
Betrouwbaarheid en validiteit van de schalen van disruptief gedrag in de oudermodule van de DB-DOS

Juni 2010

Masterthesis

Universiteit Utrecht – UMC Utrecht

Orthopedagogiek, werkveld Jeugdzorg

M. H. Mierop (3011119) & M. B. Neutel (3327205)

Begeleid door: dhr. prof. dr. W. Matthys

Tweede lezer: drs. K. Schoemaker

Voorwoord

De afgelopen negen maanden hebben wij onze masterthesis geschreven binnen het diagnostisch spreekuur voor jonge kinderen met een vermoeden van disruptieve stoornissen in het UMC Utrecht. De masterthesis is gekoppeld aan het promovendusonderzoek ‘psychiatrische diagnostiek bij jonge kinderen met gedragsproblemen’ van mw. drs. T. Bunte. We willen haar dan ook bedanken voor de mogelijkheid onze thesis aan dit onderzoek te kunnen koppelen en haar succes wensen in het promoveren. In het bijzonder gaat onze dank uit naar dhr. prof. dr. W. Matthys voor de begeleiding en feedback tijdens het schrijven van de thesis. Daarnaast willen we mw. drs. K. Schoemaker bedanken voor haar inzet als tweede lezer. In de laatste, maar zeker niet in de minste plaats willen we mw. drs. S. Laschen bedanken voor haar geduld en feedback tijdens de DB-DOS training.

Mariëtte Mierop & Marjolein Neutel

Abstract

Objective: To examine the internal consistency, the interrater reliability and the validity of the Parent Module of the Disruptive Behavior Diagnostic Observation Schedule (DB-DOS), a new observational method for assessing preschool disruptive behavior.

Method: A total of 107 coded DB-DOS sessions formed the sample, which were coded within three domains: problems in ‘anger modulation’, problems in ‘behavioral regulation’, and ‘competence’. The internal consistency was assessed with Cronbach’s alpha. The interrater reliability was assessed with Interclass Correlation Coefficient. Three groups have been composed on their diagnoses; a group of children with DBD, a group of children with DBD+ADHD and a control group. For assessing the validity from a categorical point of view, the group differences were tested using multivariate analyses of variance (MANOVA), followed by univariate analyses of variance (ANOVA). For assessing the validity from a dimensional point of view the scores of the DB-DOS were correlated with the outcomes of parent and teacher reports about aggressive behavior, global severity and effects on the family.

Results: First, the internal consistency and interrater reliability were high for the domains ‘anger modulation’ and ‘behavioral regulation’. For ‘competence’ the internal consistency was medium and the interrater reliability was low. Second, there was only an effect of group on the domain ‘behavior regulation’. Following univariate analyses of variance revealed that both the DBD group and the DBD + ADHD group scored significantly higher on this domain than the control group. Third, the Pearson correlation coefficients of the domain ‘behavioral regulation’ and CBCL, TRF, CGAS, and IFS were significant. Of the domain ‘anger modulation’ only the IFS was not significant and of the domain ‘competence’ only the IFS father was significant.

Conclusion: The results of this study are promising, considering that the sample is small and the results are from the parent module only. In next studies the other two contexts need to be included in the analyses and the domain ‘competence’ needs further improvement because it’s poor reliability.

Keywords: DB-DOS; disruptive behavior; reliability; validity

Inleiding

Vanuit verschillende epidemiologische onderzoeken is gebleken dat 10%-15% van de kleuters (van 3 tot 6 jaar) milde tot matige gedrags- en emotionele problemen vertonen (Angold & Egger, 2007; Campbell, 1995; Earls, 1980; Richman, Stevenson, & Graham, 1982). Een aanzienlijke meerderheid van de gedrags- en emotionele stoornissen, die bij peuters en kleuters zijn gediagnosticeerd, bestaat uit disruptieve stoornissen (Lavigne et al., 1998). Als gevolg van deze stoornissen kunnen zij versturende gedragingen laten zien die resulteren in een beperking in het functioneren (Keenan & Wakschlag, 2002).

In DSM-IV (American Psychiatric Association, 2000) wordt een categorie aandachtstekortstoornissen en gedragsstoornissen onderscheiden. Onder de gedragsstoornissen vallen twee stoornissen van disruptief gedrag, namelijk 'Oppositional Defiant Disorder' (ODD) en 'Conduct Disorder' (CD). Disruptieve gedragsstoornissen (DBD) worden gedefinieerd als een patroon van negatieve gedragingen die kunnen leiden tot een belemmering in de sociale contacten. Essentieel verschil tussen CD en ODD is dat CD gekarakteriseerd wordt door normoverschrijdend en antisociaal gedrag, waar ODD gekarakteriseerd wordt door terugkomende patronen van negatief, vijandig en openlijk ongehoorzaam gedrag. Onder de aandachtstekortstoornissen valt ADHD, wat wordt gekarakteriseerd door een blijvend patroon van onoplettend en/of hyperactiviteit en impulsiviteit. Er bestaan drie typen ADHD: type onoplettendheid (I), type hyperactiviteit-impulsiviteit (II) en het type gecombineerd (III) (American Psychiatric Association, 2000).

Hoewel DSM-IV CD, ODD en ADHD als drie verschillende stoornissen beschrijft, impliceert de categorisatie niet dat kinderen symptomen van enkel één stoornis vertonen. Zo werd aangetoond dat de comorbiditeit van ADHD, ODD en CD groter is dan verwacht (Loeber & Keenan, 1994). Uit onderzoek van onder andere Loeber, Green, Keenan en Lahey (1995) blijkt namelijk dat 80% van een klinische groep jongens eerder aan de criteria voor ODD voldeden. Hieruit kon worden geconcludeerd dat er een tijdelijke relatie tussen ODD en CD bestaat. Daarnaast wordt door Lahey, Loeber, Quay, Frick en Grimm (1992) een hiërarchische relatie tussen ODD en CD vastgesteld. Kinderen die voldoen aan de criteria van CD voldoen ook aan de criteria van ODD. Jongeren met CD verschillen enkel in het feit dat zij meer ernstiger antisociaal gedrag vertonen dan jongeren met ODD. Doordat er een hiërarchische relatie is, wordt echter enkel CD gediagnosticeerd (Lahey et al., 1992). Tevens voorspelt de aanwezigheid van ADHD het op jonge leeftijd voorkomen van CD (Loeber et al., 1995).

Uit longitudinaal onderzoek is gebleken dat gedrags- en emotionele stoornissen voor de leeftijd van zeven jaar beginnen (Carter, Briggs-Gowan & Davis, 2004). Voor de effectiviteit van de behandeling is het dan ook van belang dat een stoornis vroeg onderkend kan worden (Le Couteur, Haden, Hammal, & McConachie, 2008). Er bestaan echter kritische kanttekeningen bij de vraag of het mogelijk en wenselijk is om gedrags- en emotionele stoornissen te diagnosticeren op kleuterleeftijd. Zo wordt gesteld dat probleemgedrag bij jonge kinderen vaak ligt in de relatie met de ouder en de bredere omgeving en niet zozeer in het kind zelf (Egger & Angold, 2006). Daarnaast wordt genoemd dat gedurende de kleuterleeftijd snelle veranderingen in de ontwikkeling plaatsvinden waardoor het niet mogelijk is om op een valide en betrouwbare manier klinisch gedrag van gedrag binnen de norm te onderscheiden (Egger & Angold, 2006). Zo menen enkele onderzoekers dat oppositionele en agressieve gedragingen van kleuters niet als afwijkend bestempeld zouden moeten worden, gezien deze problemen vaak van voorbijgaande aard zijn. Een meerderheid van de kinderen in de kleuterleeftijd laten vroeg of laat een vorm van agressie en ongehoorzaamheid zien, maar leert dit gedrag ook weer te beheersen (Keenan & Wakschlag, 2002). Echter is er geen bewijs dat de symptomen en/of syndromen van kleuters minder stabiel zijn dan die van oudere kinderen (Angold, Egger & Carter, 2007; Carter et al., 2004). Een andere reden van kritiek is dat individuele verschillen in de normale ontwikkeling onterecht worden gezien als psychiatrische symptomen en dat de classificatiesystemen zoals DSM en ICD niet berekend zijn op deze verschillen. Kortom blijkt met name het onderscheid tussen klinisch gedrag en gedrag binnen de norm een grote rol te spelen in de mogelijkheid en wenselijkheid van vroegdiagnostiek (Egger & Angold, 2006).

Een belangrijk onderscheid tussen gedrag binnen de norm en klinisch gedrag is dat bij gedrag binnen de norm wordt aangenomen dat het ongewenste gedrag in beperkte mate en met name in de thuissituatie voorkomt, er sprake is van relatief flexibel gedrag op het gebied van emotieregulatie en er geen sprake is van disfunctioneren op school en binnen de gezinssituatie. Van klinisch gedrag daarentegen wordt verwacht dat het ongewenste gedrag in verschillende situaties voorkomt, er sprake is van hardnekkig en moeilijk beïnvloedbaar gedrag en het klinisch gedrag leidt tot disfunctioneren op school en binnen de gezinssituatie (Wakschlag & Danis, 2004). Wakschlag & Danis (2004) hebben derhalve geconstateerd dat het bij het onderscheiden van gedrag binnen de norm en klinisch gedrag van belang is de gedragingen te beoordelen op frequentie, intensiteit, hevigheid en contexten waarin het gedrag zich voor doet.

Wanneer rekening gehouden wordt met de manier van beoordelen is vroegdiagnostiek zeer wenselijk. Een belangrijk argument is dat kinderen die op jonge leeftijd externaliserende probleemgedragingen vertonen, een verhoogd risico lopen op het ontwikkelen van internaliserende probleemgedragingen als wel academische moeilijkheden (Campbell, 1995). De overlap van externaliserend en internaliserende symptomen bij kleuters is door verschillende onderzoekers opgemerkt (Campbell, 1995; Rescorla, 1986, zoals geciteerd in Thomas & Guskin, 2001). Zo beschrijft Rescorla dat de comorbiditeit van externaliserende en internaliserende symptomen meer regel dan uitzondering is. Daarnaast is externaliserende problematiek op vroege leeftijd een voorspeller van negatieve uitkomsten op latere leeftijd. Zo bleek uit onderzoek van Moffitt (1990) dat kinderen met een comorbide patroon van 'attention deficit disorder' (ADD) en delinquent gedrag (ADD/D-groep) op 13-jarige leeftijd meer chronische en ernstige gedragsproblemen vertoonden op vijfjarige leeftijd in vergelijking met kinderen die een later begin van gedragsproblemen of ADD zonder gedragsproblemen hadden. Dit onderzoek is samenhangend met onderzoeken gericht op adolescenten, welke hebben aangetoond dat het comorbide ADD/D patroon een betere voorspeller is voor later crimineel gedrag dan delinquentie alleen (Loeber, Stouthamer-Loeber, Van Kammen, & Farrington, 1991). Evenals het onderzoek van Moffitt, suggereert het onderzoek van Patterson (1992, zoals geciteerd in Shaw, Owens, Giovannelli, & Winslow, 2001) dat kinderen die op jonge leeftijd disruptief gedrag vertonen, een grotere kans hebben op meer ernstige criminele gedragingen op latere leeftijd.

Dit is eveneens met longitudinale onderzoeken aangetoond; veel vroege gedrags- en emotionele probleemgedragingen in de kindertijd zetten zich in de late kindertijd en adolescentie voort (Campbell, 1995). Volgens Campbell (1995) hebben kinderen, die tussen de drie en vier jaar probleemgedragingen vertonen, circa 50% kans op het in stand houden van de probleemgedragingen tot in de vroege adolescentie. Lavigne en collega's (1998) toonden in hun onderzoek aan dat bij meer dan 50% van de, door de kinderarts verwezen kleuters, de disruptieve stoornissen 42 tot 48 maanden voortduurden. Ook uit onderzoek van Campbell (1994) bleek dat vroege probleemgedragingen tot op latere leeftijd kunnen voortduren. In zijn onderzoek met peuters hadden de jongens die omschreven werden als 'hard to manage' 67,5% kans dat hun probleemgedragingen tot twee jaar later zouden voortduren, in tegenstelling tot een vergelijkbare groep jongens waarvan maar 18% dezelfde moeilijkheden hadden (Campbell, 1994). Vergelijkbaar hebben studies van ernstige gedrags- en emotionele stoornissen in de vroege adolescentie aangetoond dat deze stoornissen hun

oorsprong vinden in de kleuterjaren (Campbell, 1995). Bovendien is er bewijs vanuit longitudinaal onderzoek dat gedragsproblemen en emotionele probleemgedragingen op driejarige leeftijd met volwassen psychopathologie samenhangt (Caspi, Moffitt, Newman, & Silva, 1996). Vroegdiagnostiek is dan ook van belang om de problematiek vroeg te kunnen aanpakken en daarmee de problematiek op latere leeftijd te voorkomen (Egger & Angold, 2006).

Uit onderzoek van Speltz, McClellan, DeKleyn en Jones (1999) blijkt eveneens dat gedragsstoornissen bij jongens persistent zijn wanneer deze zijn gediagnosticeerd in de kleuterleeftijd (4-5.5 jaar). Zesenzeventig procent van deze kinderen, die in de kleuterleeftijd gediagnosticeerd zijn met ODD, blijkt bij follow-up onderzoek na twee jaar nog steeds te voldoen aan de DSM-III-R criteria van de stoornis. Deze bevindingen ondersteunen de predictieve validiteit van ODD wanneer deze gediagnosticeerd is in de kleuterleeftijd (Speltz et al, 1999).

Naast onderzoek naar de predictieve validiteit van ODD is er meer onderzoek gedaan naar de validiteit van ODD symptomen in de kleuterleeftijd. Zo is er door Keenan en Wakschlag (2004) onderzoek gedaan naar de discriminatieve validiteit van ODD en CD symptomen bij kleuters. Keenan en Wakschlag hebben twee groepen opgenomen in hun onderzoek, namelijk een groep kleuters met een vermoeden van ODD/CD en een groep kleuters zonder een vermoeden van ODD/CD. Beide groepen zijn vergeleken op het aantal aanwezige DSM-IV symptomen van ODD en CD. Hieruit is gebleken dat de groep met een vermoeden van ODD/CD op significant meer symptomen scoorde dan de groep zonder dit vermoeden. Geconcludeerd kan worden dat er discriminatieve validiteit bestaat van de DSM-IV symptomen voor ODD en CD in de kleuterleeftijd (Keenan & Wakschlag, 2004).

Om vast te stellen dat een kind in de kleuterleeftijd aan de DSM-IV criteria voor ODD en/of CD voldoet, zijn er verschillende screenings- en diagnostische instrumenten. Aangenomen wordt dat kinderen die jonger zijn dan tien jaar geen betrouwbare informant zijn met betrekking tot hun gedrag. Daarom is informatie van ouders erg belangrijk bij de diagnostiek van kleuters (Angold & Egger, 2006). De Child Behavior Checklist (CBCL; Achenbach & Rescorla, 2000) is een vragenlijst die hiervoor gebruikt kan worden. Deze vragenlijst wordt door de ouders ingevuld en is bestemd voor de screening van gedrags- en emotionele problemen bij kinderen. De CBCL maakt deel uit van Achenbach System of Empirically Based Assessment (ASEBA). Onder dit systeem valt ook de Teacher Report

Form (TRF; Achenbach & Rescorla, 2000), welke wordt ingevuld door de leerkracht van het kind (Rescorla, 2005).

Ondanks dat ouders veel weten over het gedrag van hun kind, zijn zij meestal moeilijk in staat hier een oordeel over te vellen (Wakschlag et al., 2005). De kennis van een clinicus is nodig om te beoordelen of het gedrag klinisch significant is of bij de leeftijd en ontwikkeling passend. Diagnostische observatie is hier een goed middel voor en vormt dan ook een belangrijke aanvulling op de informatie die ouders geven (Wakschlag, et al., 2008a).

Een bekende diagnostische observatiemethode voor het diagnosticeren van een autisme spectrum stoornis is de Autism Diagnostic Observation Schedule (ADOS). De ADOS is een gestandaardiseerde semi-gestructureerde observatie van communicatief en sociaal gedrag. Er wordt beoordeeld op sociale interactie, communicatie, spel en imaginair gebruik van materiaal (Lord et al., 2000). Bij de ADOS kan gekozen worden uit verschillende modules, waarbij ontwikkelings- en taalniveau bepalend zijn (Le Couteur et al., 2008). De ADOS is als model gebruikt voor het ontwikkelen van een diagnostische observatie methode voor disruptief gedrag, namelijk de Disruptive Behavior Diagnostic Observation Schedule (DB-DOS; Wakschlag et al., 2005).

Het doel van de DB-DOS is om te voorzien in een meetmethode voor disruptief gedrag bij jonge kinderen welke essentiële informatie oplevert voor vroege onderkenning van dit gedrag (Wakschlag et al., 2005). De DB-DOS bestaat uit drie modules, een module waarbij de onderzoeker actieve ondersteuning biedt, een module waarbij de onderzoeker weinig ondersteuning biedt en een module met ouder en kind. Met de DB-DOS worden twee domeinen onderscheiden; de wijze van het moduleren van boosheid van het kind en de wijze van gedragsregulatie van het kind (Wakschlag et al., 2008a). Wakschlag en collega's (2008a) hebben onderzoek gedaan naar de betrouwbaarheid van de DB-DOS op domeinniveau en op itemniveau. Op domeinniveau wordt gekeken naar de gedragsregulatie en het moduleren van boosheid. Op itemniveau wordt gekeken naar de verschillende gedragingen binnen deze twee domeinen. De interraterbetrouwbaarheid is op itemniveau goed te noemen. Hierbij zijn vijf items niet meegenomen, omdat deze minder dan zes keer voorkwamen. Op domeinniveau is gekeken naar de interne consistentie, deze is goed op beide domeinen. De test-hertestbetrouwbaarheid is eveneens goed op beide domeinen (Wakschlag et al., 2008a).

De validiteit van de DB-DOS is gemeten door de resultaten van de DB-DOS te vergelijken met de uitkomsten van andere meetinstrumenten en om disruptief gedrag bij jonge kinderen te meten. Hieruit blijkt dat er een significante positieve correlatie bestaat tussen de

uitkomsten van de DB-DOS en van vragenlijsten welke probleemgedrag meten, zoals de CBCL (Achenbach & Rescorla, 2000) en TRF (Achenbach & Rescorla, 2000). Tussen de DB-DOS en vragenlijsten die positieve gedragingen meten, zoals de Social Skills Rating Scale (SSRS; Gresham & Elliot, 1990) bestaat een significante negatieve correlatie. Deze bevindingen geven sterk bewijs voor de criteriumvaliditeit van de DB-DOS voor het onderscheiden van disruptief gedrag (Wakschlag et al., 2008b).

Het huidige onderzoek wordt uitgevoerd om de betrouwbaarheid en validiteit te meten van de Nederlandse vertaling van de DB-DOS. Tevens wordt dit onderzoek uitgevoerd om uitspraken te kunnen doen over enkel verwezen kinderen ten opzichte van kinderen met normatief gedrag. Dit in tegenstelling tot het bestaande onderzoek van Wakschlag en collega's (2008) waarbij de klinische groep niet enkel bestaat uit kinderen die verwezen zijn vanwege ernstig disruptief gedrag, maar ook uit kinderen die niet verwezen zijn maar wel disruptief gedrag vertonen. In het huidige onderzoek wordt enkel gekeken naar de oudermodule.

De eerste onderzoeksvraag die beantwoord wordt met dit onderzoek, luidt: is de Nederlandse vertaling van de oudermodule van de DB-DOS een betrouwbaar meetinstrument om DBD te diagnosticeren bij kinderen tussen de 3,5 en 5,5 jaar? Verwacht wordt dat zowel de interne consistentie als de interbeoordelaarsbetrouwbaarheid goed zijn, waardoor de Nederlandse vertaling van de oudermodule van de DB-DOS een betrouwbaar meetinstrument is voor het diagnosticeren van DBD bij kinderen tussen de 3,5 en 5,5 jaar in Nederland.

De tweede onderzoeksvraag heeft betrekking op de validiteit en luidt: in hoeverre verschillen de groepen van elkaar op de verschillende domeinscores van de oudermodule van de DB-DOS? Verwacht wordt dat de DBD groep en de comorbide groep (DBD+ADHD) hoger scoren op DBD-gedrag dan de controlegroep.

De derde onderzoeksvraag heeft wederom betrekking op de validiteit en luidt: in hoeverre hangen de geobserveerde problemen van de totale groep op de oudermodule van de DB-DOS samen met de beoordeling van de ouders en de leerkracht op de vragenlijsten over DBD gedrag en de gevolgen hiervan voor het functioneren van het kind en het gezin? Verwacht wordt dat er een significante samenhang is tussen de geobserveerde problemen op de oudermodule van de DB-DOS en de beoordeling van ouders en leerkracht.

Methoden

Selectie en beschrijving van de steekproef

In dit onderzoek zijn 158 Nederlandse kinderen tussen de 3,5 en 5,5 jaar opgenomen, welke zijn opgedeeld in een klinische groep en een controle groep. De klinische groep bestaat uit 113 kinderen die door de huisarts zijn doorverwezen naar het spreekuur voor jonge kinderen met een vermoeden van disruptieve stoornissen van het UMC Utrecht. Deze groep bestaat uit 94 jongens en 19 meisjes, welke zijn verdeeld in een groep met DBD en een comorbidegroep (DBD + ADHD). Deze verdeling is gemaakt op basis van consensus tussen de kinderpsychiater, een orthopedagoog en een stagiaire orthopedagogiek, waarbij gebruik gemaakt is van onder andere de DB-DOS, KDBD, CBCL/TRF en andere vragenlijsten. De klinische groep wordt vergeleken met de controlegroep, welke bestaat uit 45 kinderen waarvan 27 jongens en 18 meisjes, zonder een vermoeden van disruptieve stoornissen. De vergelijking vindt plaats op drie domeinen van de DB-DOS namelijk ‘moduleren boosheid’, ‘gedragsregulatie’ en ‘competentie’. Een deel van de kinderen die zijn gezien bij het spreekuur hebben geen diagnose gekregen. Deze kinderen vallen uit het onderzoek. Het gaat hier om een percentage van 36% (N=41). De totale klinische groep bestond aldus uit 72 kinderen. Voor beschrijvende statistieken zie tabel 1.

Beschrijving meetinstrumenten

DB-DOS

Voor het onderzoek is gebruik gemaakt van de semi-gestructureerde gedragsobservatie Disruptive Behavior Diagnostic Observation Schedule (DB-DOS; Wakschlag et al., 2008a). De DB-DOS bestaat uit drie modules, een module waarbij de onderzoeker actieve ondersteuning biedt, een module waarbij de onderzoeker minimale ondersteuning biedt en een module met ouder en kind. In alle drie de modules ligt de nadruk op gehoorzamen, frustratie en sociaal spel. Hiermee wordt druk uitgeoefend op het kind waardoor disruptief gedrag wordt uitgelokt. In dit onderzoek wordt gebruik gemaakt van de oudermodule. De oudermodule bestaat uit drie taken waarbij de ouder het kind in wisselende mate ondersteunt en één taak waarbij het kind op de ouder moet wachten en zelfstandig moet werken. De ouder wordt voor de afname geïnstrueerd door een orthopedagoog en heeft gedurende de afname een klapper met daarop een beknopte instructie (Wakschlag et al., 2008a).

Voor de beoordeling van de DB-DOS wordt gebruik gemaakt van een coderingssysteem, welke bestaat uit 27 items. Deze items zijn verspreid over drie domeinen, namelijk ‘moduleren boosheid, ‘gedragsregulatie’ en ‘competentie’. Vier onderzoekers hebben aan de hand van deze 27 items filmopnames van de oudermodule van de DB-DOS gescoord, om zo onderscheid te kunnen maken in gradaties van gedrag. De DB-DOS codes integreren zowel kwantitatieve (bijvoorbeeld frequentie) als kwalitatieve (hoe pervasief het gedrag is) dimensies van gedrag. Er wordt onderscheid gemaakt tussen normatief en klinisch probleemgedrag. Een ordinale indeling van scores is gemaakt, met een spreiding van 0 tot 3. In de probleemdomeinen wordt een scoresysteem gehanteerd, welke enerzijds onderscheid maakt tussen typisch (0- en 1-score) en anderzijds atypisch gedrag (2-en 3-score). Een score van 0 houdt in dat er geen sprake is van tekorten of atypische gedragingen, code 1 betekent een milde vorm, waartoe het normatieve wangedrag in deze leeftijdsfase gerekend wordt. Code 2 wijst op duidelijk atypische gedragingen en een score van 3 betreft opmerkelijk atypische gedragingen. Tevens worden een aantal competentie beoordeeld, waarbij code 0 aangeeft dat er geen sprake is van competentie, score 1 een milde vorm van bekwaamheid en score 2 en 3 een duidelijke bekwaamheid. Een aantal items van het domein gedragsregulatie waren op de oudermodule niet van toepassing en werden gescoord met een ‘negen’ (‘niet van toepassing’). De scores leveren een domeinscore op voor elke module (Wakschlag et al., 2008a).

CBCL en TRF

De ouders en de leerkracht van de kinderen uit de onderzoeksgroep vullen respectievelijk de Children Behavior Checklist for Ages 1,5-5 (CBCL/1.5-5; Achenbach & Rescorla, 2000) en de Teacher Report Form for Ages 1.5-5 (TRF/1.5-5; Achenbach & Rescorla, 2000) in. De CBCL en TRF zijn beiden onderdeel van het Achenbach System of Empirically Based Assessment (ASEBA). Dit is een systeem voor screening van gedrags- en emotionele problemen bij individuen. Zowel de CBCL als de TRF bestaat uit 99 vragen. De CBCL bestaat uit gesloten vragen, die ieder met een score van 0 tot en met 2 beantwoord kunnen worden. Hierbij betekent 0 ‘helemaal niet’, 1 ‘een beetje of soms’ en 2 ‘duidelijk of vaak’. In de TRF zijn diverse soorten vragen opgenomen. Er zijn open vragen, gesloten vragen met drie antwoordmogelijkheden; (0) helemaal niet, (1) een beetje of soms, (2) duidelijk of vaak en er zijn een aantal gesloten vragen met vijf antwoordmogelijkheden waarbij (1) staat voor een score onder het gemiddelde niveau en (5) staat voor een score boven het gemiddelde niveau.

De vragen van zowel de CBCL als TRF worden geclusterd tot probleemschalen. Deze schalen zijn: 'Emotioneel Reagerend', 'Angstig/Depressief', 'Lichamelijke Klachten', 'Teruggetrokken', 'Slaapproblemen', 'Aandachtsproblemen' en 'Agressief Gedrag'. De eerste vier probleemschalen vormen samen de schaal 'Internaliseren' en de laatste twee de schaal 'Externaliseren'. Alle vragen over gedrag opgeteld vormen de schaal 'Totale Problemen'. In dit onderzoek zal gebruik gemaakt worden van scores op de probleemschaal 'Agressief Gedrag'. De betrouwbaarheid en validiteit van de CBCL en TRF worden beiden als goed beoordeeld (Achenbach, 1991; Rescorla, 2005).

C-GAS

Naast het invullen van de CBCL wordt door de ouders een score toegekend aan het functioneren van het kind. Dit gebeurt aan de hand van de Children's Global Assessment Scale (C-GAS; Schaffer et al., 1983). De C-GAS is een schaal van 0 tot 100 welke de mate van psychiatrische belemmering, sociale moeilijkheden en het schools functioneren vaststelt bij kinderen tussen de 4 en 16 jaar oud. De schaal is opgedeeld in 10 categorieën welke loopt van 'extreem beperkt' bij een score tussen de 0 en 10 tot 'doet het heel goed' bij een score tussen de 91 en 100. De cut-off score van klinisch en niet-klinisch ligt bij 65 (Bird et al., 1996). De betrouwbaarheid van de C-GAS is gemeten door Dyrborg en collega's (1999) en als goed beoordeeld. De validiteit is eveneens als goed beoordeeld (Bird et al., 1996).

IFS

Ouders vullen tevens de Impact on Family Scale (IFS; Stein & Riessman, 1980) in. Dit is een vragenlijst bestaande uit 24 vragen die de mate van belasting van het gezin meet, welke veroorzaakt wordt door probleemgedrag van het kind (Kolk, Schipper, Hanewald, Casari en Fantino, 2000). De schaal bestaat uit vier subschalen, namelijk 1) financiële lasten, 2) sociale invloed of belemmeringen in het sociale leven, 3) persoonlijke gespannenheid en 4) gezag of coping-strategieën die de ouder gebruikt voor het omgaan met het probleemgedrag van het kind. De vier subschalen kunnen beantwoord worden met scores van nul tot vier; 0 = helemaal oneens, 1 = oneens, 2 = niet zeker, 3 = mee eens, 4 = helemaal mee eens. Vanuit de subschalen wordt een totale score berekend, waarbij een lagere score een hogere belasting betekent. Zowel de betrouwbaarheid en validiteit van de IFS is bij de Amerikaanse populatie als goed beoordeeld. De Cronbachs alpha is tussen de .60 en .88 (Hunfeld, Tempels, Passchier, Hazebroek en Tibboel, 1999).

Onderzoeksprocedure en dataverzameling

Voorafgaand aan de spreekurochtend is schriftelijk toestemming gevraagd aan ouders van de kinderen met en zonder een vermoeden van disruptieve stoornissen. De vragenlijsten CBCL, TRF, C-GAS en IFS zijn vooraf en tijdens de spreekurochtend afgenomen. De DB-DOS is na afloop van een uitgebreid psychologisch onderzoek afgenomen. Van iedere DB-DOS-afname is een filmopname gemaakt. De codering hiervan is gedaan door een team van onafhankelijke codeerders, welke gedurende een periode van gemiddeld vier maanden zijn getraind. Voordat aan de beoordeling van de filmopnames binnen de onderzoeksgroep begonnen kon worden, moest een betrouwbaarheid van tenminste 80% op Amerikaanse en eerdere Nederlandse opnames behaald worden. Dat wil zeggen dat tenminste 80% van de scores toegekend door trainende codeerders, overeen komen met de scores van reeds betrouwbare codeerders. Alle codeerders hebben uiteindelijk op itemniveau een betrouwbaarheid van tenminste 80% bereikt. Het coderen van de dataverzameling is blind gedaan; het was niet bekend of het om een opname van een kind uit de klinische groep of controlegroep ging. Om de interbeoordelingsbetrouwbaarheid te meten is één op de vijf banden dubbel gecodeerd door een Golden Standard. Bij verschil in beoordeling werd er op basis van consensus tot een definitieve score gekomen.

Dataverwerking en data-analyse

Voor het beoordelen van de betrouwbaarheid is allereerst de interne consistentie door middel van de Cronbach's alfa voor elk domein berekend. Vervolgens is de interbeoordelaarsbetrouwbaarheid voor elk domein berekend door het uitvoeren van een Intraclass Correlation Coëfficiënt.

Voor het beoordelen van de validiteit vanuit categoriaal oogpunt zijn voorafgaand verschillen tussen de groepen, voor leeftijd, geslacht en totaal IQ getoetst. Voor het beoordelen van significante verschillen tussen de groepen is een variantie-analyse (ANOVA) op leeftijd en totaal IQ en een Chi-kwadraat-toets voor geslacht uitgevoerd. Indien er verschillen zijn tussen de groepen wordt hiervoor bij de analyses gecorrigeerd.

Voor de validiteit vanuit categoriaal oogpunt is allereerst per domein een gemiddelde score gemaakt om de drie domeinen (moduleren boosheid, gedragsregulatie, competentie) te kunnen vergelijken met de drie groepen (controle versus DBD, controle versus comorbide, controle versus DBD en comorbide). Een gemiddelde score is gemaakt door de scores van de afzonderlijke items per domein bij elkaar op te tellen en de totaalscore te delen door het aantal

afzonderlijke items. Vervolgens is de DBD en comorbide groep samengesteld door de DBD-groep en comorbide-groep bij elkaar op te tellen. Aansluitend hierop is per groep een meervoudige variantie-analyse (MANOVA) uitgevoerd, gevolgd door afzonderlijke univariate variantie-analyses (ANOVA's) met de groep als onafhankelijke variabele en de domeinen als afhankelijke variabelen.

Voor de validiteit vanuit dimensionaal oogpunt zijn allereerst de scores op de C-GAS gehercodeerd. Vervolgens zijn over de gehele groep Pearson correlatie coëfficiënten berekend tussen de scores op de DB-DOS domeinen en de CBCL, TRF, C-GAS en IFS.

Resultaten

Tabel 1 laat zien hoe de groepen naar geslacht, leeftijd en totaal IQ zijn verdeeld. Tevens zijn de resultaten van de variantie-analyse (ANOVA) op leeftijd en totaal IQ én Chi-kwadraat-toets op geslacht in deze tabel opgenomen. Hieruit bleek een significant verschil tussen de groepen op totaal IQ.

Tabel 1. *Beschrijvende statistieken*

		DBD (N=31)	Comorbide (N=41)	Controle (N=45)	F/ X^2	<i>p</i>
Geslacht	Jongen	80.6%	80.5%	60%	5.89	.05
	Meisje	19.4%	19.5%	40%		
IQ		102.35 (11.06)	99.27 (10.02)	111.37 (11.13)	14.65	.00*
Leeftijd in maanden		52.16 (8.44)	54.54 (6.78)	55.47 (7.42)	1.83	.17

Noot. * $p < .05$.

Voor het beoordelen van de interne consistentie van de drie domeinen 'moduleren boosheid', 'gedragsregulatie' en 'competentie' is gebruik gemaakt van Cronbach's alpha. Zoals te zien is in tabel 2 bleken de domeinen 'moduleren boosheid' en 'gedragsregulatie' betrouwbaar. Het domein 'competentie' bleek een matige betrouwbaarheid te hebben.

Ter beoordeling van de interbeoordelaarsbetrouwbaarheid is gebruik gemaakt van de Interclass Correlation Coëfficiënt. Hierbij zijn de domeinscores van de eerste beoordelaar en de Golden Standard vergeleken. Gebleken is dat de interbeoordelaarsbetrouwbaarheid van de

domeinen ‘moduleren boosheid’ en ‘gedragsregulatie’ hoog is (zie tabel 2). Voor het domein ‘competentie’ daarentegen is een lage interbeoordelaarsbetrouwbaarheid gebleken.

Tabel 2. *Betrouwbaarheid*

	Interne consistentie	Interbeoordelaars Betrouwbaarheid
Moduleren boosheid	.94	.81
Gedragsregulatie	.79	.87
Competentie	.63	.46

Voor het beoordelen van de validiteit vanuit categoriaal oogpunt is er per groep (controle versus DBD, controle versus comorbide, controle versus DBD en comorbide) een multivariate toetsing gecorrigeerd voor totaal IQ (MANCOVA), gevolgd door afzonderlijke univariate toetsing (ANCOVA's), uitgevoerd.

Uit multivariate toetsing voor de groep controle versus DBD bleek een significant groepseffect op de drie domeinen, $F(3,71) = 8.55$, $p < .05$. Uit univariate toetsing bleek vervolgens een significant verschil tussen de controlegroep en de DBD-groep op het domein gedragsregulatie (zie tabel 3). Hierbij scoorde de DBD-groep hoger op gedragsregulatie dan de controlegroep.

Uit multivariate toetsing voor de groep controle versus comorbide bleek eveneens een significant verschil tussen de groepen op de drie domeinen, $F(3,81) = 7.07$, $p < .05$. Uit univariate toetsing bleek vervolgens dat de controlegroep significant verschilde van de comorbide groep op het domein gedragsregulatie (zie tabel 3). Hierbij scoorde de comorbide groep hoger op gedragsregulatie dan de controlegroep.

Uit multivariate toetsing voor de groep controle versus DBD en comorbide bleek evenals bij de andere groepen een significant verschil op de drie domeinen, $F(6,222) = 4.73$, $p = < .05$. Uit univariate toetsing bleek vervolgens een significant verschil tussen de controlegroep en de DBD en comorbide groep op gedragsregulatie (zie tabel 3). Hierbij scoorde de DBD en comorbide groep hoger op gedragsregulatie dan de controlegroep.

Tabel 3. *Verschillen tussen controle en DBD op de drie domeinen getoetst met ANCOVA*

	Controle		DBD		ANCOVA	
	M	SD	M	SD	F	<i>p</i>
Moduleren boosheid	.86	.73	1.28	.78	2.66	.107
Gedragregulatie	.50	.36	.91	.30	20.19	.000
Competentie	1.27	.36	1.42	.40	1.91	.171

Tabel 4. *Verschillen tussen controle en comorbide op de drie domeinen getoetst met ANCOVA*

	Controle		Comorbide		ANCOVA	
	M	SD	M	SD	F	<i>p</i>
Moduleren boosheid	.86	.73	1.32	.78	2.36	.129
Gedragregulatie	.50	.36	.97	.45	16.66	.000
Competentie	1.27	.36	1.36	.34	0.22	.639

Tabel 5. *Verschillen tussen controle en DBD en comorbide op de drie domeinen getoetst met ANCOVA*

	Controle		DBD en comorbide		ANCOVA	
	M	SD	M	SD	F	<i>p</i>
Moduleren boosheid	.86	.73	1.30	.78	1.97	.144
Gedragregulatie	.50	.36	.94	.37	12.41	.000
Competentie	1.27	.36	1.38	.37	.83	.437

Voor het beoordelen van de validiteit vanuit dimensionaal oogpunt zijn de drie domeinen vergeleken met een aantal vragenlijsten. Zoals blijkt uit tabel 4 zijn er weliswaar significante maar lage correlaties gevonden tussen het domein ‘moduleren boosheid’ en de CBCL Agressie, TRF Agressie, C-GAS ouder en de C-GAS leerkracht. Er zijn geen significante correlaties gevonden tussen het domein ‘moduleren boosheid’ en de overige vragenlijsten. Op het domein ‘gedragregulatie’ zijn significante correlaties gevonden met alle vragenlijsten, waarbij de correlaties met CBCL Agressie, de IFS ingevuld door moeder en de IFS ingevuld

door vader hogere correlaties hadden dan de overige vragenlijsten. Op het domein ‘competentie’ zijn met uitzondering van de IFS ingevuld door vader geen significante correlaties gevonden.

Tabel 4. *Correlaties domeinen DB-DOS en vragenlijsten*

	CBCL agression	TRF Agression	C-GAS ouder	C-GAS leerkracht	IFS Moeder	IFS vader
Moduleren boosheid	.189*	.256***	.162**	.193*	.128	.122
Gedragsregulatie	.307***	.297***	.275***	.226**	.321***	.333***
Competentie	.124	-.053	.016	.062	.116	.207*

Noot. * $p < .05$. ** $p < .01$. *** $p < .001$.

Discussie

Door middel van dit onderzoek is de betrouwbaarheid en de validiteit van de oudermodule van de DB-DOS als diagnostisch instrument voor kleuters met vermoedens van DBD nagegaan. Na de beoordeling van de interne consistentie en de interbeoordelaarsbetrouwbaarheid is onderzocht of kenmerkende disruptieve gedragingen te onderscheiden zijn bij kinderen met DBD in vergelijking met ‘normale’ controle kinderen. Hierbij zijn de scores van drie klinische groepen met DBD-gedragingen (DBD, comorbide, DBD en comorbide) vergeleken met de scores van een controlegroep.

De interne consistentie bleek goed te zijn voor de domeinen ‘moduleren boosheid’ en ‘gedragsregulatie’. Voor het domein ‘competentie’ daarentegen bleek de interne consistentie onvoldoende. Voor wat betreft de interraterbetrouwbaarheid is eveneens een hoge betrouwbaarheid voor de domeinen ‘moduleren boosheid’ en ‘gedragsregulatie’ en een lage betrouwbaarheid voor het domein ‘competentie’ gebleken. Een mogelijke verklaring voor de matige betrouwbaarheid en lage interbeoordelaarsbetrouwbaarheid van het domein ‘competentie’ is dat het systematisch observeren van positief gedrag wellicht moeilijker is dan het systematisch observeren van probleemgedrag. Positieve gedragingen blijken vaker subtiel te worden geuit dan negatieve gedragingen.

Voor wat betreft de validiteit vanuit categoriaal oogpunt is tussen alle groepen een significant verschil op het domein ‘gedragsregulatie’ gebleken, waarbij de klinische groepen hoger scoorden dan de controlegroep. Hieruit kan worden geconcludeerd dat zowel kinderen

met een DBD diagnose als kinderen met een DBD met comorbide ADHD diagnose naar verwachting meer waarneembare problemen in het reguleren van gedrag laten zien dan kinderen zonder diagnose.

Echter bleek er geen significant verschil te zijn tussen de klinische groepen en de controle groep op het domein 'moduleren boosheid'. Hieruit kan worden geconcludeerd dat kinderen met DBD en DBD met comorbide ADHD tegen de verwachting in even veel problemen laten zien in het moduleren van boosheid als kinderen zonder diagnose. Een mogelijke verklaring hiervoor is dat er slechts gebruik gemaakt is van één module. Het is hierbij een voor de hand liggende gedachte dat kinderen in zulke beperkte tijd minder problemen zullen laten zien met het moduleren van boosheid. Daarnaast is de oudermodule een één op één situatie waarbij het kind ondersteund wordt door de voor hem vertrouwde ouder.

Eveneens bleek er geen significant verschil te zijn tussen de klinische groepen en de controle groep op het domein 'competentie'. Hieruit kan worden geconcludeerd dat kinderen met DBD en DBD met comorbide ADHD, tegen de verwachting in, evenveel competentie laten zien als kinderen zonder diagnose. Een mogelijke verklaring voor deze bevindingen is dat het bij competentie voornamelijk gaat om subtiele gedragingen die bij systematische observaties wellicht minder zichtbaar zijn. Daarnaast hebben de matige interne consistentie en lage interbeoordelaarsbetrouwbaarheid negatieve invloed op de betrouwbaarheid van dit domein.

Voor de validiteit vanuit dimensionaal oogpunt bleek er samenhang te bestaan tussen het domein 'moduleren boosheid' en de CBCL Agression, de TRF Agression, de C-GAS ouder en de C-GAS leerkracht. Hieruit kan worden geconcludeerd dat zowel ouders als leerkracht naar verwachting meer agressief gedrag en een beperkte mate van functioneren rapporteren wanneer kinderen meer problemen laten zien in het moduleren van boosheid. Daarentegen bleek er geen samenhang te bestaan tussen het domein 'moduleren boosheid' en de IFS, ingevuld door zowel vader als moeder. Hieruit kan worden geconcludeerd dat ondanks dat het kind meer problemen in het moduleren van boosheid laat zien, ouders tegen de verwachting in geen hogere belasting in de opvoeding ervaren. Een mogelijke verklaring hiervoor is dat het moduleren van boosheid voornamelijk gaat om een negatieve stemming. Enkel een negatieve stemming zal door ouders niet als belastend worden ervaren.

Voor wat betreft het domein 'gedragsregulatie' bleek een samenhang met alle vragenlijsten te bestaan. Hieruit kan worden geconcludeerd dat zowel ouders als leerkracht naar verwachting meer agressief gedrag en een beperkte mate van functioneren rapporteren,

wanneer het kind meer problemen laat zien in het reguleren van gedrag. Tevens kan geconcludeerd worden dat ouders een hogere belasting in de opvoeding ervaren wanneer het kind meer problemen laat zien in het reguleren van gedrag.

Voor wat betreft het domein 'competentie' bleek een samenhang te bestaan met de IFS ingevuld door vader. Hieruit kan worden geconcludeerd dat vader een hogere belasting in de opvoeding ervaart wanneer het kind minder competent gedrag laat zien. Dit is in overeenstemming met de verwachting. Er is echter, tegen de verwachting in, geen samenhang met de andere vragenlijsten gebleken. Deze uitkomsten moeten echter, vanwege de matige interne consistentie en de lage interbeoordelaarsbetrouwbaarheid van het domein 'competentie', met voorzichtigheid worden geïnterpreteerd. Afwezigheid van samenhang, met uitzondering van de IFS vader, suggereert dat minder competent gedrag niet samenhangt met meer agressief gedrag en een beperkter functioneren. Tevens suggereert dit dat moeder geen hogere belasting in de opvoeding ervaart wanneer het kind minder competent gedrag laat zien. Een mogelijke verklaring waarom er wel samenhang wordt gevonden tussen dit domein en de belasting ervaren door vader, maar niet met de belasting ervaren door moeder is dat vaders in de opvoeding degenen zijn die met kinderen spelen terwijl moeders het zorgaspect voor hun rekening nemen. Vaders zijn dus, meer dan moeders, betrokken bij gezellige en relationele aspecten van de zorg voor de kinderen en zullen hierdoor meer belang hechten aan positief gedrag (Van Dijken & Tavecchio, 1998).

Het huidige onderzoek kent een aantal beperkingen. Een eerste beperking is de relatief kleine onderzoeksgroep (N=107), waardoor er mogelijk geen significante verschillen tussen groepen en/of samenhang tussen de vragenlijsten zijn gevonden. Daarnaast is er in het huidige onderzoek enkel gebruik gemaakt van de oudermodule, waardoor probleemgedrag enkel in interactie met de ouder wordt gesignaleerd. Zo blijkt probleemgedrag bij jonge kinderen vaak in de relatie met de ouder en niet zozeer in het kind zelf te liggen (Egger & Angold, 2006). Door het onderzoek enkel te richten op de oudermodule kan er dus geen onderscheid gemaakt worden tussen het probleemgedrag vanuit het kind zelf en het probleemgedrag in interactie met de ouder. Hierdoor is het enkel in beperkte mate mogelijk om hierover een uitspraak te doen. Een aanbeveling voor vervolgonderzoek is dan ook om de modules met de onderzoeker te betrekken in het onderzoek. Een laatste beperking van dit onderzoek is dat er slechts in beperkte mate een onafhankelijke vergelijking gemaakt kan worden tussen de diagnose en het geobserveerde gedrag. De diagnoses zijn gesteld door middel van onder andere de geobserveerde gedragingen tijdens de DB-DOS. De verdeling van de groepen is dus onder

andere aan de hand van de DB-DOS gemaakt. Aanbeveling voor vervolgonderzoek is om de gedragingen zoals gezien op de DB-DOS niet mee te nemen in de verdeling van de diagnosegroepen. Tevens zal aandacht besteed moeten worden aan de betrouwbaarheid van het domein ‘competentie’.

Tot besluit kan gesteld worden dat in de diagnostiek niet enkel kan worden volstaan met het oordeel van ouders en leerkracht. Ook de clinicus dient het gedrag te beoordelen om af te wegen of er sprake is van klinisch significant gedrag of gedrag binnen de norm (Wakschlag, et al., 2008a). In het huidige onderzoek zijn aanwijzingen gevonden dat de DB-DOS hiervoor een geschikt instrument kan zijn. In de toekomst is het echter nodig om een concrete drempel te leggen tussen klinisch disruptief gedrag en normatief gedrag. Het huidige onderzoek is hiervoor een eerste stap in die richting. Zo kan aan de hand van het huidige onderzoek gesteld worden dat de oudermodule van de DB-DOS gedeeltelijk betrouwbaar en valide is als diagnostisch instrument bij het onderscheiden van afwijkend disruptief gedrag bij kleuters. Om tot een meer betrouwbaar en valide instrument te komen zal in de toekomst allereerst een grotere onderzoeksgroep gebruikt moeten worden. Daarnaast zullen alle domeinen van de DB-DOS moeten worden meegenomen om meer valide uitspraken over het gedrag te kunnen doen. Als laatste zal een diagnose gesteld moeten worden waarbij de gedragingen op de DB-DOS niet zijn meegenomen. Hierdoor is onafhankelijk onderzoek mogelijk.

Literatuur

- Achenbach, T. M., & Rescorla, L. A. (2000). *Manual for ASEBA Preschool Forms & Profiles*. Burlington, VT: University of Vermont, Research Center for Children, Youth, & Families.
- Achenbach, T. M. (1991). *Child Behavior Checklist/4-18*. Burlington, VT: University of Vermont.
- American Psychiatric Association (APA) (2000). *Diagnostic and statistical manual of mental disorders. Fourth Edition Text Revision (DSM-IV-TR)*. Washington, DC: American Psychiatric Association.
- Angold, A., & Egger, H. L. (2007). Preschool psychopathology: lessons for the lifespan. *Journal of Child Psychology and Psychiatry*, 48, 961-966.

- Angold, A., Link Egger, H., & Carter, A. (2007). Measurement of psychopathology in children under the age of six. In: Narrow, W.E., First, M.B., Sirovatka, M.S., & Regier, D.A. *Age and Gender Considerations in Psychiatric Diagnosis: A Research Agenda for DSM-V*. Arlington, VA: American Psychiatric Association, 177-189.
- Bird, H. R., Andrews H., Schwab-Stone M., Goodman, S., Dulcan, M., & Richters, J. (1996). Global measures of impairment for epidemiologic and clinical use with children and adolescents. *International Journal of Methods in Psychiatric Research*, 6, 295-307.
- Campbell, S. B. (1994). Hard-to-manage Preschool Poys: Externalizing Behavior, Social Competence, and Family Context at Two-year Follow-up. *Journal of Abnormal Child Psychology*, 22, 147-166.
- Campbell, S. B. (1995). Behavior problems in preschool children: A review of recent research. *Journal of child psychology and psychiatry*, 36, 113-149.
- Carter, A. S., Briggs-Gowan, M. J., & Davis N. O. (2004). Assessment of young children's social-emotional development and psychopathology: recent advances and recommendations for practice. *Journal of Child Psychology and Psychiatry*, 45 (1), 109-134.
- Caspi, A., Moffitt, T. E., Newman, D. L., Silva, P. A. (1996). Behavioral observations at age 3 years predict adult psychiatric disorders. *Arch Gen Psychiatry*, 53, 1033-1039.
- Dijken, K. S. van, & Tavecchio, L. W. C. (1998). Vaders en de opvoeding en ontwikkeling van kinderen. De pedagogische betekenis van vaders. *Kind en Adolescent*, 19, 68-77.
- Dyrborg, J., Larsen, F. W., Nielsen, S., Byman, J., Nielsen, B. B., & Gautre-Delay, F. (2000). The children's global assessment scale (CGAS) and global assessment of psychosocial disability (GAPD) in clinical practice-substance and reliability as judged by intraclass correlations. *European Child & Adolescent Psychiatry*, 9 (3), 195-201.
- Earls, F. (1980). Prevalence of behavior problems in 3-year-old children. *Archives of General Psychiatry*, 37, 1153-1157.
- 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.
- Gresham F. M., & Elliot, S. N. (1990). *The social skills rating system*. Circle Pines, MN: American Guidance Services.

- Hunfeld, J. A. M., Tempels, A., Passchier, J., Hazebroek F. J. W. en Tibboel, D. (1999). Brief report: Parental burden and grief one year after the birth of a child with a congenital anomaly. *Journal of pediatric psychology*, 24, 215-520.
- Keenan, K. & Wakschlag, L. S. (2002). Can a valid diagnosis of disruptive behavior be made in preschool children? *American Journal of Psychiatry*, 159 (3), 351-358.
- Keenan, K., & Wakschlag, L. S. (2004). Are Oppositional Defiant and Conduct Disorder Symptoms Normative Behaviors in Preschoolers? A Comparison of Referred and Nonreferred Children. *American Journal of Psychiatry*, 161, 356-358.
- Keenan, K., Wakschlag, L. S., Danis, B., Hill, C., Humphries, M., Duax, J., et al. (2007). Further Evidence of the Reliability and Validity of DSM-IV ODD and CD in Preschool Children. *Journal of American Academy of Child and Adolescent Psychiatry*, 46 (4), 457-468.
- Kievit, T., Tak, J. A., Bosch, J. D. (2009). *Handboek psychodiagnostiek voor de hulpverlening aan kinderen*. Utrecht: De Tijdstroom.
- Kolk, A. M. M., Schipper, J. L., Hanewald, G. J. F. P., Casari, E. F., Fantino, A. G. (2000). The Impact-on-Family Scale: A test of invariance across culture. *Journal of Pediatric Psychology*, 25, 323-329.
- Le Couteur, A., Haden, G, Hammal, D., & McConachie, H. (2007). Diagnosing Autism Spectrum Disorders in Pre-school Children Using Two Standardised Assessment Instruments: The ADI-R and the ADOS. *Journal of Autism and Developmental Disorders*, 19 (2), 185-212.
- Lahey, B. B., Loeber, R., Quay, H. C., Frick, P. J., & Grimm, J. (1992). Oppositional defiant and conduct disorders: Issues to be resolved for DSM-IV. *Journal of the American Academy of Child and Adolescent Psychiatry*, 31, 539-546.
- Lavigne, J. V., Arend, R., Rosenbaum, D., Binns, H. J., Christoffel, K. K., Gibbons, R. D. (1998). Psychiatric Disorders with onset in the preschool years, I: stability of diagnoses. *Journal of American Academy of Child and Adolescent Psychiatry*, 37, 1246-1254.
- Loeber, R, Green, S. M., Keenan, K., & Lahey, B. B. (1995). Which boys will fare worse? Early predictors of the onset of conduct disorder in a six-year longitudinal study. *Journal of the American Academy of Child and Adolescent Psychiatry*, 34, 499-509.
- Loeber, R., & Keenan, K. (1994). Interaction between conduct disorder and its comorbid conditions: Effects of age and gender. *Clinical Psychology Review*, 14, 497-523.

- Loeber, R., Stouthamer-Loeber, M., Van Kammen, V., & Farrington, D. P. (1991). Initiation, escalation, and desistence in juvenile offending and their correlates. *Journal of Criminal Law Criminology*, 82, 36-82.
- Lord, C., Risi, S., Lambrecht, L., Cook, E. H., Leventhal, B. L., DiLavore, P. C., et al. (2000). The Autism Diagnostic Observation Schedule-Generic: A Standard Measure of Social and Communication Deficits Associated with the Spectrum of Autism. *Journal of Autism and Developmental Disorders*, 30 (3), 205-223.
- Moffitt, T. (1990). Juvenile delinquency and attention deficit disorder: boys' developmental trajectories from age 3 to 15. *Child Development*, 61, 893-910.
- Rescorla, L. S. (2005). Assesment of Young children using the Achenbach System of Empirically Based Assesment (ASEBA). *Mental Retardation and Developmental Disabilities Research Review*, 11, 226-237.
- Richman, N., Stevenson, J., Graham, P. J. (1982). *Preschool to School: A Behavioral Study*. London: Academic Press
- Shaffer, D., Gould, 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.
- Shaw, D. S., Owens, E. B., Giovannelli, J, & Winslow, E. B. (2001). Infant and Toddler Pathways Leading to Early Externalizing Disorders. *Journal of American Academy of Child and Adolescent Psychiatry*, 40 (1), 36-43.
- Speltz, M. L., McClellan, J., DeKlyen, M., Jones, K. (1999). Preschool Boys with Oppositional Defiant Disorder: Clinical Presentation and Diagnostic Change. *Journal of American Academy of Child and Adolescent Psychiatry*, 38, 838-845.
- Stein, R. E. K. en Riessman, C. K. (1980). The development of an impact-on-family scale: Preliminary findings. *Medical care*, 18, 465-472.
- Thomas, J. M., & Guskin, K. A. (2001). Disruptive Behavior in Young Children: What Does It Mean? *Journal of the American Academy of Child and Adolescent Psychiatry*, 40 (1), 44-51.
- Van Lier, P. A. C., Verhulst, F. C., Van der Ende, J. & Crijnen, A. A. M. (2003). *Classes of disruptive behaviour in a sample of young elementary school children*. *Journal of Child Psychology and Psychiatry*, 44 (3), 377-387.

- Wakschlag, L., & Danis, B. (2004). Assessment of disruptive behavior in young children: A clinical-developmental framework. In DelCarmen-Wiggins, R., & Carter, A.S. (Eds.), *Handbook of Infant, Toddler and Preschool Mental Health Assessment*. New York: Oxford University Press, 421-440.
- Wakschlag, L. S., Leventhal, B. L., Briggs-Gowan, M. J., Danis, B., Keenan, K., Hill, C., et al. (2005). Defining the “Disruptive” in Preschool Behavior: What Diagnostic Observation Can Teach Us. *Clinical Child and Family Psychology Review*, 8 (3), 183-201.
- Wakschlag, L. S., Hill, C., Carter, A. S., Danis, B., Egger, H. L., Keenan, K., et al. (2008A). Observational Assessment of Preschool Disruptive Behavior, Part I: Reliability of the Disruptive Behavior Diagnostic Observation Schedule (DB-DOS). *Journal of American Academy of Child and Adolescent Psychiatry*, 47 (6), 622-631.
- Wakschlag, L. S., Briggs-Gowan, M. J., Hill, C., Danis, B., Benneth, D., Leventhal, L., et al. (2008B). Observational Assessment of Preschool Disruptive Behavior, Part II: Validity of the Disruptive Behavior Diagnostic Observation Schedule (DB-DOS). *Journal of American Academy of Child and Adolescent Psychiatry*, 47 (6), 632-640.