T1	2.13	.47	.000	1.83	8.41	9.98
Treatment	.32	.49	.51	1.35	1.38	12.57
CD+ADHD T3						
CD+ADHD T1	1.45	.43	.001	3.37	4.27	21.00
Treatment	1.42	.57	.013	.53	4.12	3.59

Note. S.E.=Standard Estimate; Sign.=Significance; OR=Odds Ratio; CI=Confidence Interval; ODD=Oppositional Defiant Disorder; CD=Conduct Disorder; ADHD HI=attention-deficit/hyperactivity disorder predominantly hyperactivity-impulsive type; T1=Baseline; T2=9 months follow-up; T3=18 months follow-up.

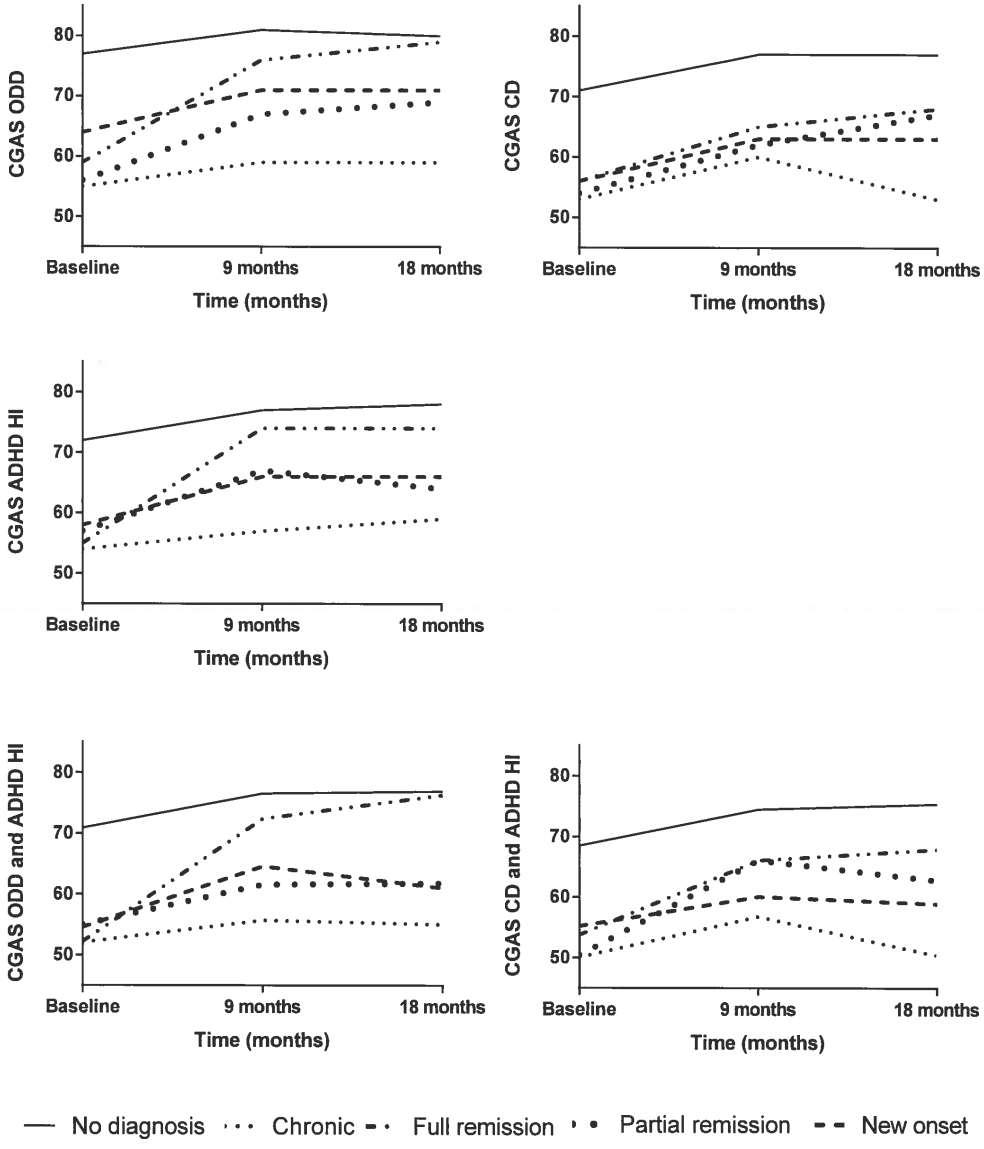


Figure 2. Average rating of impairment using CGAS as a function of diagnostic status.
CGAS=Children's Global Assessment Scale (parent); ODD=Oppositional Defiant Disorder; CD=Conduct Disorder; ADHD HI=attention-deficit/hyperactivity disorder predominantly hyperactive-impulsive type.

Table 3 *General Lineair Model Impairment (CGAS) and Total symptom score ODD, CD, ADHD HI, ODD+ADHD HI and CD+ADHD HI*

	Sum of Squares	F	Sign.
ODD CGAS			
Between groups	58474.19	45.73	.00
CD CGAS			
Between groups	41719.02	26.53	.00
ADHD HI CGAS			
Between groups	42645.25	27.40	.00
ODD+ADHD HI CGAS			
Between groups	40112.59	24.88	.00
CD+ADHD HI CGAS			
Between groups	29447.96	16.50	.00
ODD tot. symptom			
Between groups	2475.63	288.06	.00
CD tot. symptom			
Between groups	967.20	236.27	.00
ADHD HI tot. symptom			
Between groups	3348.42	203.19	.00

Note. Degrees of freedom=4; CGAS=Children's Global Assessment Scale; ODD=Oppositional Defiant Disorder; CD=Conduct Disorder; ADHD HI=attention-deficit/hyperactivity disorder predominantly hyperactivity-impulsive type; F=F value; Sign.=significance; tot=total.

ODD. For ODD, these analyses revealed significant differences between the five groups (see Table 3) with regard to impairment. In order to examine at which assessment the chronic group differed from the remission groups, we compared the chronic group with the full- en partial remission group at T1, T2 and T3. The chronic group did not significantly differ from the full- and partial remission group at T1, but both at T2 and T3 the chronic group differed significantly from these groups, see also Table 4.

CD. For CD, the analyses revealed significant differences between the five groups (see Table 3) with regard to impairment. We examined whether the chronic group differed from the remission groups at T1, T2 and T3 with regard to impairment. At T1 and T2, the chronic group did not differ significantly from the full- and partial remission group, while at T3 the chronic group differed significantly from the full remission and partial remission group on the impairment rating, see also Table 4.

ADHD. For ADHD HI, the analyses revealed significant differences between the five groups (see Table 3) with regard to impairment. We examined whether the chronic group differed from the remission groups at T1, T2 and T3 with regard to impairment. For ADHD HI, the chronic group did not significantly differ from the remission groups at T1, but both at T2

Table 4 *ANOVA Impairment (CGAS)*

	Mean Square	F	Sign.
ODD T1			
Chron- Full rem.	377.32	33.53	.07
Chron- Part. rem.	57.72	.67	.42
ODD T2			
Chron- Full rem.	5162.09	37.02	.00
Chron- Part. rem.	1372.39	9.81	.00
ODD T3			
Chron- Full rem.	7681.80	50.82	.00
Chron- Part. rem.	2618.94	18.01	.00
CD T1			
Chron- Full rem.	88.82	.54	.47
Chron- Part. rem.	16.71	.21	.65
CD T2			
Chron- Full rem.	348.38	1.83	.18
Chron- Part. rem.	70.00	.47	.50
CD T3			
Chron- Full rem.	2514.08	16.64	.00
Chron- Part. rem.	2093.22	11.89	.00
ADHD T1			
Chron- Full rem.	16.61	.19	.67
Chron- Part. rem.	183.02	1.25	.27
ADHD T2			
Chron- Full rem.	4398.45	29.92	.00
Chron- Part. rem.	1832.05	11.62	.00
ADHD T3			
Chron- Full rem.	4248.49	25.82	.00
Chron- Part. rem.	482.48	2.99	.09
ODD+ADHD T1			
Chron- Full rem.	.71	.003	.96
Chron- Part. rem.	157.11	1.37	.25
ODD+ADHD T2			
Chron- Full rem.	3136.72	21.08	.00
Chron- Part. rem.	470.65	3.33	.07
ODD+ADHD T3			
Chron- Full rem.	4867.23	29.58	.00
Chron- Part. rem.	481.71	3.20	.08
CD+ADHD T1			
Chron- Full rem.	78.65	.43	.52
Chron- Part. rem.	.47	.006	.93
CD+ADHD T2			
Chron- Full rem.	511.10	2.51	.13
Chron- Part. rem.	36.53	.17	.69
CD+ADHD T3			
Chron- Full rem.	1770.52	9.48	.01
Chron- Part. rem.	784.13	2.84	.11

Note. Degrees of freedom=1; CGAS=Children's Global Assessment Scale; ODD=Oppositional Defiant Disorder; CD=Conduct Disorder; ADHD HI=attention-deficit/hyperactivity disorder predominantly hyperactivity-impulsive subtype; Chron-Full rem.=Chronic group compared to Full remission group; Chron-Part. rem.=Chronic group compared to Partial Remission group; F=F value; Sign=Significance.

and T3 the chronic group differed significantly from the full remission group, and from the partial remission group at T2 as well, see Table 4.

ODD+ADHD. For ODD+ADHD HI, the analyses revealed significant differences between the five groups (see Table 3) with regard to impairment. In order to examine at which assessment the chronic group differed group from the remission groups, we compared the chronic group with the remission groups at T1, T2 and T3, see also Table 4. The chronic group did not significantly differ from the full- and partial remission groups at T1, but both at T2 and T3 the chronic group differed significantly from the full remission group. The chronic group did not differ significantly from the partial remission group.

CD+ADHD. For CD+ADHD HI, the analyses revealed significant differences between the five groups (see Table 3) with regard to impairment. We examined whether the chronic group differed from the remission groups at T1, T2 and T3, see Table 4. The chronic group did not significantly differ from the full- and partial remission groups at T1 and T2, but at T3 the chronic group differed significantly from the full remission group.

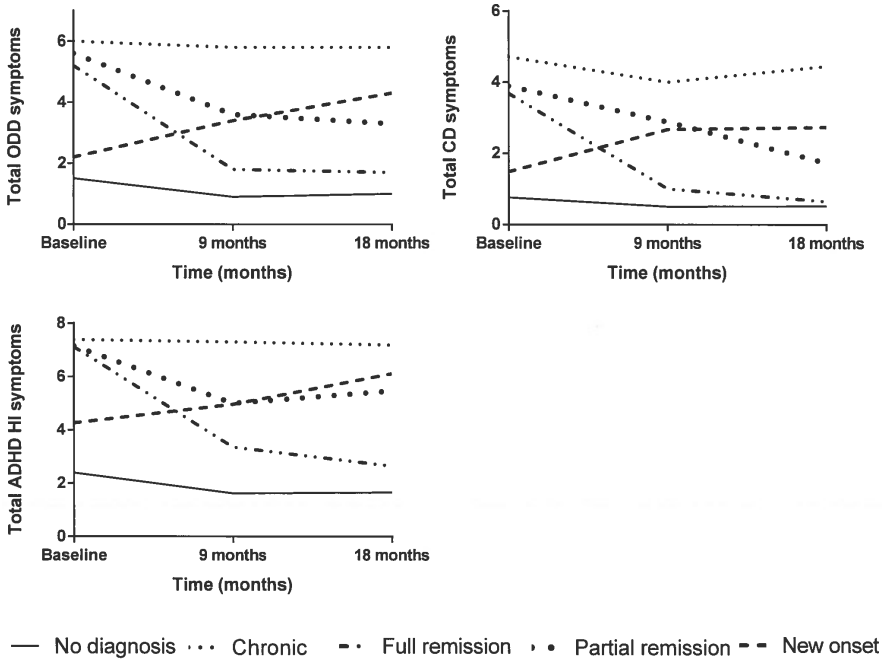


Figure 3. Average rating of total symptom score (ODD, CD, ADHD HI) as a function of diagnostic status. ODD=Oppositional Defiant Disorder; CD=Conduct disorder; ADHD HI=attention-deficit/hyperactivity disorder predominantly hyperactive-impulsive type.

Total symptom score as function of diagnostic stability of ODD, CD and ADHD. Figure 3 shows the results of average total symptoms over time for ODD, CD, and ADHD HI. To examine stability and change, the same mutually exclusive diagnostic stability groups were created as in Figure 2 to test differences for the total symptom scores for ODD, CD and ADHD HI. General linear model analyses (mixed repeated measures ANOVA) with post-hoc comparisons using Bonferroni correction were conducted in a way identical to the examination of impairment, see Table 3. ANOVA analyses, at T1, T2 and T3 with the chronic group compared to the full and partial remission groups, were also examined, see also Table 5.

ODD. For ODD, the analyses revealed significant differences between the five groups (see Table 3) with regard to number of symptoms. In order to examine at which assessment the chronic group differed from the remission groups, we compared the chronic group with the remission groups at T1, T2 and T3, see Table 5. At T1 the chronic group differed significantly from the full remission group, but not from the partial remission group. At T2 and T3 the chronic group differed significantly from the remission groups.

CD. For CD, the analyses revealed significant differences between the five groups (see Table 3) with regard to the number of symptoms. In order to examine at which assessment the chronic group differed from the remission groups, we compared the chronic group with the remission groups at T1, T2 and T3, see Table 5. There were significant differences between the chronic group and the remission groups at T1, at T2 and T3.

ADHD. For ADHD HI, analyses revealed significant differences between the five groups (see Table 3) with regard to number of symptoms. In order to examine at which assessment the chronic group differed from the remission groups, we compared the chronic group with the remission groups at T1, T2 and T3. There were no significant differences between the chronic and the remission groups at T1, but at T2 and T3 differences were significant between the chronic group and the full- and partial remission group, see Table 5.

To examine whether the chronic group differed from the remission groups (partial remission and full remission) in sex and treatment between T1 and T2, and between T1 and T3, Pearson Chi-Square tests were carried out. There were no significant differences in sex and treatment between the chronic group and the remission group for ODD, with the exception of a significant difference in treatment between T1 and T3 for ODD (Sex: Pearson X2(1)=.12, *p*=.73; Treatment T1-T2: X2(1)=3.34, *p*=.07; Treatment T1-T3: X2(1)=5.43, p=.02), for CD (Sex: X2(1)=1.06, *p*=.30; Treatment T1-T2: X2(1)=2.75, *p*=.10; Treatment T1-T3: X2(1)=1.45, *p*=.23), and for ADHD HI (Sex: X2(1).02, *p*=.90; Treatment T1-T2: X2(1)=.46, *p*=.50; Treatment T1-T3: X2(1)=1.26, *p*=.26)[a].

[a] Results of percentages as described in Figure 1, and the logistic regression analyses, are similar for the boys only sample and the boys and girls sample. Results of general linear model analyses (mixed repeated measures ANOVA) with post-hoc comparisons using Bonferroni correction and ANOVA's were similar for the boys only sample and the sample with boys and girls with regard to impairment and total symptom score.

Table 5 *ANOVA Total symptom score*

	Mean Square	F	Sign.
ODD T1			
Chron- Full rem.	11.21	7.43	.01
Chron- Part. rem.	3.10	2.26	.14
ODD T2			
Chron- Full rem.	299.44	193.92	.00
Chron- Part. rem.	110.94	49.43	.00
ODD T3			
Chron- Full rem.	310.03	204.70	.00
Chron- Part. rem.	148.77	74.57	.00
CD T1			
Chron- Full rem.	13.46	12.22	.00
Chron- Part. rem.	8.05	4.46	.04
CD T2			
Chron- Full rem.	102.49	100.36	.00
Chron- Part. rem.	13.13	7.21	.01
CD T3			
Chron- Full rem.	175.99	172.31	.00
Chron- Part. rem.	83.59	61.23	.00
ADHD T1			
Chron- Full rem.	1.43	1.27	.26
Chron- Part. rem.	.95	.73	.40
ADHD T2			
Chron- Full rem.	233.64	146.87	.00
Chron- Part. rem.	88.17	26.09	.00
ADHD T3			
Chron- Full rem.	318.34	216.51	.00
Chron- Part. rem.	48.00	23.83	.00

Note. Degrees of freedom=1; ODD=Oppositional Defiant Disorder; CD=Conduct Disorder; ADHD HI=attention-deficit/hyperactivity disorder predominantly hyperactivity-impulsivity subtype; Chron-Full rem.=Chronic group compared to Full remission group; Chron-Part. rem.=Chronic group compared to Partial Remission group; df=; F=F value; Sign.=Significance

DISCUSSION

The aim of the study was to examine stability and change of ODD, CD and ADHD diagnosis in the context of functional impairment and symptom severity in the preschool and early school period. From the perspective of stability, 62% of the children who were diagnosed with ODD at the first assessment were diagnosed again at 18 month follow-up; 35% of the children who were diagnosed with CD at the first assessment were diagnosed again at 18 months follow-up; 59% of the children who were diagnosed with ADHD HI at the first assessment were diagnosed again at 18 month follow-up. Cautiousness is needed when comparing these findings with those from other studies, such as the one by Keenan et

al. (2011) in which the same assessment method was used. For example, percentages of stability for ODD in the present study were lower than in the Keenan et al., study (2011). Likewise, in the present study the odds ratio for ODD was 8.95 at 18 month follow-up while in the US study the odds ratio was 16.31 at 12 month follow-up. Differences in stability and predictive validity between the two studies may be due to differences in SES; the sample of the US study indeed was urban and had lower SES than the sample of the present study. The study findings show that ODD, CD and ADHD are disorders that show stability but change as well in the preschool to early school period.

Stability and change were studied, first, from the perspective of the identification of children with a stable diagnosis, second, from the perspective of the identification of new cases. With regard to the first issue, mutually exclusive diagnostic stability groups were created as was done in the Keenan et al. study (2011) in order to distinguish children with a chronic trajectory from those with a remission trajectory (partial remission, full remission). Therefore, the chronic group was compared with the remission groups at the first assessment, at 9 month and 18 month follow-up. In the Keenan et al. study (2011) stability and change of ODD and CD diagnosis were examined from the perspective of impairment. In the present study we added the perspective of number of symptoms.

For all disorders, both from the perspective of impairment and number of symptoms the chronic group could be distinguished from the remission groups over time. In addition, for ODD only at 9 and 18 month follow-up did the chronic group differ from both remission groups in impairment and number of symptoms. The study results thus suggest that over time differences between the chronic group and the remission groups became more manifest. Diagnostic reassessments, therefore, were needed in order to identify children belonging to the chronic group of ODD. The addition of the perspective of number of symptoms to identify the chronic group did not appear not be strictly necessary.

For CD, only at 18 month follow-up did the chronic group differ from both remission groups in impairment. From the perspective of number of symptoms, however, the chronic group differed from the remission groups at all assessments. Thus, the chronic group already could be identified at the first assessment. These results indicate that a high number of CD symptoms at the first assessment was suggestive for a disorder that is likely to become chronic. The addition of the perspective of number of symptoms to identify the chronic group therefore was useful.

For ADHD HI, the chronic group differed from the full remission group in impairment at 9 and 18 month. Differences between the chronic group and the partial remission groups were less clear. On the contrary, results regarding number of symptoms showed that the chronic group differed from both remission groups at 9 and 18 month follow-up. The results thus suggest that over time differences between the chronic group and the remission groups became more manifest and that diagnostic reassessments using impairment or, even better, number of symptoms were needed in order to identify children belonging

to the chronic group of ADHD HI. Thus, the addition of the perspective of number of symptoms to identify the chronic group was useful.

It was not feasible to add the perspective of symptoms when studying the comorbid groups for there was no rationale for just adding up symptoms of these distinct disorders. For ODD+ADHD HI, from the perspective of impairment only at 9 and 18 month follow-up did the chronic group differ from the full remission group. Likewise, for CD+ADHD HI only at 18 month follow-up did the chronic group differ from the full remission group. The results regarding the comorbid groups thus suggest that over time differences between the chronic group and the full remission group became more manifest and that diagnostic reassessments were needed in order to distinguish children belonging to the chronic group from those belonging to the full remission group.

With regard to the perspective of the identification of children with a stable diagnosis results of the study thus suggest, first, that for both CD and ADHD HI the perspective of number of symptoms was necessary to distinguish the chronic group from the remission groups. Second, for ODD, ADHD HI and the comorbid groups diagnostic reassessments were needed in order to identify children belonging to the chronic group. Instead, for CD the chronic group already could be identified at the first assessment based on a high number of symptoms. The present study adds to the Keenan et al., study (2011) in bringing to clinicians' attention that consideration of number of symptoms is an important perspective in view of identifying children belonging to the chronic group of ODD, CD and ADHD HI, as a supplement to the perspective of impairment.

With regard to the identification of new cases, the present study shows that a substantial number of children (23) without a diagnosis of ODD at the first assessment were diagnosed with this disorder either at 9 or 18 month follow up; many of them had three symptoms at the first assessment. Likewise, a substantial number of new cases of CD (27), of ADHD HI (26), of ODD+ADHD HI (18) and of CD+ADHD (21) were identified at 9 or 18 month follow-up. Although at a group level symptoms of ODD, CD and ADHD decrease over time, at an individual level a child referred for externalizing behavior problems not diagnosed with a disorder at the first assessment may show an increase of symptoms and be diagnosed later on, possibly as a result of increasing environmental expectations over the preschool period. The study thus shows that diagnostic reassessments of preschool children referred for externalizing behavior problems are needed in order to identify new cases. At a practical level, when a preschool child referred for externalizing behavior problems is not diagnosed with a disorder, parents may be advised either psychosocial intervention (Comer, Chow, Chan, Cooper-Vince, & Wilson, 2013) or to get in touch with the clinic once behavior problems or associated impairments increase; a diagnostic reassessment may then be planned. The number of children diagnosed with CD (79) may seem high in a referred sample of 193 preschool children and 58 TD children. However, the rate of CD to ODD diagnosis

in the present study (79/142=56%) was quite lower than the rate (76/89=85%) in the Keenan et al., study (2011) in which the same assessment procedure was used. Yet in the clinical validity study of the K-DBDS only one child was diagnosed with CD according to clinical consensus diagnosis (Bunte et al., 2013b). This is quite different from the number of children (79) diagnosed according to the K-DBDS. On the one side, clinicians may be too reluctant to diagnose a preschool child with CD. On the other, it may be that the threshold for a CD diagnosis in the K-DBDS is too low.

A number of limitations should be noted. First, research questions with regard to ADHD subtypes could not be answered as in the prior study reliability and validity of the ADHD predominantly inattentive could not be studied due to small number of children diagnosed with this subtype (Bunte et al., 2013b). Second, the sample size was relatively small. Third, generalizability of findings is unclear as demographic diversity of the sample may be different in other countries. Fourth, K-DBDS and CGAS data were derived from a single informant. Fifth, same individuals administered the K-DBDS at the first assessment and 18 month follow-up and thus were aware of the child's prior diagnosis. Sixth, treatment, although controlled in the analyses, may have affected instability in diagnostic status, given that treatment is effective.

Conclusion

Previous studies have shown that preschool children's diagnosis of ODD, CD and ADHD are likely to persist into school age, which is important to support the validity of these disorders in preschool children (Keenan et al., 2011; Lahey et al., 2004; Riddle et al., 2013). As a complement to these studies, present findings point to changes in the preschool and early school period. These changes may manifest either in instability of diagnosis (i.e., children with a diagnosis are under remission), or in new cases (i.e., children without a diagnosis at the first assessment are being diagnosed at later assessments). Diagnostic reassessments are needed in order to distinguish children belonging to the chronic stability group from those belonging to the remission groups, and to identify children who did not receive a diagnosis at the first assessment but develop a disorder over time. In the clinical evaluation of children referred for externalizing behavior problems attention should be paid both to impairment and number of symptoms.

Acknowledgments

Disclosure: Ms. Bunte, Schoemaker, and Dr. Hessen, van der Heijden, and Matthys report no biomedical financial interests or potential conflicts of interest.

The authors are grateful to the parents and children who participated in this study.

REFERENCES

Achenbach, T.M., & Rescorla, L.A. (2000). *Manual for the ASEBA preschool forms and profiles.* Burlington: Research Center for Children, Youth & Families VT: University of Vermont.

American Psychiatric Association. (2000). *Diagnostic and statistical manual of mental disorders (IV-TR Edition)*. Washington, DC: Author.

American Psychiatric Association. (2013). *Diagnostic and statistical manual of mental disorders (5th Edition)*. Washington, DC: Author.

Bunte, T.L., Laschen, S., Schoemaker, K., Hessen, D.J., van der Heijden, P.G.M., & Matthys, W. (2013a). Clinical Usefulness of Observational Assessment in the Diagnosis of DBD and ADHD in Preschoolers. *Journal of Clinical Child & Adolescent Psychology, 42,* 749-61.doi:10.1080/15374416.2013.7735.

Bunte, T.L., Schoemaker, K., Hessen, D.J., van der Heijden, P.G.M., & Matthys, W. (2013b). Clinical Usefulness of the Kiddie-Disruptive Behavior Disorder Schedule in the Diagnosis of DBD and ADHD in Preschool Children. *Journal of Abnormal Child Psychology, 41*, 681-690.

Comer, J.S., Chow, C., Chan, P.T., Cooper-Vince, C., & Wilson, L.A., (2013). Psychosocial treatment efficacy for disruptive behavior problems in very young children: a meta- analytic examination. *Journal of the American Academy of Child and Adolescent Psychiatry, 52,* 26-36.

Dunn, L.M., & Dunn, L.M. (2005). *Peabody picture vocabulary test (III-NL)* [Dutch version by L. Schlichting]. Amsterdam: Harcourt Assessment.

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.

Egger, H.L., Erkanli, A., Keeler, G., Potts, E., Walter, B.K., & Angold, A. (2006). Test-retest reliability of the Preschool Age Psychiatric Assessment (PAPA). *Journal of the American Academy of Child and Adolescence Psychiatry, 45*, 538-549.

Galéra, C., Côté, S.M., Bouvard, M.P., Pingault, J.B., Melchior M., Michel, G., Boivin, M, Tremblay, R.E. (2011). Early risk factors for hyperactivity-impulsivity and inattention trajectories from age 17 months to 8 years. *Archives of General Psychiatry, 68,* 1267 -75.

Hardy, K.K., Kollins, S.H., Murray, D.W., Riddle, M.A., Greenhill, L., Cunningham, C., ….Chuang, S.Z. (2007). Factor structure of parent- and teacher-rated attention-deficit/hyperactivity disorder symptoms in the Preschoolers with Attention -Deficit/Hyperactivity Disorder Treatment Study (PATS). *Journal of Child and Adolescence Psychopharmacology, 17*, 621-34.

Howell, E.M., & Teich, J. (2008). Variations in medicaid mental health service use and cost for children. *Administration and Policy in Mental Health, 35*(3), 220-8.

Keenan, K., Boeldt, D., Chen, D., Coyne, C., Donald, R., Duax J., …Humphries, M. (2011). Predictive validity of DSM-IV oppositional defiant and conduct disorders in clinically referred preschoolers. *Journal of Child Psychology and Psychiatry, 52*, 47-55.

Keenan, K., Wakschlag, L.S., & Danis, B. (2001). Kiddie-Disruptive Behavior Disorder Schedule (version 1.1).

Keenan, K., Wakschlag, L.S., Danis, B., Hill, C., Humphries, M., Duax, J., & Donald, R. (2007). Further evidence of the reliability and validity of DSM-IV ODD and CD in preschool children. *Journal of the American Academy of Child and Adolescent Psychiatry, 26*, 457-468.

Lahey, B., Pelham, W.E., Loney, D., Kipp, H., Ehrhardt, A., Lee, S, …Massetti, G. (2004). Three-Year predictive validity of DSM-IV attention deficit hyperactivity disorder in children diagnosed at 4-6 years of age. *American Journal of Psychiatry, 161*, 2014-20.

Lahey, B.B., Pelham, W.E., Loney, J., Kipp, H., Lee, S.S. & Wilcutt, E. (2005). Instability of the DSM-IV subtypes of ADHD from preschool through elementary school. *Archives of General Psychiatry, 62*, 896-902.

Murray, D.W., Kollins, S.H., Hardy, K.K., Abikoff, H.B., Swanson, J.M., Cunningham, C., … Chuang, S.Z. (2007). Parent versus teacher ratings of attention deficit/hyperactivity disorder symptoms in the Preschoolers with Attention -Deficit/Hyperactivity Disorder Treatment Study (PATS). *Journal of Child and Adolescence Psychopharmacology,17*, 605-20.

Orvaschel, H. & Puig-Antich, J. (1995). Schedule for Affective Disorders and Schizophrenia for School-Age Children-Epidemiologic 5th Version. Ft. Lauderdale, FL: Nova University.

Raven, J.C., Court, J.H., & Raven, J. (1998). *Raven coloured progressive matrices*. Oxford: Oxford Psychologist Press.

Riddle, M.A., Yershova, K., Lazzaretto, D., Paykina, N., Yenokyan, G. Greenhill, L. …Posner, K. (2013). The preschool Attention-Deficit/Hyperactivity Disorder treatment study (PATS) 6-year follow-up. *Journal of the American Academy of Child and Adolescent Psychiatry, 52, 264-278*.

Ringel, J.S., & Sturm, R. (2001). National estimates of mental health utilization and expenditures for children in 1998. *Journal of Behavioral Health Services & Research 28*, 319-33.

Schaffer, D., Goud, M.S., Brasic, J., Ambrosini, P., Fisher, P., Bird, H., & Aluwahlia, S. (1983). A children's global assessment scale (CGAS). *Archives of General Psychiatry*, *40*, 1228-1231.

Schlichting, L. (2005). Peabody picture vocabulary test III-NL. NL: Hartcourt Assessment.

Schoemaker, K., Bunte, T., Wiebe, S.A., Espy, K.A., Dekovic, M., & Matthys, W. (2012). Executive function deficits in preschool children with ADHD and DBD. *Journal of Child Psychology and Psychiatry*, *53*, 111-119.

Shaw, D.S., Lacourse, E., Nagin, D.S. (2005). Developmental trajectories of conduct problems and hyperactivity from ages 2 to 10. *Journal of Child Psychology and Psychiatry,* 46, 931-42

Wakschlag, L. S., Hill, C., Carter, A. S., Danis, B., Egger, H. L., Keenan, K., …Briggs -Gowan, M.J. (2008a). Observational Assessment of Preschool Disruptive Behavior, Part I: Reliability of the Disruptive Behavior Diagnostic Observation Schedule (DB-DOS). *Journal of the American Academy of Child and Adolescent Psychiatry, 47*, 622-631.

Wakschlag, L. S., Briggs-Gowan, M. J., Hill, C., Danis, B., Leventhal, B.D., Keenan, K., …Carter, A.S. (2008b). Observational Assessment of Preschool Disruptive Behavior, Part II: Validity of the Disruptive Behavior Diagnostic Observation Schedule (DB -DOS). *Journal of the American Academy of Child and Adolescent Psychiatry, 47*, 632-640.

Wakschlag, L.S., Leventhal, B.L., Briggs-Gowan, M.J., Danis, B., Keenan, K., Hill, C., …Carter, A.S. (2005). Defining the "disruptive" in preschool behavior: what diagnostic observation can teach us. *Clinical Child and Family Psychology Review, 8*, 183-201.

Wilens, T.E., Biederman, J., Brown, S., Monuteaux, M., Prince, J., & Spencer, T.J. (2002). Patterns of psychopathology and dysfunction in clinically referred preschoolers. *Journal of Developmental & Behavioral Pediatrics, 23*, S31-6.

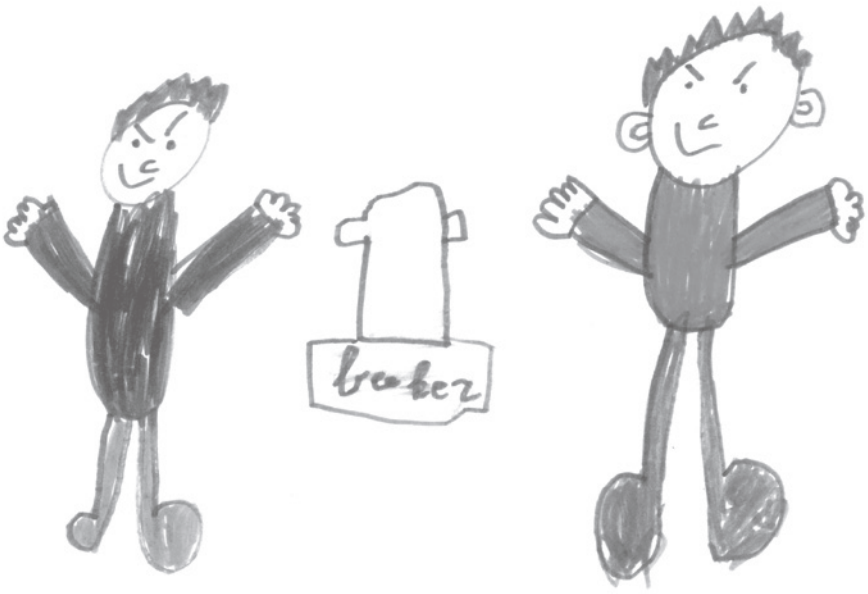
Willoughby, M.T., Pek, J., & Greenberg, M.T. (2012). Parent-Reported Attention Deficit/Hyperactivity symptomatology in Preschool-aged children: Factor structure, developmental change and early risk factors. *Journal of Abnormal Child Psychology, 40*, 1301-12.

Webster-Stratton, C., & Reid, M.J. The Incredible Years Parents, Teachers and Children Training Series: A multifaceted treatment approach for young children with conduct problems. In: Weisz J, Kazdin A, editors. Evidence-based psychotherapies for children and adolescents. 2nd edition Guilford; New York: 2010. pp. 194–210.

CHAPTER 5

Factors predicting stability of ODD, CD and ADHD diagnosis and persistence of symptoms in preschool children

Tessa L. Bunte, M.D., Kim Schoemaker, M.S., Kimberly A. Espy, Ph.D., Peter G.M. van der Heijden, Ph.D., Walter Matthys, M.D., Ph.D.



Bunte, T.L., Schoemaker, K., Espy, K.A., van der Heijden, P.G.M., Matthys, W. Factors predicting stability of ODD, CD and ADHD diagnosis and persistence of symptoms in preschool children. Manuscript submitted.

ABSTRACT

Studies on the stability of Oppositional Defiant Disorder (ODD), Conduct Disorder (CD) and Attention-Deficit/Hyperactivity Disorder (ADHD) diagnosis in preschool children not only have demonstrated stability of diagnosis but changes as well. The aim of the present study was to examine which individual and family factors predicted stability of diagnosis within an 18 months period. For diagnosing these disorders, a semi-structured diagnostic parent interview, i.e., the Kiddie-Disruptive Behavior Disorder Schedule (K-DBDS), was used. Participants were referred preschool children with externalizing behavioral problems (N=193; 83% male) and typically developing (TD) children (N=58; 71% male). Multiple logistic regression analyses showed that performance based impaired inhibition, occurrence of symptoms at daycare or school, and parental stress were associated with stability of both ODD/CD and ADHD Hyperactive Impulsive type diagnosis. We suggest that the clinical evaluation of a preschool child referred with behavior problems should include a neuropsychological assessment of inhibition, an assessment of symptoms in other settings than at home, and an assessment of parental stress.

INTRODUCTION

Epidemiological studies have demonstrated that the prevalence rates of Oppositional Defiant Disorder (ODD), Conduct Disorder (CD) and Attention-Deficit/Hyperactivity Disorder (ADHD) in preschool children are similar to those later in childhood (Egger & Angold, 2006). Moreover, the stability of ODD, CD and ADHD diagnoses in preschool children has been investigated. With regard to ODD and CD, a study on the predictive validity of ODD and CD in 3-5-year old children Keenan et al., (2011) showed that 73%, 66% and 51% of the children diagnosed with ODD at baseline met the criteria at 12-, 24-, and 36 month follow-up. Of the children who met criteria for CD at baseline, 48%, 33%, and 26% met criteria at 12-, 24-, and 36 month follow-up. With regard to ADHD, Lahey et al., (2004) examined stability of diagnosis in a three year predictive validity study in children diagnosed with ADHD at 3.8-7 years of age. This study found that 79% of the children who met criteria for ADHD at study entry received the diagnosis on at least two of three follow-up assessments. Likewise, in a follow-up of the Preschool ADHD Treatment Study (baseline mean age: 4,4 years), the rate of diagnosis irrespective of medications status at 3 year follow-up was 76% and at 6 year follow-up 77% (Riddle et al., 2013). In sum, there is evidence for the stability of ODD, CD and ADHD diagnosis in preschool children, but attention also should be paid to changes in diagnosis.

In a Dutch study on the stability and change of ODD, CD and ADHD diagnosis in preschoolers aged 3.5-5.5 years, follow-up assessments at 9 and 18 months allowed to distinguish children with a stable diagnosis from those under remission (Bunte et al., 2013c). From the perspective of stability, 62% of the children who were diagnosed with ODD at the first assessment were diagnosed again at 18 month follow-up; 35% of the children who were diagnosed with CD at the first assessment were diagnosed again at 18 months follow-up; 59% of the children who were diagnosed with ADHD at the first assessment were diagnosed again at 18 months follow-up. In addition, new cases were identified at follow-up assessments, i.e., children without a diagnosis of ODD, CD or ADHD at the first assessment were diagnosed at follow-up. In conclusion, as a complement to studies showing stability of ODD, CD and ADHD diagnosis into school age, these findings point to changes of diagnosis in the preschool period and early elementary school.

The aim of the present study was to identify factors that predict stability of ODD, CD and ADHD diagnosis in preschool children. To our knowledge, no such study has been conducted to date. A few non-clinical studies, however, examined which individual and family factors are associated with the persistence of behavior and hyperactivity/attention problems. With regard to individual factors, Campbell, March, Pierce, Ewing & Szumowski (1991) showed that in preschool boys, identified by their teacher as active, inattentive and impulsive, children's initial symptom levels predicted persistence of externalizing

problems at one-year follow-up. Moreover, externalizing problems in preschool children from the general population predicted externalizing DSM-IV (American Psychiatric Association, 1994) diagnoses in preadolescence, independent of the influence of early family risk factors such as low SES and harsh parenting (Mesman & Koot, 2001). Gender also seems to be associated with persistence of behavior problems. In a four-year longitudinal study, Beyer, Postert, Muller & Furniss (2012) showed changes in and continuity of behavioral and emotional problems in children from kindergarten to elementary school; the presence of behavioral and emotional problems at follow-up was higher among boys. Likewise, in a study of developmental trajectories of hyperactive symptoms from 2 to 7 years, male gender was a predictor of high and persistent hyperactivity (Romano, Tremblay, Farhat & Côté, 2006). Cognitive factors also have been found to be associated with the persistence of hyperactivity/attention problems. In the previously mentioned Campbell et al. study (1991), low child IQ was associated with persistent problems at two-year follow up (Campbell, 1994). Furthermore, Von Stauffenberg and Campbell (2007) showed that performance based behavioral inhibition, one of the executive functions, and attention deficits assessed at 54 months predicted teacher rated ADHD symptoms in elementary school, even after controlling for the stability of ADHD symptoms.

Among the family factors associated with the persistence of behavior and hyperactivity/attention problems maternal depression has been studied. Maternal depression discriminated two antisocial behavior trajectory groups of boys first assessed at 18 months; there were higher rates of depressive symptoms among mothers of boys in the high increasing group versus the stable low group (Shaw, Hyde and Brennan, 2012). Likewise, maternal depression was one of the factors associated with high trajectories of hyperactivity-impulsivity and inattention symptoms in a community sample of children followed up from age 5 months to 8 years (Galera et al., 2011). In the Campbell et al. study (1991) of preschool boys identified by their teacher as active, inattentive and impulsive, maternal depression predicted ratings of externalizing problems at one-year follow-up. At two-year follow-up, maternal depression scores and low child IQ were associated with persistent problems while other family stress measures did not add further to the discrimination between groups of boys with and without continuing problems (Campbell, 1994). Parental characteristics have been found to be associated with the persistence of behavior problems. For example, observed negative maternal control predicted one-year follow-up ratings of externalizing problems in the study by Campbell et al. (1991). Similarly, in a longitudinal study using a large community sample among the best predictors of the high physical aggression trajectory group at 5 months of age was mothers' self-reported coercive parenting behavior (Tremblay et al., 2004). Also, low family income at birth was one of the best predictors for the high physical aggression trajectory in this study (Tremblay et al., 2004).

The presence of symptoms in more than one context (both at home and daycare or school) is also a possible candidate for the prediction of persistence of problem behavior in preschool children. For ODD, presence of symptoms in multiple settings was not an issue in DSM-IV-TR (American Psychiatric Association, 2000), but in DSM-5 (American Psychiatric Association, 2013), severity of the disorder may be specified depending on the presence of symptoms in only one setting (mild), in at least two settings (moderate), or in three or more settings (severe). Therefore, presence of symptoms in more than one context as an expression of severity of initial symptoms (see Campbell et al., 1991; Mesman & Koot, 2001) may be a predictor of persistence of ODD symptoms. With regard to ADHD, there has been a debate about the presence of symptoms or impairment in multiple settings as a criterion for ADHD, and criteria seem to become stricter over time. While according to DSM-IV-TR (American Psychiatric Association, 2000) some impairment of the symptoms was a criterion for ADHD, according to DSM-5 (American Psychiatric Association, 2013) several symptoms must be present in two or more settings. However, in the study by Bunte et al. (2013b) on the clinical utility of the Kiddie-Disruptive Behavior Disorder Schedule in preschool children it was shown that the presence of ADHD symptoms in more than one context resulted in a worse sensitivity-specificity balance. Likewise, in a longitudinal study it was shown that a substantial number of preschool children who meet symptom criteria for ADHD but do not meet the cross-sectional impairment later, meet full diagnostic criteria for ADHD (Lahey et al., 2004). Yet studies on the occurrence of symptoms in two or more settings as a predictor of ADHD or ODD symptoms are lacking.

In summary, several individual and family factors have been shown to be associated with the persistence of preschool behavior problems and hyperactivity/attention problems into elementary school age. The main aim of the present study was to examine which individual factors (initial symptom level, gender, IQ, behavioral inhibition, occurrence of symptoms at daycare or school) and which family factors (SES, stress, maternal depression, harsh and inconsistent discipline) are predictors of stability of ODD, CD and ADHD diagnosis in the preschool and early elementary school period. We also investigated which factors predict persistence of ODD, CD and ADHD symptoms. Therefore, in a mainly referred group of preschool children aged 3.5–5.5 years we investigated whether the above mentioned factors predicted 1. ODD, CD and ADHD diagnosis according to the Kiddie-Disruptive Behavior Diagnostic Schedule (K-DBDS; Keenan et al., 2007, Bunte et al., 2013b) at 18 months follow-up, 2. the total symptom score of ODD, CD and ADHD symptoms on the K-DBDS.

METHOD

Participants

Participants were 251 children, aged at baseline 42-66 months (mean=55 months, SD=7.8); 80% were male, 86% White, 12% Turkish/Moroccans, 2% African, 0.5% Asian. In terms of educational level of the parents: primary accounted for 6%, secondary 25%, intermediate vocational 34%, higher vocational 16% and university 19%. General practitioners, pediatricians and well-baby clinics in the province of Utrecht were invited to refer children aged 3.5-5.5 with externalizing behavior problems to the Outpatient Clinic for Preschool Children with Behavioral Problems, Department of Psychiatry, University Medical Center, Utrecht, the Netherlands. Children were included if they scored at or above the 90-th percentile either on the Aggressive Behavior scale or on the Attention Problems scale of the Child Behavior Checklist (CBCL/1.5-5) or the Child-Teacher Report Form (C-TRF/1.5-5) (Achenbach & Rescorla, 2000). The typically developing (TD) group (N=58) was recruited from regular elementary schools and daycare centers. Children with a score in the clinical range either on the Attention Problems scale or on the Aggressive Behavior scale of the CBCL/1.5-5 or C-TRF/1.5-5 were excluded. All the children (referred and TD) with an IQ below 70, estimated with the average score on the Raven Coloured Progressive Matrices (Raven, Court, & Raven, 1998) and the Peabody Picture Vocabulary Test-III-NL (Dunn & Dunn, 2005; Schlichting, 2005), were excluded. Children were followed over a period of 18 months (Mean=18.1 months, range=15-24 months, SD=1.2 months). Each child was given a diagnosis of either ODD, CD or ADHD at the first assessment (T1) and the follow-up assessment 18 months later (T3) based on the K-DBDS.

Several children received treatment after the first assessment: individual parent counseling at home or at the outpatient clinic, the Incredible Years parent program (Webster-Stratton & Reid, 2010), and/or pharmacotherapy (methylphenidate in most cases, but atomoxetine and risperidone were prescribed as well). Between T1 and T3, 137 (71%) of the referred preschool children have received treatment.

Procedure

At each assessment children were evaluated in a single morning session. At T1, two measures of intellectual functioning were administered followed by executive functions (EF) tasks (Schoemaker et al., 2012). After a break the child was observed using the Disruptive Behavior Diagnostic Observation Schedule (DB-DOS; Wakschlag et al., 2008ab; Bunte et al.,, 2013a) and the K-DBDS (Keenan et al.,, 2007, Bunte et al.,, 2013b) was administered. At T3, the assessment was similar to T1 with the exception of the assessment of intellectual functioning. The K-DBDS at T1 and T3 was administrated by the same child psychiatrist/clinical child psychologist. Retention was high: 95% at T3. There were no significant

differences between those participants who were and were not seen at T3 in terms of age, sex, referral status or diagnosis at T1.
Parents received a small financial compensation for participating, children received two small gifts. Written informed consent from the parents was gathered before participating. The Medical Ethical Review Committee of the University Medical Center, Utrecht, approved this study.

Measures

Total symptom score on the K-DBDS; the K-DBDS is a semi-structured clinical interview assessing DSM-IV (American Psychiatric Association, 2000) ODD, CD and ADHD symptoms in preschool children (Keenan et al., 2007, Bunte et al., 2013b). It is based on the Schedule for Affective Disorders and Schizophrenia for School-Age Children (Orvaschel & Puig-Antich, 1995). Administrating the K-DBDS takes approximately 45 minutes, coding takes approximately ten minutes. To determine if the behavior is a symptom according to the K-DBDS, the Qualitative Coding Method (QCM) has been used (K-DBDS Preschool Version Instruction Manual and Bunte et al., 2013b). A restriction of the K-DBDS for ADHD subtypes, is that the operationalization of the ADHD predominantly inattentive type could not be examined due to low numbers of children diagnosed with this subtype (Bunte et al., 2013b).
Inhibition; three inhibition tasks, who were considered to preferentially measure inhibitory skills (Go-No-Go task, Snack Delay and Shape School) were used (Schoemaker et al., 2012); The *Go-No-Go task* is a computerized task where children were instructed to catch as many fish (Go stimuli, 75%) as possible by pressing the button when a fish appeared on the screen. They were instructed to let the shark (No-Go stimuli, 25%) swim by withholding the button press. Auditory feedback was provided when appropriately catching a fish or inappropriately catching a shark. *Modified Snack Delay* is a newly developed task that integrates the motivational context from the original Snack Delay paradigm (Kochanska, Murray, Jacques, Koenig, & Vandegeest, 1996) with the motor – inhibitory control demands of NEPSY Statue (Korkman, Kirk & Kemp, 1998). To measure pure motor inhibition only hand movement was used, where no movement was assigned a score of 1, some hand movement .5 and lots of movement 0. The *Shape School-Inhibit Condition* is a computerized task with cartoon figures with different shapes, colors, and expressions, where the naming rule differs in varying conditions. In the Inhibit condition, participants had to name the color of the figures with happy faces and suppress the prepotent color naming response when the figure had a sad/frustrated face. In this study we used the z-scores of the three tasks for the inhibition score.
Daycare or school context: Child-Teacher Report Form/1.5-5 (C-TRF/1.5-5.; Achenbach & Rescorla, 2000); disruptive behavior symptoms and attention problems in the daycare or

school context were assessed using the scores of the Attention Problem scale and the Aggressive Behavior scale of the C-TRF/1.5-5 completed by teachers or daycare caregivers;
Social Economic Status (SES); for the social economic status we used the educational level of the father; Nine different educational levels were used: from 0=no education, 1=primary education, for example 5=high school, till 8=university.
Parenting Stress Index (PSI; Abidin, 1983); this questionnaire was designed to measure parental stress perception with raising their child for parents of children between 2-13 years old. The parental characteristics domain evaluates the source of stress and potential dysfunction on the parent-child system as related to parental functioning. Mothers answered the 26 questions on a six-point scale ranging from "totally disagree" to "totally agree". It yields a total stress score with a possible range from 26-156. In this study a mean score was used. Cronbach alpha in the present study was 0.92.
Beck Depression Inventory (BDI; Beck, Ward, Mendelson, Mock & Erbaugh, 1961); this inventory was designed to assess general depressive symptoms. Mothers reported on 21 items regarding their depressive feelings on a four-point scale (0-3). A total score was derived. Cronbach alpha in the present study was 0.84.
Parent Practices Interview (PPI; Webster-Stratton, 2001); this interview was designed to measure dimensions of parenting practices of parents of young children. The PPI consisted of 64 items on a seven-point scale, ranging from "not (likely) at all" to "always/very likely". Mothers reported on their response to misbehavior or appropriate behavior. To assess negative control we only used the scale: "Harsh and inconsistent discipline" (e.g., threatening, but not punishing, 15 items, Cronbach alpha=0.83). Summary scales were computed, consequently each scale ranged from 0 to 7.

Statistical analyses

For all scales used, the percentage of missing data for the different questionnaires was: 0-1.6%. We made sure mothers filled out the questionnaires on the day of assessment, consequently there were very few missing data for these questionnaires. An exception was the educational level of the father with 6.4% missing data. For each analysis children with missing data were excluded. First, we computed bivariate correlations between ODD, CD and ADHD HI symptoms on the K-DBDS at baseline and 18 month follow-up, and individual factors at baseline (gender, IQ, inhibition, C-TRF attention problem scale/ aggressive behavior scale) and family factors at baseline (SES, PSI, BDI, and PPI [HID]) in the total group (N=238 at follow-up) (referred and TD children). Second, using simple and multiple hierarchical linear regression analyses the association between individual and family factors at baseline and the total symptom score on the K-DBDS at follow-up after 18 months was examined in the total group (N=238) (referred and TD children), with the total symptom score as the first step, and individual factors and family factors as second step.

Third, using simple and multiple logistic regression analyses we examined the association between individual and family factors at baseline, and a DBD and ADHD HI diagnosis at 18 months follow-up compared to the TD children. There were 102 children who had an ODD diagnosis at T3 based on the K-DBDS and 46 children who had a CD diagnosis at T3, based on the K-DBDS. Because most of the CD diagnosed children (N=44) scored both on ODD and CD we used a DBD group (N=104) instead of a separate ODD and CD group. A child was placed in the DBD group with and without an ADHD HI diagnosis. Likewise, a child was placed in the ADHD HI group (N=76) with and without a DBD diagnosis.

RESULTS

The TD, DBD and ADHD HI group were first compared on demographic, individual-, including daycare or school context, and family characteristics, see Table 1. The differences between the clinical groups and the TD group were as expected. Bivariate correlations between individual and family factors displayed in Table 2 were in general as expected.
Simple linear regression analyses showed that for ODD, CD and ADHD HI symptom scores at baseline were associated with the symptom scores at T3 (Table 3). With regard to individual factors, low inhibition scores at baseline were associated with ADHD HI symptom score at T3, and high TRF aggressive behavior/attention problem scale scores at baseline were associated with the symptom scores for ODD, CD and ADHD at T3. With regard to the family factors, analyses showed that scores for family factors (low SES, and high PSI, high BDI, and high PPI) at baseline were associated with symptom scores for ODD, CD and ADHD HI at T3. Multiple hierarchical linear regression analyses showed that for ODD, CD and ADHD HI, symptom scores at baseline were associated with the symptom scores at T3 (Table 3). For ODD, TRF aggression scores at baseline were associated with the symptom scores at T3. For ADHD, IQ at baseline was associated with the symptom score at T3, based on the level of association (B=.033), meaning that 30 points (more) on the IQ scale is associated with one ADHD symptom. With regard to the family factors, for CD and ADHD, high PSI scores at baseline were associated with the symptom scores at T3. For ODD, the high PPI (HID) scores at baseline were associated with the symptom score at T3. Simple logistic regression analyses showed that scores for both individual and family factors were all associated with a DBD and ADHD HI diagnosis at T3 (Table 4). Multiple logistic regression analyses showed that for both DBD and ADHD HI, low inhibition scores, high TRF scores and high PSI scores at baseline were associated with a DBD and ADHD HI diagnosis at T3.

Table 1 *Means (and SD) for Demographic, Individual and Family Factors at baseline and follow-up*

	TD (N=58)	DBD (N=104)	ADHD HI (N=76)
	Mean (SD)	Mean (SD)	Mean (SD)
Age (months) T1	56.0 (7.3)	55.1 (7.6)	54.0 (8.0)
% Boys	71%	83%*	83%*
IQ	112.4 (10.4)	103.3 (11.6)***	104.4 (11.5)***
ODD T1	1.2 (1.3)	5.4 (1.6)***	5.1 (1.8)***
ODD T3	1.0 (1.4)	5.4 (1.4)***	4.8 (1.8)***
CD T1	.4 (.7)	2.8 (1.8)***	2.7 (1.9)***
CD T3	.4 (.6)	2.3 (1.7)***	2.1 (1.9)***
ADHD T1	1.2 (1.3)	5.7 (2.1) ***	6.6 (1.8)***
ADHD T3	1.2 (1.7)	5.3 (2.5) ***	7.0 (1.9)***
INH T1	.6 (.4)	-.1 (.7)***	-.1 (.7)***
TRF agg T1	52.0 (3.1)	64.1 (10.6)***	63.0 (10.6)***
TRF att T1	52.0 (3.8)	62.9 (11.3)***	65.8 (11.3)***
SES T1	6.5 (2.0)	4.4 (2.2)***	4.2 (2.2)***
PSI T1	12.4 (3.8)	20.0 (5.0)***	20.0 (5.6)***
BDI T1	4.3 (4.0)	10.7 (6.1)***	10.9 (5.6)***
PPI (HID) T1	2.5 (.5)	3.0 (.8)***	3.0 (.8)***

Note. TD=typically developing; DBD=disruptive behavior disorder (with and without ADHD HI); ADHD HI=attention-deficit/hyperactivity disorder predominantly hyperactive-impulsive type (with and without DBD); IQ[a] estimate based on mean Raven and Peabody; ODD=Oppositional Defiant Disorder (total symptom score on K-DBDS); CD=Conduct Disorder (total symptom score on K-DBDS); T1=Baseline; T3=18 months follow-up; INH=Inhibition; TRF=Teacher Report Form; agg=aggressive behavior scale; att=attention problem scale; SES=Social Economic Status; PSI=Parenting Stress Index; BDI=Becks Depression Inventory; PPI=Parenting Practices Interview; (Harsh and Inconsistent Discipline [HID]);.

*p <.05. ***p <.001. (compared with the TD Group).

Table 2 *Bivariate correlation between Individual Factors and Environmental Factors at baseline and follow-up*

	1.	2.	3.	4.	5.	6.	7.	8.	9.	10.	11.	12.	13.	14.	15.
1. ODD T1	-														
2. CD T1	.60**	-													
3. ADHD T1	.51**	.43**	-												
4. ODD T3	.62**	.49**	.39**	-											
5. CD T3	.47**	.42**	.30**	.60**	-										
6. ADHD T3	.41**	.34**	.68**	.53**	.40**	-									
7. Gender T1	-.06	-.15*	-.13*	-.05	-.05	-.12	-								
8. IQ[a] T1	-.19**	-.13*	-.21**	-.09	-.12	-.05	.07	-							
9. INH T1	-.19**	-.18**	-.33**	-.06	-.07	-.16*	.07	.50**	-						
10. TRF agg T1	.26**	.35**	.30**	.19**	.15*	.22**	-.13*	-.20**	-.25**	-					
11.TRF att T1	.19**	.24**	.36**	.08	.05	.23**	-.10	-.31**	-.33**	.68**	-				
12. SES T1	-.24**	-.12	-.30**	-.15*	-.15*	-.24**	.01	.23**	.28**	-.12	-.18**	-			
13. PSI T1	.45**	.33**	.45**	.45**	.32**	.43**	-.04	-.10	-.11	.18**	.11	-.21**	-		
14. BDI T1	.35**	.27**	.37**	.41**	.23*	.35**	-.02	-.09	-.12	.09	.09	-.25**	.71**	-	
15. PPI T1	.25**	.18**	.23**	.31**	.15*	.22**	.01	-.10	-.05	.11	.11	-.19**	.32**	.28**	-

Note. ODD=Oppositional Defiant Disorder (total symptom score on the K-DBDS); CD=Conduct Disorder (total symptom score on the K-DBDS); ADHD HI=attention-deficit/hyperactivity disorder predominantly hyperactive-impulsive type (total symptom score on the K-DBDS); T1=Baseline; T3=18 months follow-up; IQ[a]=estimate based on mean Raven and Peabody; INH=Inhibition; TRF=Teacher Report Form; agg=aggressive behavior scale; att=attention problem scale; BDI=Becks Depression Inventory; PSI=Parenting Stress Index, SES=Social Economic Status; PPI=Parenting Practices Interview; (Harsh and Inconsistent Discipline [HID]).

*p<.05. **p<.01.

Table 3 *Simple and Multiple Linear Regression with the total group (N=238)*

	Simple			Multiple	
	B (SE)	β	r^2	B (SE)	Beta
ODD T3					
Step 1.					
ODD T1				.64 (.05)	.64***
Step 2.					
ODD T1	.62 (.05)	.62***	.38	.52 (.06)	.51***
Gender	-.27 (.37)	-.05	.00	-.17 (.28)	-.03
IQ[a]	-.02 (.01)	-.09	.01	.01 (.01)	.03
Inhibition	-.18 (.20)	-.06	.00	.15 (.18)	.05
TRF agg	.04 (.02)	.19**	.04	.03 (.01)	.11*
SES	-.17 (.07)	-.15*	.02	.03 (.06)	.03
PSI	.18 (.02)	.45***	.20	.03 (.03)	.07
BDI	.16 (.02)	.41***	.17	.05 (.03)	.14
PPI (HID)	1.00 (.20)	.31***	.10	.42 (.17)	.13*
CD T3					
Step 1					
CD T1				.37 (.05)	.44***
Step 2.					
CD T1	.36 (.05)	.42***	.18	.30 (.06)	.36***
Gender	-.18 (.24)	-.05	.00	.03 (.23)	.01
IQ[a]	-.02 (.01)	-.12	.01	-.01 (.01)	-.04
Inhibition	-.12 (.13)	-.07	.00	.13 (.14)	.07
TRF agg	.02 (.01)	.15*	.02	.01 (.01)	.05
SES	-.11 (.04)	-.15*	.02	-.06 (.04)	-.09
PSI	.08 (.02)	.32***	.10	.06 (.02)	.21*
BDI	.06 (.02)	.23***	.05	-.01 (.02)	-.05
PPI (HID)	.31 (.14)	.15*	.02	.04 (.14)	.02

	Simple			Multiple	
	B (SE)	β	r^2	B (SE)	Beta
ADHD HI T3					
Step 1.					
ADHD HI T1				.70 (.05)	.68***
Step 2.					
ADHD HI T1	.69 (.05)	.68***	.46	.61 (.06)	.59***
Gender	-.79 (.44)	-.12	.01	-.34 (.33)	-.05
IQ[a]	-.01 (.02)	-.05	.00	.03 (.01)	.14*
Inhibition	-.58 (.23)	-.16*	.03	-.04 (.21)	-.01
TRF att	.05 (.02)	.23***	.05	.01 (.01)	.05
SES	-.29 (.08)	-.24***	.06	-.04 (.06)	-.03
PSI	.21 (.03)	.43***	.18	.07 (.03)	.15*
BDI	.17 (.03)	.35***	.12	.02 (.03)	.03
PPI (HID)	.84 (.25)	.22***	.05	.12 (.20)	.03

Note. ODD=Oppositional Defiant Disorder (total symptoms); CD=Conduct Disorder (total symptoms); ADHD HI=attention-deficit/hyperactivity disorder predominantly hyperactive-impulsive type (with and without ODD, CD); T1=Baseline; T3=18 months follow-up; IQ[a]=estimate based on mean Raven and Peabody; TRF=Teacher Report Form; agg=aggressive behavior scale; att=attention problem scale; SES=Social Economic Status; PSI=Parenting Stress Index, BDI=Becks Depression Inventory; PPI=Parenting Practices Interview; (Harsh and Inconsistent Discipline [HID])

Δ R^2=.41 for Step 1 ODD (p<.001), Δ R^2=.07 for Step 2 ODD (p<.001), Δ R^2=.19 for Step 1 CD (p<.001), Δ R^2=.05 for Step 2 CD (p<.001), Δ R^2=.47 for Step 1 ADHD HI (p<.001), Δ R^2=.05 for Step 2 ADHD HI (p<.001).

*p<.05. **p<.01. ***p<.001.

Table 4 *Simple and Multiple Logistic Regression*

	Simple				Multiple			
		95% CI for Exp b						
	B (SE)	Lower	Exp b	Upper	B (SE)	Lower	Exp b	Upper
DBD T3 N=104								
Gender	-.83 (.38)*	1.09	2.30	4.86	-.93 (.94)	.40	2.50	16.10
IQ[a]	-.07 (.02)***	1.04	1.07	1.11	-.05 (.04)	.98	1.05	1.13
Inhibition	-2.20 (.44)***	3.82	9.04	21.38	-2.20 (.86)*	1.64	8.83	47.63
TRF agg	.26 (.05)***	.70	.77	.85	.34 (.10)***	.59	.71	.86
SES	-.46 (.09)***	1.32	1.58	1.89	-.16 (.18)	.82	1.17	1.68
PSI	.35 (.06)***	.63	.70	.78	.43 (.13)***	.50	.65	.85
BDI	.27 (.05)***	.69	.76	.84	.05 (.12)	.74	.95	1.21
PPI (HID)	1.47 (.34)***	.12	.23	.45	.87 (.73)	.10	.42	1.75
ADHD HI T3 N=76								
Gender	-.86 (.41)*	1.05	2.36	5.31	-1.94 (1.70)	.24	6.98	200.58
IQ[a]	-.06 (.02)***	1.03	1.07	1.10	.04 (.05)	.87	.96	1.07
Inhibition	-2.40 (.49)***	4.25	11.01	28.54	-4.49 (1.79)*	2.67	89.3	2990.5
TRF att	.28 (.05)***	.69	.76	.84	.48 (.16)***	.46	.62	.85
SES	-.48 (.10)***	1.34	1.61	1.95	-.19 (.27)	.72	1.21	2.04
PSI	.30 (.05)***	.67	.74	.82	.71 (.35)*	.25	.49	.98
BDI	.30 (.06)***	.67	.74	.83	-.28 (.33)	.70	1.32	2.51
PPI (HID)	1.50 (.37)***	.11	.22	.46	1.17 (1.14)	.03	.31	2.92

Note. CI=Confidence Interval; DBD=disruptive behavior disorder (with and without ADHD HI); ADHD HI=attention-deficit/hyperactivity disorder predominantly hyperactive-impulsive type (with and without DBD); T3=18 months follow-up; IQ[a]=estimate based on mean Raven and Peabody; TRF=Teacher Report Form; agg=aggressive behavior scale; att=attention problem scale; SES=Social Economic Status; PSI=Parenting Stress Index, BDI=Becks Depression Inventory; PPI=Parenting Practices Interview; (Harsh and Inconsistent Discipline [HID]).
*p<.05. **p<.01. ***p<.001

DISCUSSION

Studies on the stability of ODD, CD and ADHD diagnosis in preschool children not only have demonstrated stability of diagnosis but changes as well (Keenan et al., 2011; Lahey et al., 2004; Riddle et al., 2013; Bunte et al., 2013b). The aim of the present study was to examine which individual and family factors predicted stability of diagnosis within an 18 months period. Multiple logistic regression analyses showed that impaired inhibition, occurrence of symptoms at daycare or school, and parental stress were associated with stability of both ODD/CD (or DBD) and ADHD HI diagnoses. We also investigated which

factors predicted persistence of ODD, CD and ADHD symptoms. Multiple linear regression analyses showed that occurrence of symptoms at daycare or school, and harsh and inconsistent discipline were associated with persistence of ODD symptoms, while parental stress was associated with persistence of CD and ADHD HI symptoms.

The present study is the first one to show that performance based poor inhibition is a predictor of DBD and ADHD HI diagnostic stability. Findings are in line with the study by Von Stauffenberg and Campbell (2007) demonstrating that performance based behavioral inhibition and attention deficits assessed at 54 months predicted teacher rated ADHD symptoms in elementary school. Over the last decade there has been an increasing interest in executive functions among which inhibition in young children. In a meta-analysis of performance based inhibition studies in preschoolers with ADHD symptoms or diagnosis a medium to large effect size was found for inhibition (Pauli-Pott & Becker, 2011). Likewise, in a meta-analysis of preschoolers with externalizing behavior problems a medium correlational effect size was obtained for inhibition (Schoemaker, Mulder, Dekovic & Matthys, 2013). Finally, when compared to TD children, preschool children with ADHD and DBD showed inhibition deficits (Schoemaker et al., 2012). In sum, poor inhibition performance not only seems to be a characteristic of preschool children with ADHD and DBD symptoms or diagnoses, but also to play a role in the stability of ADHD and DBD diagnoses.

The occurrence of DBD or ADHD HI symptoms at daycare or school was another predictor of diagnostic DBD and ADHD HI stability. In addition, symptoms at daycare or school predicted ODD symptoms at follow-up. Presence of symptoms in more than one setting is an understudied area in preschool children with ODD. Yet, according to DSM-5 (American Psychiatric Association, 2013) severity of the disorder may be specified depending on the presence of symptoms in only one setting (mild), in at least two settings (moderate), or in three or more settings (severe). Results of the study support the need to consider symptoms at daycare or school in preschool children, possibly an expression of severity of initial symptoms (Campbell et al., 1991; Mesman & Koot, 2001). It might be that in preschool children ODD symptoms occurring only at home are too mild to be indicative of an ODD diagnosis. Further research on this important area is needed. With regard to ADHD, recently the presence of symptoms in multiple settings has become part of the criteria for ADHD. While according to DSM-IV-TR (American Psychiatric Association, 2000) some impairment of the symptoms in multiple settings was a criterion for ADHD, according to DSM-5 (American Psychiatric Association, 2013) several symptoms must be present in two or more settings. From the perspective of the stability of the ADHD diagnosis in preschool children, results of the study support the need to include presence of symptoms not only at home but at daycare or school as well.

Parental stress predicted both the stability of DBD and ADHD HI diagnosis, and the persistence of CD and ADHD HI symptoms, while harsh and inconsistent discipline only predicted the persistence of ODD symptoms. It may be that parental stress was a more

global adverse family characteristic than maternal depression or harsh and inconsistent discipline, and the result of many negative factors among which the child's symptoms. Correlations between parental stress on the one hand, and both parent-reported symptoms and caregiver/teacher reported problems indeed were high. Harsh and inconsistent discipline as a specific predictor of the persistence of ODD symptoms probably refers to the typical coercive interactions (Patterson, Reid & Dishion, 1992) characteristic of the frequently occurring but relatively mild ODD symptoms. The lack of association between SES and persistence of symptoms or stability of diagnosis probably may be due to the limited range of SES in the sample, consisting mainly of middle class families.

A number of limitations should be noted. First, research questions with regard to ADHD subtypes could not be answered as in the prior study reliability and validity of the ADHD predominantly inattentive could not be studied due to small number of children diagnosed with this subtype (Bunte et al., 2013b). Second, the sample size was relatively small, and it was not possible to take comorbidity groups as separate groups. Third, generalizability of findings is unclear as demographic diversity of the sample may be different in other countries. Fourth, it was not possible to control for treatment as treatment was closely connected with group membership (one of the clinical groups versus the typically developing group). Findings of this naturalistic study therefore are valid for preschool children visiting an outpatient clinic.

Clinical implications

With regard to clinical implications, first, findings of the occurrence of symptoms at daycare or school as a predictor of the stability of diagnosis support the notion that the clinical evaluation of a preschool child referred with behavior problems must include a careful assessment of symptoms in other settings than at home. In addition, interventions should target these symptoms. Second, results of poor inhibition as a predictor of diagnostic stability suggest that the neuropsychological assessment of inhibition and possibly other executive functions such as working memory become part of the clinical evaluation of preschool children. The study also supports efforts at developing and evaluating training methods of executive functions for preschool children. Third, the clinical evaluation of preschool children referred with behavior problems should include the assessment of parental stress which is strongly associated with ODD, CD and ADHD symptoms at home and at daycare or school.

Acknowledgments.

Disclosure: Ms. Bunte, Schoemaker, Espy, and Dr. van der Heijden, and Matthys report no biomedical financial interests or potential conflicts of interest.

The authors are grateful to the parents and children who participated in this study.

REFERENCES

Abidin, R.R. (1983). Parental Stress Index: Manual. Charlottesville: Pediatric Psychology Press. Achenbach, T.M., & Rescorla, L.A. (2000). *Manual for the ASEBA preschool forms and profiles.* Burlington: Research Center for Children, Youth & Families VT: University of Vermont.

American Psychiatric Association. (1994). *Diagnostic and statistical manual of mental disorders (4 thEdition)*. Washington, DC: Author.

American Psychiatric Association. (2000). *Diagnostic and statistical manual of mental disorders (4 thEdition, text-revised)*. Washington, DC: Author.

American Psychiatric Association. (2013). *Diagnostic and statistical manual of mental disorders (5 th Edition)*. Washington, DC: Author.

Beck, A., Ward, C., Mendelson, M., Mock, J., & Erbaugh, J. (1961). An inventory for measuring depression. *Archives of general psychiatry, 4,* 561-571.

Beyer, T., Postert, C., Müller, J.M., & Furniss, T. (2012). Prognosis and continuity of child mental health problems from preschool to primary school: results of a four-year longitudinal study. *Child Psychiatry Human Development, 43,*533-543.

Bunte, T.L., Laschen, S., Schoemaker, K., Hessen, D.J., van der Heijden, P.G.M., & Matthys, W. (2013a). Clinical Usefulness of Observational Assessment in the Diagnosis of DBD and ADHD in Preschoolers. *Journal of Clinical Child & Adolescent Psychology*, *42,* 749-61. doi:10.1080/15374416.2013.7735.

Bunte, T.L., Schoemaker, K., Hessen, D.J., van der Heijden, P.G.M., & Matthys (2013b). Clinical Usefulness of the Kiddie-Disruptive Behavior Disorder Schedule in the Diagnosis of DBD and ADHD in Preschool Children. *Journal of Abnormal Child Psychology, 41*, 681-690.

Bunte, T.L., Schoemaker, K., Hessen, D.J., van der Heijden, P.G.M., & Matthys (2013c). Stability and Change of ODD, CD and ADHD diagnosis in Referred Preschool Children. Under revision; *Journal of Abnormal Child Psychology.*

Campbell, S.M. (1994). Hard-to-manage preschool boys: Externalizing behaviour, social competence, and family context a two-year follow-up. *Journal of Abnormal Child Psychology, 22*, 147-166.

Campbell, S.B., March, C.L., Pierce, E.W., Ewing, L.J., & Szumowski, E.K. (1991). Hard-to -manage preschool boys: family context and the stability of externalizing behavior. *Journal of Abnormal Child Psychology, 19,* 301-318.

Dunn, L.M., & Dunn, L.M. (2005). *Peabody picture vocabulary test (III-X)* [X version by L. Schlichting]. X: Harcourt Assessment.

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.

Galéra, C., Côté, S.M., Bouvard, M.P., Pingault, J.B., Melchior M., Michel, G., Boivin, M, Tremblay, R.E. (2011). Early risk factors for hyperactivity-impulsivity and inattention trajectories from age 17 months to 8 years. *Archives of General Psychiatry, 68,* 1267 -75.

Keenan, K., Boeldt, D., Chen, D., Coyne, C., Donald, R., Duax J., …Humphries, M. (2011). Predictive validity of DSM-IV oppositional defiant and conduct disorders in clinically referred preschoolers. *Journal of Child Psychology and Psychiatry, 52*, 47-55.

Keenan, K., Wakschlag, L.S., & Danis, B. (2001). Kiddie-Disruptive Behavior Disorder Schedule (version 1.1).

Keenan, K., Wakschlag, L.S., Danis, B., Hill, C., Humphries, M., Duax, J., & Donald, R. (2007). Further evidence of the reliability and validity of DSM-IV ODD and CD in preschool children. *Journal of the American Academy of Child and Adolescent Psychiatry, 26*, 457-468.

Kochanska, G., Murray, K., Jacques, T. Y., Koenig, A. L., & Vandegeest, K. A. (1996). Inhibitory control in young children and its role in emerging internalization. *Child Development, 67,* 490–507. doi: 10.1111/j.1467-8624.1996.tb01747.x

Korkman, M., Kirk, U., & Kemp, S. (1998). NEPSY: A developmental neuropsychological assessment. Boston, MA: The Psychological Association.

Lahey, B., Pelham, W.E., Loney, D., Kipp, H., Ehrhardt, A., Lee, S, …Massetti, G. (2004). Three-Year predictive validity of DSM-IV attention deficit hyperactivity disorder in children diagnosed at 4-6 years of age. *American Journal of Psychiatry, 161*, 2014-20.

Mesman, J. & Koot, H.M. (2001). Early preschool predictors of preadolescent internalizing and externalizing DSM-IV diagnoses. *Journal of the American Academy of Child and Adolescent Psychiatry, 40,* 1029-1036.

Orvaschel, H. & Puig-Antich, J. (1995). Schedule for Affective Disorders and Schizophrenia for School-Age Children-Epidemiologic 5th Version. Ft. Lauderdale, FL: Nova University.

Patterson, G. R., Reid, J. B., & Dishion, T. J. (1992). *Antisocial boys*. Eugene, OR: Castalia.

Pauli-Pott, U., & Becker, K. (2011). Neuropsychological basic deficits in preschoolers at risk for ADHD: a meta-analysis. *Clinical Psychological Review, 31*, 626-637.

Raven, J.C., Court, J.H., & Raven, J. (1998). *Raven coloured progressive matrices*. Oxford: Oxford Psychologist Press.

Riddle, M.A., Yershova, K., Lazzaretto, D., Paykina, N., Yenokyan, G. Greenhill, L. …Posner, K. (2013). The preschool Attention-Deficit/Hyperactivity Disorder treatment study (PATS) 6-year follow-up. *Journal of the American Academy of Child and Adolescent Psychiatry, 52,* 264-278.

Romano, E., Tremblay, R.E., Abdeljelil, F. & Côté, S. (2006). Development and prediction of hyperactive symptoms from 2 to 7 years in a population based sample. *Pediatrics, 117*, 2101-2110.

Schlichting, L. (2005). Peabody picture vocabulary test III-NL.NL: Hartcourt Assessment.

Schoemaker, K., Bunte, T., Wiebe, S.A., Espy, K.A., Dekovic, M., & Matthys, W. (2012). Executive function deficits in preschool children with ADHD and DBD. *Journal of Child Psychology and Psychiatry*, *53*, 111-119.

Schoemaker, K., Mulder, H., Dekovic, M., & Matthys, W. (2013). Executive functions in preschool children with externalizing behavior problems: a meta-analysis. *Journal of Abnormal Child Psychology, 41,* 457-471.

Shaw, D.S., Hyde, L.W., & Brennan, L.M. (2012). Early predictors of boys' antisocial trajectories. *Development and psychopathology, 24*, 871-888. doi: 10.1017/S0954579412000429.

Tremblay, R.E., Nagin, D.S., Séguin, J.R., Zoccolillo, M., Zelazo, P.D., Boivin, M., Pérusse, D., & Japel, C. (2004). Physical aggression during early childhood: trajectories and predictors. *Pediatrics, 114*, 43-50.

Von Stauffenberg, C. & Campbell, S.B. (2007). Predicting the early developmental course of symptoms of attention deficit hyperactivity disorder. *Journal of Applied Developmental Psychology, 28*, 536-552. doi: 10.1016/j.appdev. 2007.06.011

Wakschlag, L. S., Hill, C., Carter, A. S., Danis, B., Egger, H. L., Keenan, K., …Briggs -Gowan, M.J. (2008a). Observational Assessment of Preschool Disruptive Behavior, Part I: Reliability of the Disruptive Behavior Diagnostic Observation Schedule (DB-DOS). *Journal of the American Academy of Child and Adolescent Psychiatry, 47*, 622-631.

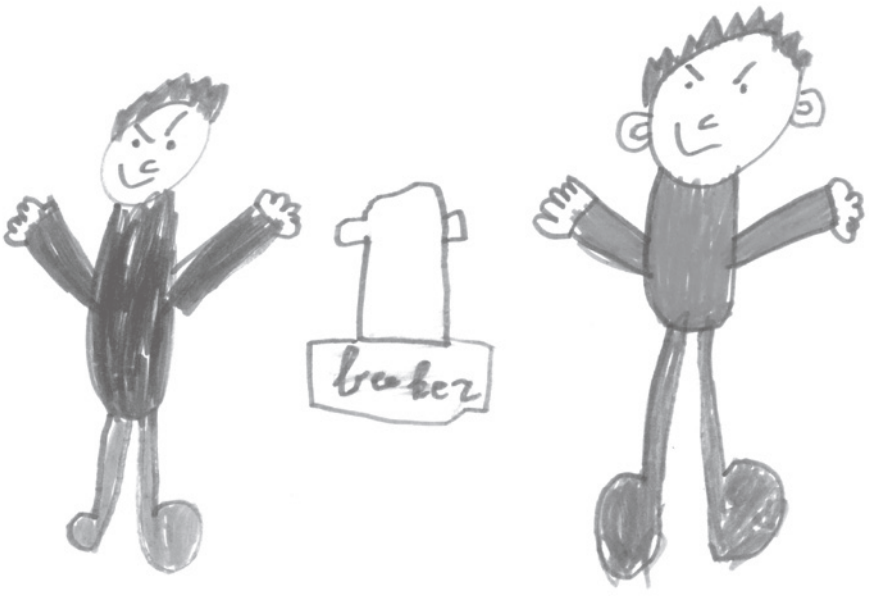
Wakschlag, L. S., Briggs-Gowan, M. J., Hill, C., Danis, B., Leventhal, B.D., Keenan, K., …Carter, A.S. (2008b). Observational Assessment of Preschool Disruptive Behavior, Part II: Validity of the Disruptive Behavior Diagnostic Observation Schedule (DB -DOS). *Journal of the American Academy of Child and Adolescent Psychiatry, 47*, 632-640.

Webster-Stratton, C. (2001). Parenting practices interview. Unpublished assessment instrument. (http://www.son.washington.edu/centers/parenting-clinic/forms.asp).

Webster-Stratton, C., & Reid, M.J. The Incredible Years Parents, Teachers and Children Training Series: A multifaceted treatment approach for young children with conduct problems. In: Weisz J, Kazdin A, editors. Evidence-based psychotherapies for children and adolescents. 2nd edition Guilford; New York: 2010. pp. 194–210.

CHAPTER 6

General Discussion



The first aim of this dissertation was to study the clinical usefulness of the DB-DOS (chapter 2) and the K-DBDS (chapter 3) at the preschool age, to diagnose young children with behavioral problems. The second aim was to study the stability and change of ODD, CD and ADHD diagnosis at the preschool age (chapter 4) and the factors predicting stability of ODD, CD and ADHD diagnosis and persistence of symptoms in preschool children (chapter 5). In this concluding chapter we address the main findings of the studies and provide a general discussion, including some recommendations for future research and interventions.

PSYCHIATRIC ASSESSMENTS OF PRESCHOOL CHILDREN

As outlined in the introduction, there are several challenges related to the diagnostic process of behavioral problems in preschoolers. Therefore we choose to use an observational assessment and a semi-structured interview that were specifically developed for preschool children, i.e., the DB-DOS and the K-DBDS.

DB-DOS

With regard to the DB-DOS (Wakschlag et al., 2008ab), the results of reliability (internal consistency, interrater reliability, test retest reliability) and validity (convergent and divergent) of DBD symptom scores and newly developed ADHD symptom scores are satisfactory for the DB-DOS. This newly developed ADHD scale for the DB-DOS is important, as in the clinical assessment of preschool children with externalizing behavior problems, comorbidity with ADHD should be investigated. The DB-DOS coding system for behavior problems consists of problems in the domains of Behavior Regulation (15 items), and Anger Modulation (6 items). The ADHD domain consists of 10 items organized in terms of problem behaviors related to inattention (4 items), hyperactivity (4 items) and impulsivity (2 items).

While exploring the clinical cutoff point for DBD in the context of the study of the clinical validity on an individual level, the Behavior Regulation total score appeared to have slightly better sensitivity/specificity than the total DBD score. In particular, using the Behavioral Regulation total score, four out of five children, who according to clinical consensus were diagnosed with DBD, will receive a DBD diagnosis. In contrast, the Anger Modulation total score was unsatisfactory. Further, the clinical cutoff point for ADHD symptom scores appeared to be good. Indeed, almost nine out of ten children, who according to clinical consensus were diagnosed with ADHD, will receive an ADHD diagnosis. We should keep in mind that in everyday clinical practice the decision of diagnosing a child with ODD, CD or ADHD is not based on the results of a single measure but on the combination of results

from multiple measures such as a standardized parent and teacher/caregiver rating scales and a (semi-)structured DSM orientated interview with the parents (Matthys & Lochman, 2010). Our main hypothesis was that the DB-DOS may help support the presumption of a diagnosis generated by the information from parents, teachers (or other caregivers) and cognitive assessment. Results suggest that the DB-DOS indeed may be used in the clinical evaluation of children referred with behavior problems.

Coding the videotaped DB-DOS takes about 120 minutes. For clinical purposes, a more practical and less time consuming way of coding should be developed. This seems feasible. For DBD we may limit ourselves to the Behavior Regulation domain. In addition in this domain 5 out of 15 items (verbal aggression (two items), directed aggression, spiteful and sneaky behavior) had an occurrence too low to be evaluated statistically. As a result, these items may be excluded, leaving 10 Behavior Regulation items to be scored in three different contexts. Likewise, for ADHD 10 items should be scored in three different contexts. Thus, 20 items (10 items for DBD and 10 items for ADHD), in three contexts (Parent, Examiner Engaged, Examiner Busy) should be coded. This is similar to the Autism Diagnostic Observation Schedule (ADOS) developed by Lord (Lord et al., 1989). Indeed, the ADOS consists of 28 or 29 items, depending on the language level. We expect that coding the DB-DOS will take about 45 minutes.

K-DBDS

With regard to the clinical usefulness of the K-DBDS (Keenan et al., 2007), the results of reliability (internal consistency, interrater reliability, test retest reliability) and validity (convergent and divergent) are satisfactory. We compared the Qualitative Coding Method, with regard to **ODD**, using “often” as threshold for frequency of behaviors, with a more specific frequency threshold (Specific Coding Method). Apparently, if the threshold is “often” and the follow up probes specify (severe) clinical concern as within the Specific Coding Method, a cutoff score of two ODD symptoms is generated. It may be that the inclusion of clinical concern refers not so much to the presence of symptoms as to the severity of symptoms. Thus, a cutoff score of four ODD symptoms using “often” (according to the Qualitative Coding Method) as threshold for frequency of behavior seems satisfactory.

In DSM-5 (American Psychiatric Association, 2013) for children younger than 5 years a refinement has been made as a note: the behavior should occur on most days. This is less strict than the Specific Coding Method. Indeed, according to the latter a behavior is coded as a symptom not only when it often occurs but also based on follow up probes on the frequency of occurrence, i.e., either many times per day or 1-2 times a day, the latter in combination with a qualitative indicator specifying clinical concern (severity, duration, support and pervasiveness). We agree with the specification in DSM-5 with regard to the

frequency of behavior, i.e., in order to consider behaviors as symptoms they should occur on most days of the week, but not necessarily many times per day or 1-2 times as severe misbehavior.

For **ADHD** in preschool children, results with regard to ADHD in the present study were based on the presence of the numbers of hyperactive-impulsive symptoms only; ADHD subtypes (predominantly hyperactive-impulsive, predominantly inattentive, and combined subtype) may not be the best descriptors of the disorder (Lahey et al., 2005; Lahey & Willcutt, 2010). Lahey et al., (2005) reported that ADHD remains stable over time, but there was considerable instability in the subtypes of ADHD. The distinction between the hyperactive-impulsive and combined subtypes disappeared over time, as most children in the former group shifted to the latter. The distinction between the inattentive and combined subtypes also demonstrated instability, with children shifting from one to another over time. This may be related, in part, to a decline in hyperactive-impulsive symptoms and an increase in inattention symptoms with age, resulting in a later diagnosis of combined subtype for the majority of those initially classified as hyperactive-impulsive subtype (Hardy et al., 2007). Consistent with these findings, the base rate of inattentive symptoms in clinically referred preschoolers has been found to be lower and more variable than the base rate of hyperactive-impulsive symptoms (Murray et al., 2007). It may be difficult to evaluate inattentive symptoms in preschoolers (e.g., "often has difficulty organizing tasks and activities") as parents differ in how they assist their children in organizing tasks and activities. It is also likely that parents notice hyperactive and impulsive behavior at home more readily due to the more salient nature of these behaviors as compared to inattention problems; as a result, parents do not develop a concern about problems with inattention (Massetti et al., 2008).

In summary, the present study shows that ADHD in preschool children may be diagnosed based on a symptom threshold of "a lot" for five ADHD symptoms, while the pervasiveness threshold that the behavior must be endorsed in more than one context does not need to be included. According to DSM-IV criteria for ADHD, a cross situational requirement is needed, although Lahey et al., argued that preschool children with ADHD could be missed because they only have ADHD symptoms in one situation (Lahey et al., 2005).

With regard to DSM-5 (American Psychiatric Association, 2013), the same 18 symptoms are used as in DSM-IV TR (American Psychiatric Association, 2000) and are divided into two symptom domains (inattention and hyperactivity-impulsivity), of which at least six symptoms in one domain are required for diagnosis. One change is that the cross-situational requirement has been strengthened to "several" symptoms in each setting. The present study, however, shows that ADHD in preschool children may be diagnosed while the pervasiveness threshold that the behavior must be endorsed in more than one context does not need to be included.

With regard to **CD** at the preschool age it is unclear how early this disorder can be identified in children (Rolon-Arroy, Arnold & Harvey, 2013). According to Kim-Cohen and colleagues (2005), the DSM-IV criteria for CD diagnosis are sensitive enough to identify young children with CD. The K-DBDS has also been developed to assess CD symptoms. Some of the CD symptoms are adapted with more developmentally appropriate language, whereas four CD symptoms (breaking into a house, car or building; running away from home overnight; often staying out late; truancy) are not assessed because of a lack of face validity in preschool children.

In our study, the number of children diagnosed with CD using the K-DBDS may seem high in a referred sample of preschool children. Yet in the clinical validity study of the K-DBDS only one child was diagnosed with CD according to clinical consensus diagnosis. On the one side, we may have been too reluctant to diagnose children referred with behavior problems with CD. On the other, it may be that the threshold for a CD diagnosis in the K-DBDS is too low; five symptoms might be more appropriate than three symptoms. Another possibility is to describe the CD symptoms within the K-DBDS in a more age appropriate way using the Specific Coding Method for CD (e.g., the frequency thresholds and qualitative indicators). For example for the DSM-IV symptom "often initiates physical fights", the K-DBDS uses the definition "is s/he often physically aggressive such as biting, hitting, kicking or spitting at others". This definition could be more specified using the qualitative indicator "towards others than siblings/parents". This would be in line with Kim-Cohen and colleagues who argued that clinicians wishing to minimize the risk of treating "false positives" can adopt a conservative approach by applying more stringent diagnostic criteria for moderate-to-severe conduct disorder (Kim-Cohen et al., 2005). In DSM-5 (American Psychiatric Association, 2013), however, the criteria for conduct disorder are largely unchanged from DSM-IV and do not consider developmental issues, in contrast to ODD.

STABILITY AND CHANGE OF DIAGNOSIS

The second aim of the study was to examine stability and change of ODD, CD and ADHD diagnoses in the context of functional impairment and symptom severity in the preschool and early school period.

With regard to the perspective of the identification of children with a stable diagnosis results of the study suggest, first, that for both CD and ADHD HI the perspective of number of symptoms was necessary to distinguish the chronic group from the remission groups. Second, for both ODD and ADHD HI diagnostic reassessments were needed in order to identify children belonging to the chronic group. Instead, for CD the chronic group already

could be identified at the first assessment based on a high number of symptoms. The present study is bringing to clinicians' attention that consideration of number of symptoms is an important perspective in view of identifying children with a stable diagnosis of ODD, CD and ADHD HI, as a supplement to the perspective of impairment. Gadow, Kaat and Lecavalier (2013) studied agreement between impairment and other illness parameters (number and severity of symptoms, as well as categorical and dimensional scoring); they showed considerable variation as a function of type of parameter, disorder, and informant. Findings of the present study are consistent with the notion that each illness parameter represents a unique conceptual construct (Gadow, Kaat and Lecavalier, 2013).

With regard to the identification of new cases, the study shows that diagnostic reassessments of preschool children referred for externalizing behavior problems are needed in order to identify new cases. The preschool period is characterized by rapid changes with regard to behavior of the preschool child but also with regard to the environmental demands. For example, the demands at the end of the first year in preschool are different than the demands at the end of the second year. At a practical level, when a preschool child referred for externalizing behavior problems is not diagnosed with a disorder, parents may be advised either psychosocial intervention (Comer, Chow, Chan, Cooper-Vince, & Wilson, 2013) or to get in touch with the clinic once behavior problems or associated impairments increase; a diagnostic reassessment may then be planned.

Previous studies have shown that preschool children's diagnosis of ODD, CD and ADHD are likely to persist into school age, which is important to support the validity of these disorders in preschool children (Keenan et al., 2011; Lahey et al., 2004; Riddle et al., 2013). As a complement to these studies, present findings point to changes in the preschool and early school period. These changes may manifest either in instability of diagnosis, i.e., children with a diagnosis are under remission, or in new cases, i.e., children without a diagnosis at the first assessment are being diagnosed at later assessments. Diagnostic reassessments are needed in order to distinguish children belonging to the chronic stability group from those belonging to the remission groups, and to identify children who did not receive a diagnosis at the first assessment but develop a disorder over time. In the clinical evaluation of children referred for externalizing behavior problems attention should be paid both to impairment and number of symptoms.

ROLE OF INDIVIDUAL AND FAMILY FACTORS

Part of the second aim of this dissertation was to examine which individual and family factors predicted stability of diagnosis within an 18 months period. Schoemaker and colleagues (2012) showed that deficits in executive functions (EF) especially inhibition are

associated with ADHD and DBD in preschool children. The present study shows that poor inhibition performance plays a role in the stability of ADHD and DBD diagnosis. On the other hand, van Lieshout and colleagues conducted a systematic review and showed that there is no evidence that either automatically controlled (requiring little mental effort; lower level), or more consciously controlled (requiring high levels of mental effort; higher level) neurocognitive functions differentiate ADHD persistence from remittance. These authors concluded that neurocognitive deficits can be best seen as epiphenomena, i.e., related to the same etiological factors as the ADHD symptoms, but not directly mediating between etiological factors and phenotype (van Lieshout et al., 2013).

The occurrence of DBD or ADHD HI symptoms at daycare or school was another predictor of diagnostic DBD and ADHD HI stability. In addition, symptoms at daycare or school predicted ODD symptoms at follow-up. Presence of symptoms in more than one setting is an understudied area in preschool children with ODD. Yet, according to DSM-5 (American Psychiatric Association, 2013) severity of the disorder may be specified depending on the presence of symptoms in only one setting (mild), in at least two settings (moderate), or in three or more settings (severe). With regard to ADHD, recently the presence of symptoms in multiple settings has become part of the criteria for ADHD. While according to DSM-IV-TR (American Psychiatric Association, 2000) some impairment of the symptoms in multiple settings was a criterion for ADHD, according to DSM-5 (American Psychiatric Association, 2013) several symptoms must be present in two or more settings. From the perspective of the stability of the ADHD diagnosis in preschool children, results of the study support the need to include presence of symptoms not only at home but at daycare or school as well. Interestingly, the study of the clinical usefulness of the K-DBDS showed that ADHD in preschool children may be diagnosed based on a symptom threshold of “a lot” for five ADHD symptoms, while the pervasiveness threshold that the behavior must be endorsed in more than one context does not need to be included. However, with regard to the prediction of the stability of diagnosis careful assessment of symptoms in other settings than at home is needed. Maybe in our study children with an instable diagnosis had symptoms only in one context.

The clinical evaluation of preschool children referred with behavior problems should include the assessment of parental stress which is strongly associated with ODD, CD and ADHD symptoms at home and at daycare or school. Harsh and inconsistent discipline as a specific predictor of the persistence of ODD symptoms probably refers to the typical coercive interactions (Patterson, Reid & Dishion, 1992) characteristic of the frequently occurring but relatively mild ODD symptoms.

METHODOLOGICAL CONSIDERATIONS

Strenghts, limitations, and future directions

When we started this study we considered several issues. We aimed to capture preschool children with severe behavioral problems and a typically developing group. We also used a longitudinal design to examine the fast changing period in the preschool years with regard to diagnoses and symptom levels but also with regard to individual and environmental factors. Despite these methodological strengths, there are some limitations worth mentioning. First, research questions with regard to clinical validity of the K-DBDS to diagnose CD and ADHD subtypes could not be answered as only one child with CD and only six children with ADHD predominantly inattentive type were diagnosed using the best estimate diagnosis. Second, the age range of the children was quite large (3.5 to 5.5 years); at follow-up assessments, the younger children were still in preschool whereas the older ones were already at elementary school. Third, the relatively small sample size of each group resulted in limited power to detect significant differences between ages or diagnostic groups.

Further research is required to study the predictive validity of the DB-DOS and the stability and change of ODD, CD and ADHD diagnosis over a longer follow-up period than 18 months. It should be considered whether a more practical and less time-consuming way of coding may be developed in view of using the DB-DOS in clinical practice.

OPPORTUNITIES FOR INTERVENTION

Findings of the occurrence of symptoms at daycare or school as a predictor of the stability of diagnosis support the notion that therapeutic interventions at daycare or school are important. Interventions to reduce parental stress and harsh and inconsistent discipline are very essential as well; consensus guidelines recommend that psychosocial interventions constitute first-line treatment of ADHD and DBD in preschool children (Gleason et al., 2007). These interventions have been shown to be effective in young children with disruptive behavior problems (Comer, Chow, Chan, Cooper-Vince, & Wilson, 2013). For example, a recent meta-analysis showed that the Incredible Years parent program is successful in decreasing disruptive child behavior (Menting, Orobio de Castro & Matthys, 2013). Likewise, the effectiveness of the Incredible Years parent and child training programs is established in children diagnosed with oppositional defiant disorder and recently ADHD; treatment effects were reported for children's externalizing, hyperactivity, inattentive and oppositional behaviors, and emotion regulation and social competence. (Webster-Stratton et al., 2011). In addition, Mulqueen, Bartley & Block (2013) showed in

their meta-analysis preliminary evidence that parental interventions are an efficacious treatment for preschool ADHD.
The efficacy of methylphenidate in preschool children with ADHD and comorbid disorders among which ODD has been demonstrated in the Preschool ADHD Treatment Study (Greenhill et al., 2006). The effect size in preschool children who have ADHD and low comorbidity (e.g., only ODD) is similar to the effect size in school-age children (Ghuman et al., 2007).
As an alternative for psychosocial interventions and pharmacological interventions, there has been increasing efforts to target EF deficits in children and adolescents with ADHD. The Cogmed Working Memory Training for elementary school children and adolescents with ADHD was evaluated applying evidence based treatment criteria. This review shows mixed findings and concludes that the training is a "possible efficacious treatment" (Chacko et al., 2012). At this moment it is not advised as a treatment possibility. To our knowledge there is only one study (Halperin et al., 2013) that shows preliminary evidence for play-based intervention (training executive, attention, and motor skills) for preschoolers with ADHD.

GENERAL CONCLUSION

The aim of this dissertation was to study 1. the clinical usefulness of the DB-DOS, the K-DBDS; 2. the stability and change of ODD, CD and ADHD diagnoses at the preschool age; 3. the factors that predict stability of ODD, CD and ADHD in the preschool and early school period. First, the DB-DOS, including the newly developed ADHD part and using a less time-consuming way of coding, and the K-DBDS using the Qualitative Coding Method, are promising assessment procedures in the clinical evaluation of preschoolers referred with behavioral problems. Second, diagnostic reassessments are needed in order to distinguish children belonging to the chronic stability group of ODD and ADHD from those belonging to the remission groups, and to identify children who did not receive a diagnosis at the first assessment but developed a disorder over time. Third, in the clinical evaluation of children referred for externalizing behavior problems attention should be paid both to impairment and number of symptoms. Fourth, results of poor inhibition as a predictor of diagnostic stability suggest that the neuropsychological assessment of inhibition and possibly other executive functions such as working memory become part of the clinical evaluation of preschool children. Fifth, the clinical evaluation of a preschool child referred with behavior problems must include a careful assessment of symptoms in other settings than at home. Sixth, the clinical evaluation of preschool children referred with behavior problems should include the assessment of parental stress and harsh and inconsistent discipline.

REFERENCES

American Psychiatric Association. (2000). *Diagnostic and statistical manual of mental disorders (4th Edition, text-revised)*. Washington, DC: Author.

American Psychiatric Association. (2013). *Diagnostic and statistical manual of mental disorders (5th Edition)*. Washington, DC: Author.

Chacko, A., Bedard, A.C., Marks, D.J., Feirsen, N., Uderman, J.Z., Chimiklis, …Ramon, M. (2012). A randomized clinical trial of Cogmed Working Memory Training in school -age children with ADHD: a replication in a diverse sample using a control condition. *Journal of Child Psychology and Psychiatry*. doi: 10.1111/jcpp.12146.

Comer, J.S., Chow, C., Chan, P.T., Cooper-Vince, C.,, & Wilson, L.A., (2013). Psychosocial treatment efficacy for disruptive behavior problems in very young children: a meta- analytic examination. *Journal of the American Academy of Child and Adolescent Psychiatry, 52,* 26-36.

Gadow, K.D., Kaat, A.J., and Lecavalier, L. (2013). Relation of symptom-induced impairment with other illness parameters in clinic-referred youth. *Journal of Child Psychology and Psychiatry, 54*, 1198-1207.

Ghuman, J. K., Riddle, M. A., Vitiello, B., Greenhill, L. L., Chuang, S. Z., Wigal, S. B., … Skrobala, A. M. (2007). Comorbidity moderates response to methylphenidate in the Preschoolers Attention-Deficit/Hyperactivity Disorder Treatment Study (PATS). *Journal of Child and Adolescent Psychopharmacology, 17*, 563-579.

Gleason, M.M., Egger, H.L., Emslie, G.J., Greenhill, L.L., Kowatch, R.A., Lieberman, A.F., …Zeanah, C.H. (2007). Psychopharmacological treatment for very young children: contexts and guidelines. *Journal of the American Academy of Child and Adolescent Psychiatry, 52,* 26-36.

Greenhill, L., Kollins, S., Abikoff, H., McCracken, J., Riddle, M., Swanson, J., …Cooper, T. (2006). Efficacy and safety of immediate-release methylphenidate treatment for preschoolers with ADHD. *Journal of the American Academy of Child and Adolescent Psychiatry, 45*, 1284-93.

Halperin, J.M., Marks, D.J., Bedard, A.V., Chacko, A., Curchack, J.T., Yoon, C.A., & Healy, D.M. (2013). Training executive, attention, and motor skills: a proof-of-concept study in preschool children With ADHD. *Journal of attention disorders, 17,* 711-21. doi: 10.1177/1087054711435681.

Hardy, K.K., Kollins, S.H., Murray, D.W., Riddle, M.A., Greenhill, L., Cunningham, C., ….Chuang, S.Z. (2007). Factor structure of parent- and teacher-rated attention-deficit/hyperactivity disorder symptoms in the Preschoolers with

Keenan, K., Boeldt, D., Chen, D., Coyne, C., Donald, R., Duax J., …Humphries, M. (2011). Predictive validity of DSM-IV oppositional defiant and conduct disorders in clinically referred preschoolers. *Journal of Child Psychology and Psychiatry, 52*, 47-55.

Keenan, K., Wakschlag, L.S., Danis, B., Hill, C., Humphries, M., Duax, J., & Donald, R. (2007). Further evidence of the reliability and validity of DSM-IV ODD and CD in preschool children. *Journal of the American Academy of Child and Adolescent Psychiatry, 26*, 457-468.

Kim-Cohen, J., Arseneault, L., Caspi, A., Tomás, M.P., Taylor, A., Moffitt, T.E. (2005). Validity of DSM-IV conduct disorder in 41/2-5-year-old children: a longitudinal epidemiological study. *American Journal of Psychiatry, 162*, 1108-17.

Lahey, B.B., Pelham, W.E., Loney, J., Kipp, H., Lee, S.S. & Wilcutt, E. (2005). Instability of the DSM-IV subtypes of ADHD from preschool through elementary school. *Archives of General Psychiatry, 62*, 896-902.

Lahey, B.B., & Willcutt, E.G. (2010). Predictive validity of a continuous alternative to nominal subtypes of attention-deficit/hyperactivity disorder for DSM-V. *Journal of Clinical Child and Adolescent Psychology, 39,* 761-75. doi: 10.1080/ 15374416. 2010.517173

Lahey, B., Pelham, W.E., Loney, D., Kipp, H., Ehrhardt, A., Lee, S, …Massetti, G. (2004). Three-Year predictive validity of DSM-IV attention deficit hyperactivity disorder in children diagnosed at 4-6 years of age. *American Journal of Psychiatry,161*, 2014-20.

van Lieshout, M., Luman, M., Buitelaar, J., Rommelse, N.N., & Oosterlaan, J. (2013). Does neurocognitive functioning predict future or persistence of ADHD? A systematic review. *Clinical Psychological Review, 33,* 539-60. doi:10.1016/j.cpr.2013.02.003.

Lord, C., Rutter, M., Goode, S., Heemsbergen, J., Jordan, H., Mawhood, L., & Schopler, E. (1989). Autism diagnostic observation schedule: a standardized observation of communicative and social behavior. *Journal of Autism and Developmental Disorders*, *19*, 185-212.

Massetti, G.M., Lahey, B.B., Pelham, W.E., Loney, J., Ehrhardt, A., Lee, S.S., & Kipp, H. (2008). Academic achievement over 8 years among children who met modified criteria for Attention-Deficit/Hyperactivity Disorder at 4-6 years of age. *Journal of Abnormal Child Psychology, 36*, 399-410.

Matthys, W., & Lochman, J.E. (2010). *Oppositional defiant disorder and conduct disorder in childhood*. Chichester: Wiley-Blackwell.

Menting, A.T., Orobio de Castro, B., & Matthys, W. (2013). Effectiveness of the Incredible Years parent training to modify disruptive and prosocial child behavior: A meta -analytic review. *Clinical Psychology Review, 33*, 901-913.

Mulqueen, J.M., Bartley, C.A., Bloch, M.H. (2013). Meta-Analysis: Parental Interventions for Preschool ADHD. *Journal of Attention Disorders*, doi: 10.1177/1087054713504135.

Murray, D.W., Kollins, S.H., Hardy, K.K., Abikoff, H.B., Swanson, J.M., Cunningham, C., … Chuang, S.Z. (2007). Parent versus teacher ratings of attention deficit/hyperactivity disorder symptoms in the Preschoolers with Attention -Deficit/Hyperactivity Disorder Treatment Study (PATS). *Journal of Child and Adolescence Psychopharmacology,17*, 605-20.

Patterson, G. R., Reid, J. B., & Dishion, T. J. (1992). *Antisocial boys*. Eugene, OR: Castalia.

Rolon-Arroyo, B., Arnold, D.H., & Harvey, E.A. (2013). The Predictive Utility of Conduct Disorder Symptoms in Preschool Children: A 3-Year Follow-Up Study. *Child Psychiatry Human Development, Aug*. 24. Epub ahead of print.

Riddle, M.A., Yershova, K., Lazzaretto, D., Paykina, N., Yenokyan, G. Greenhill, L. …Posner, K. (2013). The preschool Attention-Deficit/Hyperactivity Disorder treatment study (PATS) 6-year follow-up. *Journal of the American Academy of Child and Adolescent Psychiatry, 52, 264-278*.

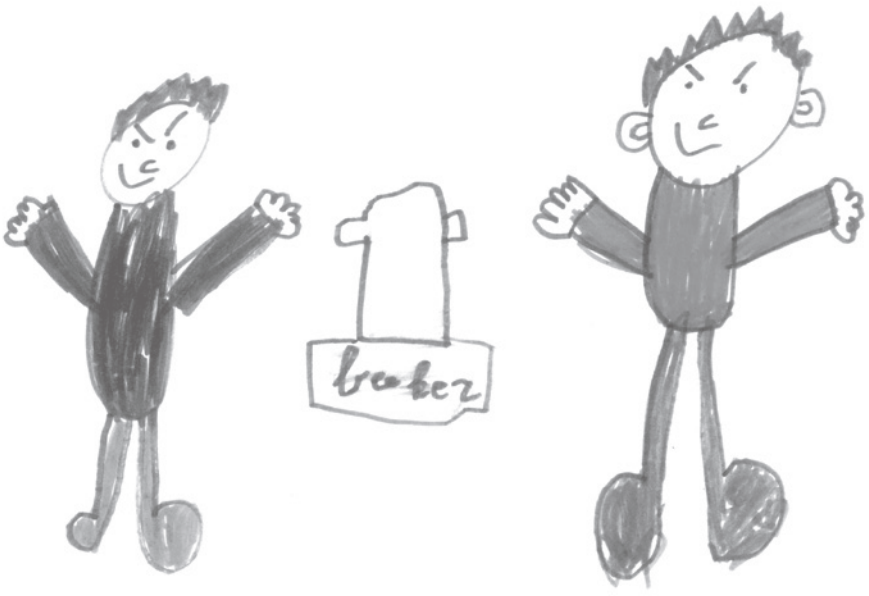
Schoemaker, K., Bunte, T., Wiebe, S.A., Espy, K.A., Dekovic, M., & Matthys, W. (2012). Executive function deficits in preschool children with ADHD and DBD. *Journal of Child Psychology and Psychiatry*, *53*, 111-119.

Wakschlag, L. S., Hill, C., Carter, A. S., Danis, B., Egger, H. L., Keenan, K., …Briggs -Gowan, M.J. (2008a). Observational Assessment of Preschool Disruptive Behavior, Part I: Reliability of the Disruptive Behavior Diagnostic Observation Schedule (DB-DOS). *Journal of the American Academy of Child and Adolescent Psychiatry, 47*, 622-631.

Wakschlag, L. S., Briggs-Gowan, M. J., Hill, C., Danis, B., Leventhal, B.D., Keenan, K., …Carter, A.S. (2008b). Observational Assessment of Preschool Disruptive Behavior, Part II: Validity of the Disruptive Behavior Diagnostic Observation Schedule (DB -DOS). *Journal of the American Academy of Child and Adolescent Psychiatry, 47*, 632-640.

Webster-Stratton, C.H., Reid, M.J., Beauchaine, T. (2011). Combining parent and child training for young children with ADHD. *Journal of Clinical Child and Adolescent Psychology, 40*, 191-203. doi: 10.1080/15374416.2011.546044.

Samenvatting



De laatste jaren is er een toename van verwijzingen van jonge kinderen naar de kinderpsychiatrie. Het is moeilijk om op de kleuterleeftijd (leeftijd van vier tot ongeveer zes jaar; voor het eerst naar groep 1 gaan tot overgang naar groep 3) ongewenst gedrag passend bij de leeftijd te onderscheiden van gedrag dat past bij een kinderpsychiatrische stoornis. Clinici zijn lang terughoudend geweest om op de kleuterleeftijd een kinderpsychiatrische stoornis te diagnosticeren, ook vanuit het oogpunt van mogelijke overdiagnostiek. Dit geldt met name voor de kinderpsychiatrische stoornissen: oppositioneel-opstandige stoornis (ODD, Oppositional Defiant Disorder), (antisociale) gedragsstoornis (CD, Conduct Disorder) en aandachtstekortstoornis met hyperactiviteit (ADHD, Attention-Deficit/Hyperactivity Disorder). De psychiatrische stoornissen ODD en CD worden samen ook wel disruptieve gedragsstoornissen genoemd (DBD, Disruptive Behavior Disorder).
De meeste symptomen van ODD (patroon van opstandig en ongehoorzaam gedrag, driftig zijn, boos zijn en anderen ergeren) zijn zeker niet ongewoon op deze jonge leeftijd; bijvoorbeeld driftbuien komen bij veel zich normaal ontwikkelende kleuters voor. CD wordt gekarakteriseerd door een repetitief en volhardend patroon van gedrag waarin de basisrechten van anderen en/of belangrijke leeftijdspassende sociale normen of regels worden overtreden, zoals vechtpartijen beginnen, liegen, stelen, mensen of dieren mishandelen. Vechtpartijen kunnen op de kleuterleeftijd voorkomen bij zich normaal ontwikkelende kleuters. ADHD wordt gekenmerkt door een patroon van hyperactiviteit (niet stil kunnen zitten, niet kunnen stoppen met praten), impulsiviteit (niet denken maar doen, moeite hebben met op beurt wachten) en/of aandachtsproblemen (snel afgeleid zijn, dingen niet afmaken, alles tegelijk doen); juist deze symptomen passen ook bij zich normaal ontwikkelende kleuters.
Omdat het onderscheid tussen ongewenst gedrag passend bij de kleuterleeftijd en een kinderpsychiatrische stoornis (ADHD, ODD en CD) op deze leeftijd ingewikkeld is, is het belangrijk om gebruik te maken van een combinatie van informatie over het gedrag van het kind van ouders en leerkrachten/verzorgers, en een klinische gedragsobservatie. Deze dissertatie gaat specifiek over het gebruik van een klinische gedragsobservatie (Disruptive Behavior Diagnostic Observation Schedule; DB-DOS) en een semi-gestructureerd interview (Kiddie-Disruptive Behavior Disorder Schedule; K-DBDS) gericht op ODD, CD en ADHD symptomen bij jonge kinderen.
De kleutertijd wordt gekenmerkt door snelle veranderingen in zowel de ontwikkeling van de vaardigheden van het kind als van de eisen/verwachtingen van de omgeving; van een kleuter die net in groep 1 begint, verwacht je minder vaardigheden en stel je minder eisen dan aan een kleuter die bijna naar groep 3 gaat. Het is om die reden belangrijk om na te gaan of diagnosen gesteld in deze ontwikkelingsfase stabiel zijn of niet, en welke kind- en omgevingsfactoren aan die stabiliteit ten grondslag liggen.

Doel van deze dissertatie

Het doel van deze dissertatie was om het klinisch gebruik van de DB-DOS en de K-DBDS te onderzoeken op de kleuterleeftijd voor de diagnostiek van gedragsproblemen. Daarnaast is de stabiliteit en verandering van ODD, CD en ADHD diagnosen op de kleuterleeftijd onderzocht gedurende 18 maanden, evenals welke factoren ten grondslag liggen aan de stabiliteit van ODD, CD en ADHD diagnosen op de kleuterleeftijd.

Spreekuur voor jonge kinderen met gedragsproblemen

Bij de afdeling kinder- en jeugdpsychiatrie van het Universitair Medisch Centrum Utrecht hebben we in 2007 een spreekuur opgezet voor jonge kinderen met gedragsproblemen. Kinderen tussen de leeftijd van 3 ½-5 ½ jaar werden verwezen door kinderartsen, consultatie bureau artsen of huisartsen wanneer er vanwege gedragsproblemen een vermoeden bestond van ODD, CD of ADHD. Deze kinderen werden in een ochtend door een multidisciplinair team onderzocht en werden heronderzocht na 9 en 18 maanden. Binnen dit spreekuur zijn ook normaal ontwikkelende kinderen onderzocht, om deze kinderen te vergelijken met de kinderen die verwezen werden. Deze kinderen werden geworven via kinderdagverblijven en scholen.

Het onderzoek op de eerste onderzoeksochtend bestond uit psychiatrisch en psychologisch onderzoek. Het psychiatrisch onderzoek bestond uit een korte inventarisatie van de klachten met de ouders, de afname van de ontwikkelingsgeschiedenis, vervolgens de gedragsobservatie (DB-DOS) en het semigestructureerd ouderinterview (de K-DBDS). Het psychologisch onderzoek bestond uit de afname van twee taken om de intelligentie in te schatten en zeven executieve functie taken (EF taken); dit zijn taken die de hogere controle functies van de hersenen onderzoeken. De bruikbaarheid in de praktijk van deze executieve functietaken zijn in de dissertatie van Kim Schoemaker beschreven. Verder vulden de ouders een aantal vragenlijsten in over hun eigen opvoedingsvaardigheden, symptomen van stress, depressie en ADHD. De tweede onderzoeksochtend na 9 maanden was korter en bestond voor het kind uit het uitvoeren van EF taken en voor de ouders wederom de K-DBDS en een aantal vragenlijsten (behalve depressie en ADHD vragenlijst voor de ouders). Het derde onderzoeksmoment was gelijk aan het eerste onderzoek met uitzondering van de intelligentietaken en een aantal vragenlijsten (depressie en ADHD vragenlijst voor de ouders).

DB-DOS

De DB-DOS is een in de Verenigde Staten ontwikkelde gestandaardiseerde gedragsobservatie om op de kleuterleeftijd ODD en CD te kunnen vaststellen. De DB-DOS is ontwikkeld om verschillende kwalitatieve en kwantitatieve aspecten van het gedrag van kleuters te onderzoeken. Het gedrag wordt in drie afzonderlijke contexten onderzocht; één ouder

context en twee onderzoeker contexten (onderzoeker is actief en passief). In de drie contexten worden verschillende domeinen van symptomen gescoord (Gedrags Regulatie (Behavior Regulation), Boosheids Modulatie (Anger Modulation) en ADHD). Het domein ADHD is door ons zelf ontwikkeld als toevoeging aan de bestaande domeinen omdat DBD en ADHD vaak samen voorkomen. In **hoofdstuk 2** hebben we onderzocht of de DB-DOS een betrouwbaar en valide onderzoeksinstrument is om op jonge leeftijd DBD en/of ADHD vast te stellen. Er werd onder meer vastgesteld dat, gebruik makend van de Gedrags Regulatie score, vier van de vijf kinderen een DBD diagnose kregen die volgens de klinische consensus diagnose ook een DBD diagnose hadden gekregen. De consensusdiagnose voor de DB-DOS was gebaseerd op scores van vragenlijsten ingevuld door ouders en leerkrachten, het ouderinterview (K-DBDS), de scores van het psychologisch onderzoek en opgenomen beeldmateriaal van het kind tijdens het psychologisch onderzoek. Op grond van deze informatie kenden 2 clinici onafhankelijk van elkaar een diagnose toe (of niet); bij verschil van mening bereikten ze consensus. Voor ADHD gold dat bijna negen van de tien kinderen een ADHD diagnose kregen die volgens de klinische consensus diagnose ook een ADHD diagnose hadden gekregen. Dit onderzoek bevestigt dat de DB-DOS gebruikt kan worden ter ondersteuning bij de diagnostiek van jonge kinderen met gedragsproblemen als zinvolle aanvulling op informatie van ouders en leerkrachten/verzorgers.

K-DBDS

Naast een gedragsobservatie kan gebruik worden gemaakt van een semi-gestructureerd interview met de ouders, de K-DBDS, welke eveneens in de Verenigde Staten is ontwikkeld speciaal voor jonge kinderen met gedragsproblemen passend bij ODD, CD en ADHD. In **hoofdstuk 3** is onderzocht dat de K-DBDS een betrouwbaar en valide onderzoeksinstrument is om de diagnose ODD, CD en ADHD vast te stellen op de kleuterleeftijd. We hebben twee verschillende scoringsmethoden vergeleken; de kwalitatieve scoringsmethode (Qualitative Coding Method), gebaseerd op “vaak” voor de frequentie van gedrag en de specifieke scoring methode (Specific Coding Method), gebaseerd op “vaak” voor de frequentie van gedrag met vervolgvragen om de klinische ernst te specificeren. Daarnaast is gekeken bij hoeveel symptomen het afkappunt zou moeten liggen om wel of geen diagnose (gebaseerd op de klinische consensus diagnose) te stellen voor de beide scoringsmethoden. Voor het stellen van de diagnose ODD op de kleuterleeftijd bleek de kwalitatieve scoringsmethode (met het gebruik maken van “vaak” voor de frequentie van gedrag) en een afkappunt van 4 symptomen voldoende te zijn. Dit komt overeen met het criterium van het aantal symptomen voor ODD zoals vastgesteld in het internationale classificatie systeem voor psychiatrische stoornissen: the Diagnostic and Statistical Manual of Mental Disorders 4th edition (DSM-IV) and 5th edition (DSM-5). Voor ADHD is alleen gekeken naar de hyperactieve en impulsieve symptomen omdat op de jonge leeftijd het lastig

bleek goed onderscheid te maken tussen de verschillende subtypes (hyperactief/impulsief type, onoplettendheid type (aandachtsproblemen) en gecombineerde type). Voor de hyperactieve en impulsieve symptomen bleek dat het afkappunt bij 5 symptomen lag (volgens de DSM-IV en DSM-5 is voor ADHD het minimum aantal symptomen 6) waarbij het niet noodzakelijk bleek te zijn dat de symptomen op meer dan één gebied (bv thuis en school) voorkwamen, hetgeen gedeeltelijk verschilt van de DSM-5 criteria waarbij er enkele symptomen op meer dan 1 gebied moeten voorkomen. Met betrekking tot de diagnose CD kon bovengenoemde niet worden vastgesteld in dit onderzoek omdat er maar bij 1 kind, volgens de klinische consensus diagnose, CD was vastgesteld.

Stabiliteit en verandering van diagnosen

In de studie besproken in **hoofdstuk 4** hebben we de stabiliteit en verandering van ODD, CD en ADHD diagnosen op de kleuterleeftijd en vroege schoolleeftijd onderzocht gedurende 18 maanden. Het bleek dat, van de kinderen aanvankelijk gediagnosticeerd met ODD, bij 62% na 18 maanden nog steeds sprake was van een ODD diagnose; van de kinderen aanvankelijk gediagnosticeerd met CD, bleek bij 35% na 18 maanden nog steeds sprake te zijn van een CD diagnose en voor ADHD bleek bij 59% van de kinderen, waarbij aanvankelijk ADHD was vastgesteld, na 18 maanden nog steeds sprake te zijn van een ADHD diagnose. Naast stabiliteit van diagnosen is tevens sprake van verandering van diagnosen hetzij omdat de diagnose niet meer kon worden gesteld (remissie), hetzij omdat kinderen die bij aanvang van het onderzoek geen diagnose hadden gekregen bij vervolg onderzoek wel een stoornis ontwikkeld bleken te hebben. Om onderscheid te maken tussen een kind met een stabiele diagnose en een diagnose in remissie, bleek het algemeen niveau van (dis)functioneren, zoals ingeschat door de ouders, en het aantal symptomen van belang. Diagnostisch heronderzoek is op deze jonge leeftijd dus van belang om de kinderen te onderscheiden die in de groep zitten met een stabiele diagnose en degenen die remissie van de diagnose hebben. Daarnaast bleek herdiagnostiek van belang omdat er een groep kinderen is, die bij het eerste onderzoek geen diagnose kreeg maar bij vervolg onderzoek wel een stoornis bleek te hebben ontwikkeld.

Rol van individuele en omgevingsinvloeden

Een ander belangrijk doel van deze dissertatie was om te onderzoeken welke individuele en omgevingsinvloeden bij het eerste onderzoek een rol spelen in de stabiliteit van diagnosen na 18 maanden (**hoofdstuk 5**). Tekort aan remming van impulsen zoals vastgesteld met EF taken bleek een rol te spelen in de stabiliteit van ADHD en DBD diagnosen. Daarnaast was de aanwezigheid van DBD of ADHD symptomen op school of op een kinderdagverblijf/peuterspeelzaal een voorspeller voor stabiliteit van DBD en ADHD diagnosen. Stress bij ouders voorspelde eveneens DBD en ADHD diagnosen. Ook werd onderzocht

welke factoren samenhingen met het voorkomen van symptomen over de tijd. Stress bij ouders voorspelde de aanwezigheid van CD en ADHD symptomen, evenals bepaalde opvoedingsvaardigheden (hard en inconsequent opvoeden) de aanwezigheid van ODD symptomen na 18 maanden voorspelden. Voor de hand liggend was de uitkomst dat het aantal symptomen de aanwezigheid van ODD, CD en ADHD na 18 maanden voorspelde.

Aanbevelingen

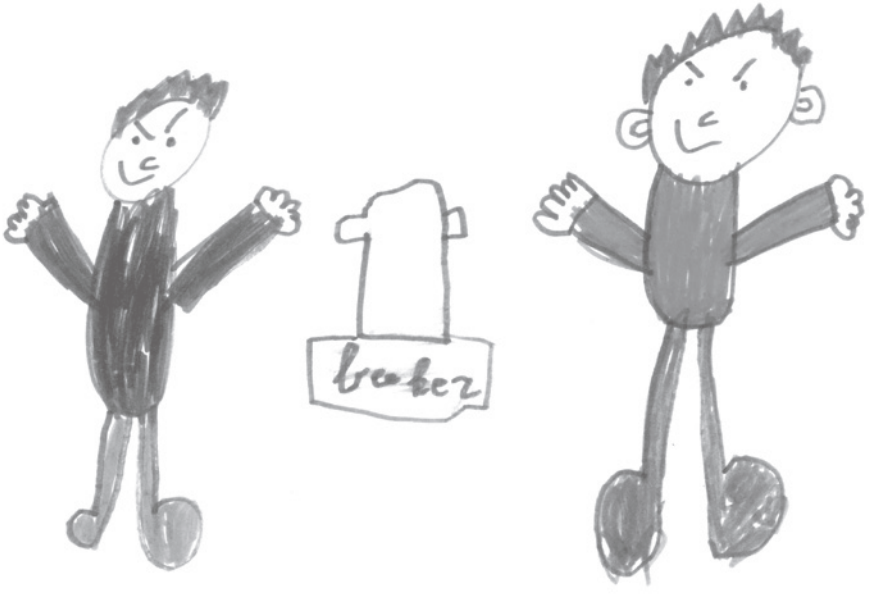
In **hoofdstuk 6** hebben we de resultaten samengevat, bediscussieerd en vergeleken met de resultaten van andere studies. Tevens hebben we sterke punten en beperkingen van het onderzoek besproken. Verder noemen we twee suggesties voor verder onderzoek: de voorspellende waarde van de DB-DOS en de stabiliteit en verandering van diagnosen over een langere periode (3 tot 5 jaar). Dit onderzoek is relevant voor de klinische praktijk. Ten tweede, het ontwikkelen van een meer praktische en tijdbesparende manier van scoren is belangrijk voor de DB-DOS.

De bevinding dat de aanwezigheid van symptomen op school of op het kinderdagverblijf/ de peuterspeelzaal een voorspeller is voor de stabiliteit van diagnosen maakt dat het belangrijk is om ook aldaar interventies te doen. Interventies die stress bij ouders en bepaalde opvoedingsvaardigheden (hard en inconsequent opvoeden) verminderen zijn eveneens van belang. Naast psychosociale interventies en farmacotherapeutische interventies (medicatie voor ADHD) kan mogelijk training van EF een bijdrage leveren aan de behandeling.

Conclusie

Op de kleuterleeftijd zijn de DB-DOS, inclusief het nieuw ontwikkelde ADHD deel, en de K-DBDS veelbelovende onderzoeksinstrumenten voor de diagnostiek van jonge kinderen met gedragsproblemen. Diagnostisch heronderzoek is van belang om onderscheid te maken tussen kinderen die behoren tot de groep met een stabiele diagnose en de kinderen die tot de remissie groep behoren, maar ook met het oog op de identificatie van de kinderen die bij het eerste onderzoek geen diagnose kregen maar een stoornis ontwikkelen over de tijd. Tevens is het belangrijk oog te hebben voor het algemeen niveau van (dis) functioneren en het aantal symptomen. Voor de stabiliteit van diagnosen is het belangrijk om te kijken naar de mate van remming van impulsen, de aanwezigheid van symptomen elders dan de thuissetting, stress bij de ouders, hard en inconsequent opvoeden en het aantal symptomen.

Dankwoord



In de eerste plaats wil ik alle kinderen en hun ouders/verzorgers bedanken die hebben meegewerkt aan het onderzoek. Zonder hen was er geen onderzoek mogelijk geweest!

Walter, mijn eerste ervaring met de kinderpsychiatrie was onder jouw supervisie op Vosseveld. Je enthousiasme voor "het vak" maar ook voor het wetenschappelijk onderzoek naar gedragsstoornissen heb je uitstekend overgebracht! Jouw vermogen tot reflectie, het bekijken van een probleem van meerdere kanten, maar bovenal het altijd bereid zijn tot overleg heb ik zeer gewaardeerd.
Peter, het was goed dat jij vanuit een "statistische" invalshoek meedacht. Jouw heldere kijk op zaken heeft er mede voor gezorgd dat de voortgang erin bleef.

Geachte leden van de ***leescommissie***, professor Buitelaar, professor Dekovic, professor Durston, professor Jongmans en professor Swaab, graag wil ik u hartelijk danken voor uw bereidheid om zitting te nemen in de leescommissie.

I would like to thank ***Laurie Wakschlag*** and ***Kate Keenan*** for the opportunity to use the DB-DOS and the K-DBDS and also for your hospitality in Chicago. ***Kim Espy***, thank you for the opportunity to use the EF tasks.
Dave, dank je wel voor je geduldige uitleg en hulp met de analyses, figuren en tabellen.

Kim, wat heb jij veel geïnvesteerd in het spreekuur en het onderzoek. Zonder jou waren er geen kinderen en controle kinderen geïncludeerd en was er geen databestand geweest! Dank ook voor de fijne samenwerking en de latte machiato's die we hebben gedronken tijdens ons "werkoverleg". Ik heb veel bewondering voor jouw optimisme en doorzettingsvermogen. Ik vind het heel fijn dat je mijn paranimf wilt zijn.
Lidewij, al meer dan de helft van ons leven kennen wij elkaar! Onze eerste kennismaking was in Groningen en hoewel de geografische afstand de laatste jaren alleen maar groter is geworden, zijn jij en Gunnar er altijd voor ons om lief en leed te delen. Heel fijn dat jij als paranimf aan mijn zijde wilt staan!

Sarah, natuurlijk verdien jij het ook om hier apart genoemd te worden! Wat heb jij ook enorm veel werk verzet voor het spreekuur. Onze studiereis naar Chicago zal ik niet snel vergeten! Natuurlijk wil ik ook graag alle onderzoeksassistenten bedanken die hebben meegewerkt om alle data te verzamelen en alle DB-DOS banden te coderen: ***Eva, Justa, Perrine, Susanne en Malou***, veel dank daarvoor. Ook veel dank aan alle 26 ***stagiaires*** die hebben meegewerkt aan het spreekuur en aan de dataverzameling voor dit proefschrift. Het ***spreekuurteam*** wil ik graag bedanken voor hun fijne samenwerking: Aafke, Claudine, Frederieke, Geeke, Lucet, Marte, Rob en Saskia.

Graag wil ik mijn directe ***collegae kinderpsychiaters*** bedanken voor hun collegialiteit en steun in de afgelopen roerige reorganisatie jaren: Elemi, Emma, Floor, Hilgo, Jacob, Manon, Patricia, Peter en oud-collegae: Annerieke, Babette, Bertine, Els, Jeannette, Lidewij, Wieske en Wouter. Een apart woord van dank gaat uit naar Herman van Engeland, Patricia en Peter die de mogelijkheid hebben geboden om naast klinisch werk een promotie onderzoek op te zetten/te doen.
Natuurlijk wil ik het team van de ***dagbehandeling, de poli van het jonge kind en het secretariaat*** bedanken voor hun interesse de afgelopen jaren met de vorderingen van dit proefschrift en hun interesse voor gebeurtenissen buiten dit proefschrift!
Darinka, Jessika, Lidewij, Marieke en Trude, veel dank voor de waardevolle ***intervisie*** momenten.

Lieve ***vrienden/vriendinnen***, ook al noem ik jullie niet allemaal apart, jullie zijn daarom niet minder belangrijk!

Hind, waar zouden wij en de kinderen zijn zonder jou…dat je nog heel veel jaar bij ons wilt oppassen!

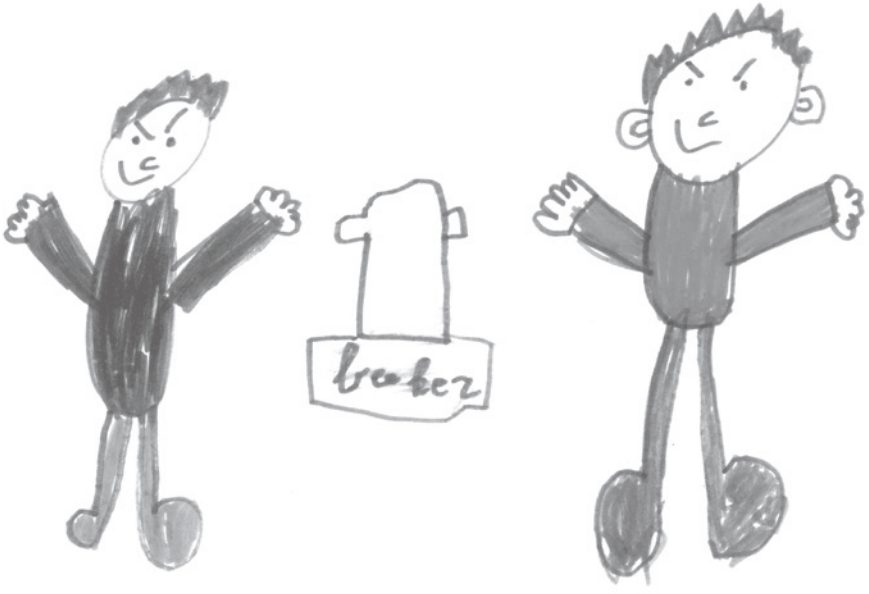
Dick en Tineke, ik had me geen betere schoonouders kunnen wensen! ***Beatrijs en Thomas***, dank voor alle logeerpartijtjes van de kinderen en wat fijn dat jullie wel een pony hebben!
Florine en Andrew, dank jullie wel voor jullie hulp met de Engelse vertaling…op nog veel logeerweekenden in Engeland! ***Andert en Tanya***, wat zijn wij blij dat jullie weer in Nederland zijn komen wonen! Dat we met zijn allen nog heel veel taart gaan eten in Amsterdam en Soest!

Papa en Mama, mede dankzij jullie is het mogelijk geweest een druk gezin, klinisch werk en een promotie onderzoek te combineren! Bedankt voor alle steun, liefde en vertrouwen.

En last but zeker niet least: ***Hidde***: “waarom zou je promoveren, je bent toch al slim genoeg”; wat zijn wij trots op jou als oudste zoon! ***Merle***: wat hebben wij een bewondering voor hoe jij in het leven staat en wat zijn we blij met onze enige dochter! ***Jurre***: ons “middelste mannetje”! Dank je wel voor de prachtige tekening op de voorkant! ***Arne:*** wat zijn we blij dat jij en ***Tibbe*** “uit één ei zijn gekomen” zoals jullie dat zelf noemen…. Wat genieten we van jullie apart maar ook samen!

Lieve Eelko, de laatste zinnen uit het dankwoord zijn natuurlijk voor jou! Jouw support en relaxte levensinstelling helpen mij relativeren en bedenken waar het echt om gaat in het leven!

Curriculum Vitae



Tessa Bunte-Rosingh is op 31 juli 1974 geboren in Curaçao. Zij deed in 1992 eindexamen aan het Stedelijk Gymnasium te Leeuwarden. In 1992 is zij gestart met haar studie medicijnen aan de Rijks Universiteit Groningen. Zij volgde haar coschappen in het St. Elisabeth Hospitaal te Curaçao. In 1999 kreeg zij haar artsenbul (cum laude). Van 1999-2000 heeft zij haar keuzejaar kinderpsychiatrie in het UMC Utrecht bij de afdeling kinder- en jeugdpsychiatrie gevolgd (opleider prof. dr. H. van Engeland), waarna zij aansluitend van 2000-2004 haar opleiding tot psychiater bij stichting Adhesie te Deventer heeft gedaan (opleider dr. L. Timmerman).

In 2004 kwam ze terug naar het UMC Utrecht en werd aldaar staflid, waar ze in 2005 haar aantekening tot kinder- en jeugdpsychiater behaalde (opleider prof. dr. H. van Engeland). Momenteel werkt Tessa als kinder- en jeugdpsychiater bij de poli van het jonge kind en de dagbehandeling van de zorglijn ontwikkelingsstoornissen van de afdeling psychiatrie van het UMC Utrecht.

Tessa is getrouwd met Eelko en samen hebben zij vijf kinderen: Hidde (2003), Merle (2005), Jurre (2007), Arne (2010) en Tibbe (2010).