

Evaluatie van Samen sterker Terug Op Pad (STOP4-7): Een Multimodale Interventie voor Kinderen met Gedragsproblemen

Brechtje de Mooij¹, Laura Beurskens-Claessens^{2,3}, Pien Geertse², Wim De Mey⁴ en Geertjan Overbeek⁵

¹Universiteit van Amsterdam, Research Institute of Child Development and Education, programma-groep Forensische Orthopedagogiek & research priority area Urban Mental Health | ²Entrea Lindenhout, Instelling voor Specialistische Jeugdhulp, Speciaal Onderwijs en Onderzoek | ³Praktikon, wetenschappelijk praktijkgericht onderzoeksbureau op gebied van (jeugd)zorg en onderwijs | ⁴Universiteit Gent, Vakgroep Ontwikkelings-, Persoonlijkeids- en Sociale Psychologie; ckg Het Open Poortje | ⁵Universiteit van Amsterdam, Research Institute of Child Development and Education, programma-groep Preventieve Jeugdhulp en Opvoeding

Samenvatting In deze studie onderzochten wij of de multimodale interventie voor kinderen met ernstige gedragsproblemen 'Samen sterker Terug Op Pad' (STOP4-7) zou kunnen leiden tot een afname van internaliserende en externaliserende gedragsproblemen bij kinderen, en een afname van opvoedstress bij ouders. Deze studie gebruikte een herhaalde metingen-design met een voormeting, nameting, en follow-up (6 maanden). De huidige steekproef bestond uit $N = 836$ kinderen van drie t/m acht jaar en maakte gebruik van zowel leerkracht- als ouder rapportage van probleemgedrag. De kinderen waren aangemeld bij entrea lindenhout in Nederland en bij verscheidene Vlaamse jeugdhulpinstellingen. Analyses toonden aan dat STOP4-7 leidde tot significante, klinisch relevante, en blijvende afnames van zowel externaliserend als internaliserend probleemgedrag, en ook leidde tot een significante afname van opvoedstress. De gerapporteerde afnames waren even sterk voor ouders van jongere en oudere kinderen, voor jongens en meisjes, en voor kinderen die bij aanvang van STOP4-7 minder of juist meer problemen hadden.

Trefwoorden STOP4-7, interventie, externaliserend, internaliserend, opvoedstress

Artikelgeschiedenis

Ontvangen: 30 augustus 2021

Geaccepteerd: 24 januari 2022

Online: 3 maart 2022

Contactpersoon

Brechtje de Mooij, l.s.demooij@uva.nl

Copyright

© Author(s); licensed under Creative Commons Attribution 4.0. This allows for unrestricted use, as long as the author(s) and source are credited.

Financiering Onderzoek

Geen.

Belangen

Wim De Mey is ontwikkelaar en uitvoerder van Samen sterker Terug Op Pad. Echter, de analyses werden uitgevoerd door Brechtje de Mooij en Geertjan Overbeek. De eerste versie van dit manuscript werd opgesteld door Brechtje de Mooij. Zodoende is er geen sprake van belangenverstrengeling.

1 Introductie

1.1 De Ontwikkeling van Probleemgedrag

Gedragsproblemen worden gekenmerkt door ongehoorzaamheid, opstandigheid, prikkelbare stemmingen, en door verbale of fysieke agressie richting anderen (DSM-V; APA, 2013). Kinderen die op jonge leeftijd al ongewenst gedrag vertonen, lopen een verhoogd risico op de ontwikkeling van klinische niveaus van internaliserende en externaliserende problematiek op latere leeftijd. Zo liet onderzoek onder volwassenen met een klinische diagnose (variërend van angst en depressie tot schizofrenie) zien dat bij 25–60% al in de kindertijd ongewenst gedrag werd geïdentificeerd (Kim-Cohen et al., 2003). Kinderen met gedragsproblemen hebben vaak moeite met het omgaan met hun emoties, en deze vroege marker van impulsiviteit en een laag zelfregulerend vermogen (Bridgett et al., 2015; Eisenberg et al., 2009) is weer voorspellend voor de ontwikkeling van middelenmisbruik, gezondheidsproblemen, financiële problemen, en delinquentie (Moffitt et al., 2011; Von Stumm et al., 2011) later in het leven. Dit toont het belang van het vroeg identificeren en ingrijpen in mogelijk ongewenst gedrag bij kinderen.

Vroegtijdig ingrijpen bij ongewenst gedrag kan middels het multimodale programma Samen sterker Terug Op Pad (STOP4-7; De Mey & Merlevede, 2013). Het huidige onderzoek draagt bij aan de huidige literatuur middels een evaluatie van de korte- en langetermijneffecten van STOP4-7 op het externaliserend en internaliserend probleemgedrag van kinderen en opvoedstress van ouders. Vernieuwend daarbij is dat gebruik werd gemaakt van zowel ouder- als leerkrachtrapportage om veranderingen in het probleemgedrag van de kinderen in kaart te brengen. Eveneens vernieuwend is bovendien dat wordt gekeken naar de klinische relevantie van waargenomen veranderingen en naar *voor wie* dit programma beter of minder goed werkt.

1.2 Dwingende Ouder-Kind Interactiepatronen

Niet alle jonge kinderen met gedragsproblemen zullen later in de adolescentie antisociaal gedrag vertonen. De kans op blijvende gedragsproblemen neemt met name toe wanneer er in het gezin sprake is van een proces van wederzijds dwingend, in stand houdend gedrag tussen ouder en kind, ofwel een *coercive process*. Wanneer ouders en kinderen elkaar wederzijds negatief bekrachtigen, kan er een neerwaartse spiraal van elkaar in stand houdende interactiepatronen ontstaan. Deze negatieve, dwingende (i.e., coërcieve) interactiepatronen kunnen worden verklaard vanuit de sociale leertheorie zoals uitgewerkt door Patterson (Patterson, 1982; Patterson, 2002; Patterson et al., 1992) en is één van de vroegste en meest krachtige antecedenten van later externaliserend probleemgedrag. Ouders bekrachtigen een kind door het diens zin te geven, en kinderen bekrachtigen het toegefelijke gedrag van hun ouders door te kalmeren nadat zij hun zin hebben gekregen. Ook een gebrek aan warmte in de ouder-kindrelatie en inconsistentie

in het begrenzen van kinderen is gerelateerd aan de ontwikkeling van gedragsproblemen (Mingebach et al., 2020).

Als gevolg van deze dwingende, negatieve interactiepatronen in de gezinscontext missen kinderen bij aanvang van hun schoolloopbaan belangrijke sociaal-cognitieve vaardigheden, zoals het reguleren van emoties, delen van aandacht, rekening houden met anderen, samen spelen, maar ook volhouden en afmaken van een taak (Smith et al., 2014). Bovendien blijkt dat kinderen die ongewenst (i.e., dwars, dwingend, opstandig, agressief) gedrag vertonen vaak over gebrekkige mogelijkheden beschikken om sociale signalen op te vangen en te interpreteren. Zij gaan in sociale situaties weinig op zoek naar feiten en denken al gauw dat anderen vijandige bedoelingen hebben. Voor sociale problemen bedenken ze minder oplossingen en vaker een agressieve oplossing (Orobio de Castro et al., 2002).

Verschillende studies lieten zien dat oudertrainingen een effectief middel zijn om dergelijk gedrag bij kinderen te verminderen of voorkomen (Kazdin & Weisz, 1998; Lundahl et al., 2006; Webster-Stratton & Taylor, 2001). Uit deze literatuur blijkt verder dat met name die programma's behulpzaam zijn die zijn gestoeld op de sociale leertheorie en cognitieve gedragstherapie. Ook bleek dat programma's die zich richten op het ecologische systeem (dus zowel op ouders als kind en leerkracht) effectiever zijn dan wanneer slechts een schakel uit het systeem wordt aangepakt. Multimodale programma's, waarbij naast de ouders ook het kind en de leerkracht betrokken worden, blijken ook effectiever te zijn dan programma's die zich uitsluitend richten op ofwel de ouders, dan wel de kinderen of de leerkrachten (Webster-Stratton & Taylor, 2001).

1.3 Een Multimodale Interventie: Samen sterker Terug Op Pad (STOP4-7)

Een van deze multimodale programma's is Samen sterker Terug Op Pad (STOP4-7; De Mey & Merlevede, 2013). STOP4-7 werd in Vlaanderen ontwikkeld, en wordt al vanaf 2003 in Nederland geïmplementeerd. Het programma, gericht op vier- tot zevenjarige kinderen met gedragsproblemen, is gebaseerd op een vertaling van het Amerikaanse universele preventieprogramma *Linking the Interests of Families and Teachers* (LIFT; Eddy et al., 2000) naar een geïndiceerd programma voor kinderen met ernstige gedragsproblemen, aangepast aan de Vlaamse situatie (De Mey, 2010). Het programma bevat zowel een kind- als een oudertraining, waarin verschillende technieken en vaardigheden worden aangeleerd die helpen gewenst gedrag van kinderen te stimuleren en negatieve interactiepatronen te doorbreken. Ook de leerkracht van het kind neemt deel aan een training en er wordt maatwerk geboden door het geprotocolleerde programma aan te vullen met huis- en schoolbezoeken. In Nederland wordt het programma geïmplementeerd door entrea lindenhout, een instelling voor specialistische jeugdhulp, onderwijs en onderzoek, en door het Kabouterhuis, een orthopedagogisch centrum voor jeugd- en gezinshulp en -ondersteuning. In Vlaanderen wordt het programma door verschillende jeugdinstituten (Centra voor Kinderzorg en Gezinsondersteuning) uitgevoerd.

Zowel in Vlaanderen als in Nederland worden sinds 2003 voormeting-, nameting- en follow-up-gegevens verzameld bij gezinnen en leerkrachten die meededen aan STOP4-7. Uit eerder onderzoek naar STOP4-7 onder 201 Vlaamse kinderen tussen de vier en acht jaar oud met klinische gedragsproblemen bleek dat het externaliserende gedrag zowel thuis als op school was verbeterd na deelname aan STOP4-7. Ook waren ouders minder inconsequent in hun disciplineren (De Mey & Braet, 2011). Een kleinschaliger onderzoek onder 90 kinderen en hun ouders liet vergelijkbare resultaten zien en liet bovendien zien dat veranderingen na STOP4-7 significant groter waren in vergelijking met een wachtlijstcontrolegroep (De Mey, 2010). In Nederland werden soortgelijke resultaten voor STOP4-7 gevonden. Uit onderzoek (Van Brussel et al., 2009) onder 69 kinderen tussen de vier en acht jaar, bleek dat de kinderen en ouders die het programma volgden significante vooruitgang boekten. Na de behandeling vertoonden kinderen volgens de ouders significant minder externaliserend, minder internaliserend en meer prosociaal gedrag. Ouders rapporteerden daarnaast consequenter te zijn geworden in hun opvoeding en minder opvoedbelasting te ervaren. Leerkrachten zagen minder vooruitgang op externaliserend gedrag, maar wel op internaliserend en prosociaal gedrag. Ander onderzoek van Delsing et al. (2010) naar STOP4-7 onder 284 gezinnen met kinderen in de leeftijd van vier tot en met acht jaar (aangemeld bij vier ouder-en-kindcentra in Nederland) liet zien dat de aan STOP4-7 verbonden afnames in probleemgedrag stabiel bleven tot aan een follow-up na zes maanden.

1.4 Langetermijneffecten, Comorbiditeit, en Klinische Relevantie

De eerdere bevindingen impliceren dat STOP4-7 mogelijk leidt tot een afname van externaliserend probleemgedrag bij kinderen, en tot een afname van opvoedbelasting bij ouders. Echter werd niet eerder onder een substantiële steekproef van gezinnen afkomstig uit zowel Nederland als Vlaanderen (vanuit alle jeugdhulplocaties die STOP4-7 aanbieden) middels een multi-informant-benadering gekeken naar de korte- en langetermijneffecten van STOP4-7. Een recente meta-analyse naar – al dan niet multimodale – oudertrainingen (Van Aar et al., 2016) liet zien dat deze over het algemeen significante langetermijneffecten opleveren, tot aan een gemiddelde follow-up-periode van 2,5 jaar. Ook voor STOP4-7 is het van belang te onderzoeken in hoeverre het programma langdurige effecten heeft.

Ondanks dat niet veel onderzoek zich hierop richt, is het van belang om de bredere effecten van de interventie te onderzoeken – niet alleen op externaliserend maar ook op internaliserend probleemgedrag. Programma's als STOP4-7 zijn meestal met name gericht op het voorkomen of verbeteren van externaliserend gedrag. Dit is niet verbazend gezien het feit dat externaliserend gedrag zich manifesteert in interactie (i.e., conflict) met anderen, en het omgaan met ongewenst, naar buiten gericht gedrag mogelijk meer stress oplevert voor ouders dan het omgaan met internaliserend gedrag. Externaliserende gedragsproblemen zijn een van de voornaamste redenen waarom kinderen en

hun ouders worden doorverwezen naar behandeling (Shelleby & Shaw, 2014). Het is echter ook bekend dat comorbiditeit van externaliserend gedrag met internaliserend gedrag eerder de norm is dan de uitzondering voor kinderen met ernstige gedragsproblemen (Overbeek et al., 2001; Van der Giessen et al., 2013). Gebaseerd op hun 'failure model' stellen Capaldi (1992) en Patterson en Capaldi (1990), bijvoorbeeld, dat agressief gedrag en storend gedrag kunnen leiden tot afwijzing en een gebrek aan steun door anderen (ouders, leerkrachten en leeftijdgenoten), wat vervolgens weer kan leiden tot faalervaringen en marginalisatie in sociale situaties en op school. Deze faalervaringen geven een kwetsbaarheid voor depressieve stemmingen. Het zou dus kunnen dat effectief ingrijpen op externaliserend probleemgedrag ook leidt tot een afname van internaliserend probleemgedrag doordat kinderen zich meer gesteund en geaccepteerd voelen. Het zou ook kunnen dat STOP4-7 een direct en uniek effect heeft op internaliserende problemen, bijvoorbeeld omdat kinderen beter met hun gevoelens weten om te gaan en omdat zij zich beter begrepen voelen door hun ouders nadat die de interventie hebben doorlopen.

'Bredere' effecten (i.e., op zowel internaliserend als externaliserend probleemgedrag) van oudertrainingen worden echter niet systematisch gevonden. Zo liet onderzoek naar het *New Beginnings*-programma – gericht op het aanleren van effectief ouderschap en creëren van positieve gezinsinteracties – een vermindering van emotionele problemen bij kinderen zien (Sandler et al., 2011), maar vonden verschillende studies naar *Incredible Years*, een soortgelijke oudertraining, geen effect op internaliserende problemen (Leijten et al., 2018; Overbeek et al., 2021). In het huidige onderzoek wordt gekeken in hoeverre STOP4-7, naast het verwachte effect op externaliserend probleemgedrag, een uniek effect heeft op internaliserend probleemgedrag van kinderen, ondanks dat STOP4-7 zich niet direct richt op internaliserend probleemgedrag.

Verder is het van belang te onderzoeken in welke mate een statistisch significante afname van probleemgedrag ook daadwerkelijk klinisch relevant is (Kazdin, 1999; La Greca et al., 2009). Alhoewel klinische relevantie in het overgrote merendeel van interventiestudies niet wordt vastgesteld, gaf eerder onderzoek naar STOP4-7 aan dat voor 40% van de ouders het probleemgedrag van hun kind afneemt tot onder een klinische grenswaarde, en voor 60% van de ouders de opvoedbelasting afneemt tot een niet-klinisch niveau (Delsing et al., 2010).

Onderzoekers riepen al eens op tot onderzoek naar moderatoren van interventie-effectiviteit voor programma's die zich richten op het verminderen van probleemgedrag bij kinderen (zie bijvoorbeeld Shelleby & Shaw, 2014). Zodoende is het ook van belang te onderzoeken *voor wie* STOP4-7 de sterkste – of juist de minst sterke – effecten laat zien. Met zulke resultaten kan de interventie beter op maat worden gesneden voor verschillende subgroepen van kinderen en gezinnen. Een recente meta-review (Overbeek et al., 2017) concludeerde dat de effecten van preventieve opvoedinterventies – hierin werden ook geïndiceerde interventies bij klinische risicogroepen meegenomen – duidelijk generaliseerbaar zijn over verschillende doelgroepen naar sociaaleconomische

status, opleidingsniveau, etniciteit, en het geslacht en de leeftijd van kinderen. De effecten van opvoedinterventies bleken wel sterker te zijn voor gezinnen met verhoogde risiconiveaus. Zo liet een recente *individual participant data*-meta-analyse (op basis van gepoolde datasets) zien dat de ernst van de problematiek voorafgaand aan de interventie de interventie-effecten van een oudertraining modereerde, met sterkere effecten voor gezinnen waarbij meer opvoedproblematiek of gedragsproblemen aanwezig waren (Leijten et al., 2018). Het is dus van belang, en een duidelijke meerwaarde van dit onderzoek, om te kijken naar modererende factoren die samenhangen met gedragsverbeteringen na een interventie zoals STOP4-7, zoals het geslacht en de leeftijd van kinderen, en met name ook de ernst van gedragsproblemen bij de start van de interventie. Eerder onderzoek suggereert dat jongens mogelijk een groter risico lopen op de ontwikkeling van externaliserende gedragsproblemen dan meisjes (zie Lahey et al., 2006 voor een overzicht) en mogelijk geldt ook dat multimodale interventies differentiële effecten laten zien voor jongens en meisjes. Kinderen met hogere niveaus van gedragsproblemen op jonge leeftijd lopen het grootste risico voor een problematische ontwikkeling op de lange termijn, dus met name voor deze subgroep van kinderen is het van belang te onderzoeken in hoeverre een programma als STOP4-7 effectief is (Shelleby & Shaw, 2014).

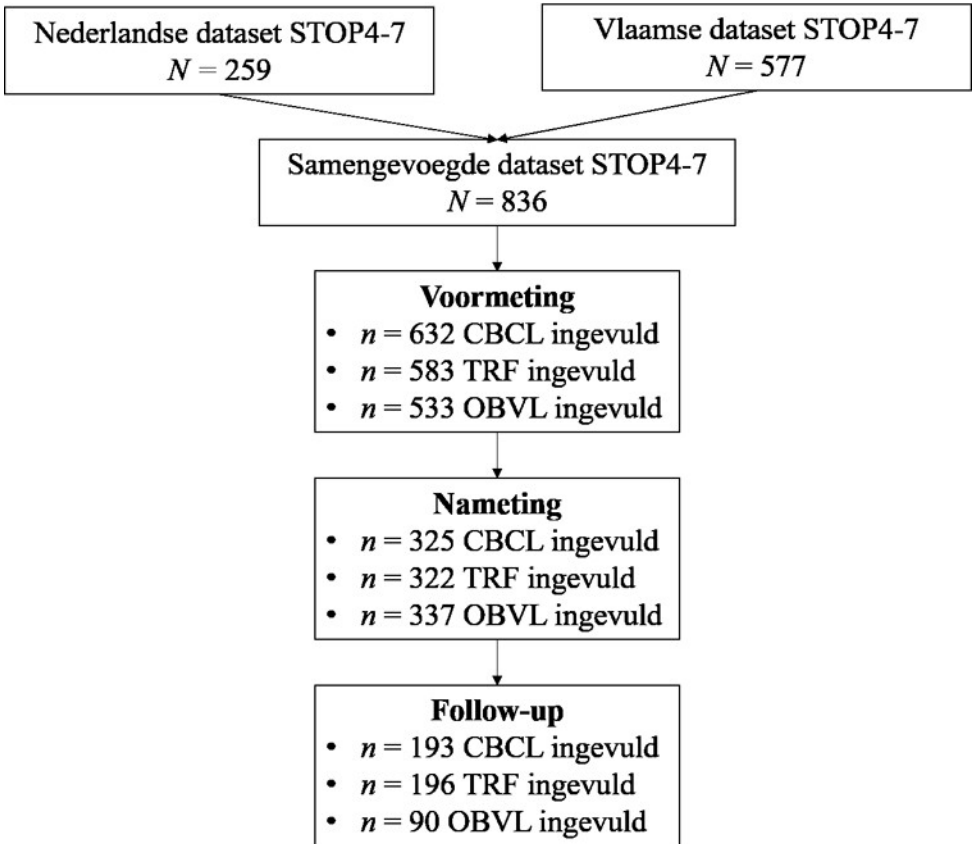
1.5 Onderzoeksvragen en Hypothesen

Op basis van eerder onderzoek verwachten wij dat de multimodale interventie STOP4-7 leidt tot een afname van zowel externaliserend als internaliserend probleemgedrag en de door ouders gerapporteerde opvoedstress van voor- naar nameting. Ook wordt verwacht dat veranderingen na het programma in stand blijven tot aan de follow-up na zes maanden. Verder verwachten wij dat STOP4-7 mogelijk differentiële – relatief sterke of juist minder sterke – effecten heeft voor subgroepen van gezinnen, afhankelijk van het geslacht en de leeftijd van kinderen. Bovendien wordt verwacht dat gerapporteerde veranderingen in gedragsproblemen groter zijn bij gezinnen die bij aanvang van de interventie ernstigere problemen ervaren. De mogelijke klinische relevantie van afnames in probleemgedrag na STOP4-7 wordt exploratief onderzocht, omdat er voor STOP4-7 nog geen of slechts beperkte eerdere resultaten op dit vlak beschikbaar zijn. Omdat onze onderzoeksgegevens uit Nederland en Vlaanderen kwamen, hebben we daarnaast gecontroleerd of de effecten van STOP4-7 vergelijkbaar waren in deze twee landen.

2 Methode

2.1 Design

Om het effect van STOP4-7 te analyseren, werden twee bestaande datasets samengevoegd. Deze datasets bevatten data die in Nederland en in Vlaanderen werd verzameld in het



Figuur 1 Flowchart van beschikbare data op de drie meetmomenten

kader van *routine outcome monitoring*. Data was beschikbaar voor drie meetmomenten: voormeting, nameting en follow-up. De voormeting vond één maand voor implementatie van STOP4-7 plaats. De nameting vond plaats op het eind van de groepstraining (drie tot vier maanden na de start van de interventie), en de follow-up vond zes tot zeven maanden na implementatie van STOP4-7 plaats. Dit onderzoek bevatte geen controlegroep. Figuur 1 geeft het verloop van de studie en de respons op de drie meetmomenten weer.

2.2 Participanten

De huidige studie analyseerde de gegevens van $N = 836$ ouders met kinderen in de leeftijd van drie tot en met acht jaar die de interventie STOP4-7 hebben doorlopen. Van de totale dataset werd $n = 259$ (31%) verzameld in Nederland en $n = 577$ (69%) in Vlaanderen. Alle ouders gaven passieve toestemming voor het gebruik van hun geanonimiseerde gegevens voor wetenschappelijk onderzoek. In Vlaanderen werd eenzelfde toestem-

mingsprocedure gebruikt, goedgekeurd door het Ethisch Comité van de Universiteit Gent. De groep kinderen bestond voor het merendeel uit jongens ($n = 683$, 82%). De gemiddelde leeftijd van de kinderen was 5,50 jaar ($SD = 1.14$). Precieze data over de leeftijd van kinderen was voor slechts 226 Nederlandse kinderen (27% van de gehele dataset) bekend. Wel is bekend dat de leeftijd van alle kinderen in de steekproef varieerde van drie tot acht jaar. Van kinderen waarvoor data op voormeting beschikbaar was, rapporteerden de ouders van iets meer dan de helft van de kinderen in de steekproef (57,6%, $n = 364$) bij aanvang van de studie klinische niveaus van probleemgedrag (CBCL T-score totaal probleemgedrag ≥ 63). Demografische gegevens van ouders en leerkrachten was niet beschikbaar voor de huidige studie. Precieze data over het aantal ouders dat het STOP4-7 programma niet geheel doorliep is niet bekend voor de huidige dataset. Wel is bekend dat gemiddeld genomen slechts één op tien tot zestien gezinnen vroegtijdig stopt met STOP4-7 en het programma niet afmaakt (De Mey & Geertse, persoonlijke communicatie).

2.3 Samen Terug Op Pad (STOP4-7)

De doelgroep van STOP4-7 bestaat uit kinderen van vier tot en met zeven jaar en hun opvoeders. Het gaat om kinderen met ernstige gedragsproblemen zoals liegen, vechten, ongecontroleerde driftbuien, en een ernstige mate van ongehoorzaam en opstandig gedrag. Er is in deze gezinnen vaak sprake van een dwingende cyclus tussen ouder en kind, waarbij ouders en kinderen in een wederzijds dwingend gedragspatroon verstrikt zijn geraakt en er geen duidelijke gezagsrelatie meer is binnen het gezin. Ouders rapporteren hierbij vaak onvoldoende opvoedvaardigheden en ervaren vaak een hoge opvoedingsbelasting. Het programma STOP4-7 is gebaseerd op leertheoretische principes, welke stellen dat het gedrag van kinderen wordt gevormd door belonings- en strafsignalen in de opvoedingscontext waar zij aan blootgesteld worden (Reid et al., 2002).

STOP4-7 is een multimodale groepstraining voor kinderen, ouders en leerkrachten. De interventie is geprotocolleerd, met een aantal doelstellingen en methoden die per bijeenkomst nauwkeurig omschreven zijn. De verfijning 'op maat' wordt aangebracht met een aantal huis- en schoolbezoeken (De Mey & Merlevede, 2013). De kindtraining vindt plaats gedurende tien weken, één dag per week en in een groep van vier tot acht kinderen. De activiteiten verschillen niet veel van die in een gewone kleuterklas. Het verschil zit in de gedragstherapeutische attitude van de trainers (leerkrachten en/of pedagogisch medewerkers) en de gerichtheid op het aanleren van concrete vaardigheden aan de kinderen. Met behulp van specifieke opvoedvaardigheden, zoals positieve bekrachtiging (complimenten, belonen), modellering, negeren en het toepassen van time-outs, wordt gewenst gedrag van kinderen versterkt en hun ongewenst gedrag uitgedoofd. In de training wordt de kinderen verder gewenst gedrag aangeleerd, zoals gevoelens herkennen en benoemen, rekening houden met de ander, problemen oplossen, zelfcontrole (sociaal-cognitieve vaardigheden), luisteren en samenspelen.

De oudertraining bestaat uit tien wekelijkse groepssessies van twee uur en beoogt het vergroten van de opvoedvaardigheden van ouders. Aan de orde komen onder meer operante leerprincipes, modelling, complimenten geven, beloningssystemen gebruiken, grenzen stellen, negeren en time-outs toepassen, evenals actief luisteren. Ouders worden aangemoedigd zoveel mogelijk zelf te proberen en te oefenen. De leerkrachten (en indien wenselijk andere opvoeders van het kind, zoals groepsleiders) krijgen vier sessies waarin een vergelijkbaar programma wordt doorlopen.

2.4 Meetinstrumenten

2.4.1 Gedragsproblemen – Ouderrapportage

Gedragsproblemen bij kinderen werden gemeten middels de Nederlandse versie van de *Child Behavior Checklist* (CBCL; Achenbach, 1991a; Verhulst et al., 1996). De CBCL werd door een van de ouders of verzorgers (i.e., pleegmoeder) of door ouders samen ingevuld. De CBCL meet zowel internaliserend als externaliserend probleemgedrag en sociale competentie, en wordt ingevuld door ouders. De CBCL bevat 120 items, welke worden beantwoord op een driepuntschaal (0 = *helemaal niet* tot 2 = *duidelijk of vaak*). De psychometrische kwaliteit van de CBCL is in voorgaand onderzoek aangetoond (Evers et al., 2001); de interne consistentie in de huidige dataset was goed (Totaal probleemgedrag: $\alpha = 0,84$; Internaliserend probleemgedrag: $\alpha = 0,87$; Externaliserend probleemgedrag: $\alpha = 0,80$).

Wanneer beide ouders afzonderlijk rapporteerden over het probleemgedrag van hun kind, werd voor de analyses die ouder geselecteerd voor wie de data het meest compleet was (i.e., voor wie er op meer meetmomenten data beschikbaar waren). In die gevallen werd er per kind dus informatie van één opvoeder meegenomen in de huidige analyses. Wanneer beide ouders evenveel meetmomenten invulden werd er willekeurig een keuze gemaakt tussen ouders. Op deze manier werd onnodig verlies van data voorkomen en ontstond er geen afhankelijkheid in de data. Voor $n = 153$ kinderen werd door beide ouders over gedragsproblemen gerapporteerd, voor $n = 468$ kinderen werd door de moeder gerapporteerd en voor $n = 84$ kinderen werd door de vader gerapporteerd. Voor $n = 9$ kinderen werd door een andere verzorger over gedragsproblemen gerapporteerd (bijvoorbeeld oma of een gezinshuismoeder) en voor $n = 122$ kinderen was geen ouderrapportage beschikbaar.

2.4.2 Gedragsproblemen – Leerkrachtrapportage

Naast de CBCL werd ook informatie bij leerkrachten opgehaald over het probleemgedrag van kinderen middels de *Teacher's Report Form* (TRF; Achenbach, 1991b; Verhulst et al., 1997). Dit is de leerkrachtversie van de CBCL. Van de TRF werd eveneens de subschaal Internaliserend probleemgedrag (35 items) en Externaliserend probleemgedrag (34 items) meegenomen. Items worden beantwoord op een driepuntschaal (0 = *helemaal niet* tot 2 = *duidelijk of vaak*). De psychometrische kwaliteit van de TRF is in voorgaand onderzoek aangetoond (Evers et al., 2001); de interne consistentie in de huidige dataset

was acceptabel (Totaal probleemgedrag: $\alpha = 0,76$; Internaliserend probleemgedrag: $\alpha = 0,74$; Externaliserend probleemgedrag: $\alpha = 0,77$). Data van door leerkrachten gerapporteerde gedragsproblemen was beschikbaar voor in totaal $n = 710$ kinderen. Voor $n = 595$ van deze kinderen was er ook data van door ouders gerapporteerde gedragsproblemen beschikbaar. Voor $n = 115$ kinderen was er enkel data beschikbaar voor door leerkrachten rapporteerde gedragsproblemen.

2.4.3 Opvoedstress

De door ouders ervaren opvoedbelasting werd gemeten middels de *Opvoedbelasting Vragenlijst* (OBVL; Vermulst et al., 2015). De OBVL bevat 34 items verdeeld over 5 subschalen: Problemen in de ouder-kindrelatie, Problemen in de opvoeding, Depressieve stemmingen, Rolbeperking, en Gezondheidsklachten. In deze studie gebruikten wij de eerste drie subschalen, evenals een totaalscore van deze drie schalen. De schalen Rolbeperking en Gezondheidsklachten werden in de huidige studie buiten beschouwing gelaten omdat deze minder indicatief zijn voor aan de opvoeding gerelateerde stress en de beleving van de ouder-kindrelatie. Een voorbeelditem van de OBVL is "Ik voel me gelukkig met mijn kind". Items worden beantwoord op een vierpuntschaal (1 = *Geldt niet* tot 4 = *Geldt helemaal*). De psychometrische kwaliteit van de OBVL is in voorgaand onderzoek aangetoond (Vermulst et al., 2015); de interne consistentie in de huidige dataset was acceptabel (Problemen in de ouder-kindrelatie: $\alpha = 0,79$; Problemen in de opvoeding: $\alpha = 0,79$; Depressieve stemmingen: $\alpha = 0,75$; Totale opvoedstress: $\alpha = 0,77$).

Voor $n = 593$ ouders was data over opvoedstress beschikbaar. Daarvan rapporteerden $n = 564$ ouders ook over het probleemgedrag van hun kind. Voor $n = 29$ ouders was enkel data over opvoedstress beschikbaar en $n = 148$ ouders rapporteerden over het probleemgedrag van hun kind, maar niet over hun ervaren opvoedstress.

2.5 Statistische Analyses

Ten eerste werden de cross-sectionele en longitudinale bivariate relaties tussen internaliserende en externaliserende gedragsproblemen bekeken middels Pearson's r -correlaties tussen de schalen van het door ouders gerapporteerd probleemgedrag (i.e., CBCL) en de schalen van het door leerkrachten gerapporteerde probleemgedrag (i.e., TRF). Ook werd de cross-sectionele relatie tussen ouder- en leerkrachtrapportage van gedragsproblemen bekeken.

Vervolgens werd de data geanalyseerd middels Repeated Measures ANOVA in SPSS. Eerst werd per uitkomstmaat (i.e., CBCL/TRF/OBVL) het hoofdeffect van voormeting naar follow-up berekend (i.e., *within-subjects effect*). Bij een significant *within-subjects effect* (i.e., een significante afname van probleemgedrag of opvoedstress van voormeting naar follow-up) voerden we *paired sample t-tests* uit om vast te stellen in welke tijdspanne (i.e., van voor- naar nameting, of van nameting naar follow-up) de significante verandering plaatsvond. De effectgroottes werden bekeken aan de hand van de *partial eta*

squared (η_p^2) waarbij $\eta_p^2 = 0,01$ als een klein effect wordt geïnterpreteerd, $\eta_p^2 = 0,06$ als een middelgroot effect en $\eta_p^2 = 0,14$ als een groot effect (Field, 2013).

Daarna werd bekeken in hoeverre de resultaten afhankelijk waren van de leeftijd van het kind, het geslacht van het kind, het niveau van problematiek bij aanvang (borderline-klinisch/niet-klinisch) en de locatie van implementatie (Vlaanderen/Nederland; i.e., *within x between-subjects effects*). Om de klinische relevantie van de resultaten in kaart te brengen, werd tenslotte middels Chi-kwadraattoetsen bekeken in hoeverre kinderen die voorafgaand aan STOP4-7 in de klinische range van probleemgedrag scoorden (T-score ≥ 63 op CBCL/TRF; van Dam & de Meyer, 2010; Nelson et al., 2002) na de interventie onder de klinische grenswaarde scoorden. Ook werd de klinische relevantie van verandering in opvoedstress middels Chi-kwadraattoetsen bekeken. Een T-score ≥ 63 werd beschouwd als indicatief voor klinische niveaus van opvoedstress (Vermulst et al., 2015).

3 Resultaten

3.1 Beschrijvende Analyses

De gemiddelden en standaarddeviaties voor alle uitkomstmaten en alle meetmomenten zijn weergegeven in Tabel 1. Er waren geen verschillen in de door ouders gerapporteerde gedragsproblemen voor jongens en meisjes voor aanvang van STOP4-7 ($t[629] = 0,286$, $p = 0,775$). Wel rapporteerden ouders minder gedragsproblemen voor jongere kinderen dan voor oudere kinderen (≥ 6 jaar; $t[214] = -2,728$, $p < 0,01$). Leerkrachten rapporteerden evenveel gedragsproblemen voor jongens als voor meisjes ($t[580] = 0,707$, $p = 0,480$), en voor jongere en oudere kinderen ($t[175] = 0,881$, $p = 0,379$) voor aanvang van STOP4-7. Er waren geen verschillen voor aanvang van STOP4-7 in ervaren opvoedbelasting voor ouders van jongens en meisjes ($t[531] = 0,136$, $p = 0,892$), of ouders van jongere en oudere kinderen ($t[171] = 1,298$, $p = 0,196$).

Zowel de samenhang tussen de subschalen Internaliserend probleemgedrag en Externaliserend probleemgedrag als de samenhang tussen meetmomenten werd bekeken middels Pearson's r -correlaties (zie Tabel 2). De correlaties tussen de meetmomenten van door ouder gerapporteerd internaliserend probleemgedrag van kinderen waren matig tot sterk positief (r_s van 0,648 tot 0,682, $p < 0,001$). Hetzelfde gold voor de autocorrelaties van door ouders gerapporteerd externaliserend probleemgedrag (r_s van 0,488 tot 0,652, $p < 0,001$). Een vergelijkbaar patroon was zichtbaar voor het door leerkrachten gerapporteerd internaliserend en externaliserend probleemgedrag van kinderen (r_s van 0,454 tot 0,534, $p < 0,001$, en r_s van 0,454 tot 0,653, $p < 0,001$, respectievelijk).

Analyse van de correlaties tussen de schalen Internaliserend probleemgedrag en Externaliserend probleemgedrag zoals gerapporteerd door ouders liet een matig sterke, positieve correlatie zien op alle meetmomenten (r_s van 0,604 tot 0,642, $p < 0,001$). De correlatie tussen internaliserend en externaliserend probleemgedrag zoals gerapporteerd

Tabel 1 Gemiddelden en standaarddeviaties: door ouders en leerkracht gerapporteerde gedragsproblemen en opvoedstress van voormeting naar follow-up

	Voormeting			Nameting			Follow-up		
	<i>n</i>	<i>M</i>	<i>SD</i>	<i>n</i>	<i>M</i>	<i>SD</i>	<i>n</i>	<i>M</i>	<i>SD</i>
Gedragsproblemen – ouderrapportage									
Totaal	632	63,73	9,68	325	58,21	10,58	193	59,39	10,75
Internaliserend		59,24	10,31		54,94	10,77		55,42	11,43
Externaliserend		67,03	9,87		60,46	10,02		61,34	10,18
Gedragsproblemen – leerkrachtrapportage									
Totaal	583	64,47	8,94	322	61,44	8,36	196	61,98	9,25
Internaliserend		57,71	9,43		55,52	8,36		56,13	10,61
Externaliserend		66,73	9,55		63,43	8,33		63,62	8,79
Opvoedstress									
Totaal	533	65,15	10,15	337	60,28	10,51	90	60,48	10,70
Ouder-kindrelatie		63,94	9,93	337	60,71	9,59	90	62,48	9,98
Problemen in opvoeding		66,18	9,70	337	60,96	9,53	89	60,52	9,21
Depressieve stemming		61,80	9,04	336	59,19	8,83	90	59,80	8,76

door leerkrachten was eveneens matig sterk en positief (r_s van 0,454 tot 0,483, $p < 0,001$). Deze resultaten geven aan dat er op alle meetmomenten sprake was van comorbiditeit tussen internaliserend en externaliserend probleemgedrag van kinderen. Deze correlatie nam niet af na deelname aan STOP4-7.

Tevens werd de samenhang tussen het door ouders en leerkrachten gerapporteerde probleemgedrag bekeken middels Pearson's r -correlaties (Tabel 2). Enkel de correlatie tussen door ouders en leerkrachten gerapporteerd internaliserend probleemgedrag ten tijde van de follow-up-meting was significant ($r = 0,253$, $p < 0,05$). De correlatie tussen door ouders en leerkracht gerapporteerd internaliserend probleemgedrag op de voormeting en nameting was niet significant (r_s van 0,073 en 0,141, $p > 0,05$). Ook de correlatie tussen door ouders en leerkrachten gerapporteerd externaliserend probleemgedrag was niet significant (r_s van $-0,003$ tot 0,142, $p > 0,05$).

3.2 Veranderingen in door Ouders Gerapporteerd Probleemgedrag

Ouders rapporteerden een grote, significante afname van totaal probleemgedrag ($\eta_p^2 = 0,31$), internaliserend probleemgedrag ($\eta_p^2 = 0,20$) en externaliserend probleemge-

Tabel 2 Pearson correlaties tussen schalen en meetmomenten van door ouders gerapporteerd probleemgedrag (i.e., CBCL) en door leerkrachten gerapporteerde probleemgedrag (i.e., TRF)

	1.	2.	3.	4.	5.	6.	7.	8.	9.
1. INT-TO	,073	,457***	,454***	,454***	,291***	,203**	,704**	,341**	,292**
2. INT-T1	,648**	,141	,534***	,225***	,447***	,143	,329**	,655**	,301**
3. INT-T2	,681**	,682**	,253*	,199*	,178	,483**	,337**	,368**	,728**
4. EXT-TO	,604**	,380**	,425**	,053	,653**	,529**	,895**	,568**	,474**
5. EXT-T1	,391**	,625**	,421**	,652**	-,003	,454**	,650**	,911**	,437**
6. EXT-T2	,349**	,382**	,642**	,488**	,550**	,142	,485**	,403**	,896**
7. TOT-TO	,834**	,572**	,582**	,870**	,627**	,460**	-,016	,636**	,513**
8. TOT-T1	,552**	,830**	,566**	,581**	,901**	,514**	,696**	,016	,493**
9. TOT-T2	,521***	,568**	,832**	,512**	,569**	,901**	,590**	,651**	,171

Noot. Correlaties voor CBCL onder de diagonaal, voor TRF boven de diagonaal; Lichtgrijs gekleurde vakken geven de correlatie tussen subschalen weer. Donker grijs gekleurde vakken op de diagonaal geven de correlatie tussen ouder- en leerkrachtrapportage weer. TO = voormeting, T1 = nameting, T2 = follow-up meting; * $p < 0,05$, ** $p < 0,01$.

drag ($\eta_p^2 = 0,38$) van kinderen van voormeting tot follow-up (Tabel 3). Het door ouders gerapporteerde totale probleemgedrag van kinderen nam significant af van voormeting naar nameting ($M_{TO-T1} = 5,736$, $SD = 7,702$, $t[268] = 12,214$, $p < 0,001$) en stabiliseerde van nameting naar follow-up ($M_{T1-T2} = -0,273$, $SD = 8,842$, $t[142] = -0,369$, $p = 0,731$). Wanneer specifiek naar het type probleemgedrag werd gekeken, was een vergelijkbaar patroon zichtbaar. Het door ouders gerapporteerde internaliserende probleemgedrag van kinderen nam significant af van voormeting naar nameting ($M_{TO-T1} = 4,145$, $SD = 8,753$, $t[268] = 7,67$, $p < 0,001$), en bleef stabiel van nameting naar follow-up ($M_{T1-T2} = 0,195$, $SD = 9,066$, $t[142] = 0,258$, $p = 0,797$). Dit was eveneens het geval voor het door ouders gerapporteerde externaliserende probleemgedrag van kinderen ($M_{TO-T1} = 6,907$, $SD = 8,001$, $t[268] = 14,158$, $p < 0,001$; $M_{T1-T2} = -0,440$, $SD = 9,377$, $t[142] = -0,562$, $p = 0,575$).

De door ouders gerapporteerde veranderingen op zowel het totale probleemgedrag als het internaliserende en het externaliserende probleemgedrag van kinderen waren niet afhankelijk van geslacht of leeftijd van de kinderen, het totale probleemgedrag bij aanvang of de locatie van de interventie (zie Tabel 3). Met andere woorden, wanneer ouders op baseline meer probleemgedrag rapporteerden vergeleken met leeftijdgenoten van hun kind, rapporteerden ouders ook op de nameting en follow-up meer probleemgedrag. De door ouders gerapporteerde afname van probleemgedrag was even sterk voor jongens als voor meisjes, en was ook even sterk in Vlaamse als in Nederlandse jeugdhulplocaties.

Tabel 3 Uitkomsten Repeated Measures ANOVA: van voormeting (To) tot follow-up (T2) op CBCL en TRF

	Ouderrapportage			Leerkrachtrapportage		
	<i>F</i>	<i>p</i>	η^2	<i>F</i>	<i>p</i>	η^2
Totaal probleemgedrag	55,927	,000	,314	3,130	,046*	,035
– Geslacht kind	,077	,926	,001	1,451	,238*	,017
– Leeftijd kind	,050	,951	,001	,712	,472*	,018
– Baseline gedragsproblemen	2,764	,065	,022	9,618	,000	,101
– Implementatie locatie	,665	,520	,005	,340	,695*	,004
Internaliserend gedrag	30,449	,000	,200	,033	,968	,000
– Geslacht kind	,276	,759	,002	4,216	,016	,047
– Leeftijd kind	1,387	,254	,022	,321	,686*	,008
– Baseline gedragsproblemen	2,675	,072	,021	4,799	,009	,053
– Implementatie locatie	,197	,822	,002	,582	,560	,007
Externaliserend gedrag	74,204	,000	,378	5,022	,010*	,055
– Geslacht kind	,572	,565	,005	,539	,567*	,006
– Leeftijd kind	,339	,713	,005	,579	,447*	,019
– Baseline gedragsproblemen	2,935	,055	,024	9,779	,000*	,102
– Implementatie locatie	2,456	,088	,020	,034	,695*	,004

Noot. Dikgedrukte resultaten zijn statistisch significant. *Greenhouse-Geisser-correctie toegepast wegens schenden voorwaarde sphericiteit.

3.3 Veranderingen in door Leerkrachten Gerapporteerd Probleemgedrag

Leerkrachten rapporteerden een kleine, significante afname van het totale probleemgedrag ($\eta_p^2 = 0,03$) en externaliserende probleemgedrag ($\eta_p^2 = 0,05$) van kinderen van voormeting naar follow-up, maar rapporteerden geen significante afname van het internaliserende probleemgedrag (zie Tabel 3). Het totale probleemgedrag van kinderen nam volgens leerkrachten af van voormeting naar nameting ($M_{T0-T1} = 2,259$, $SD = 7,285$, $t[242] = 4,834$, $p < 0,001$), en stabiliseerde van nameting naar follow-up ($M_{T1-T2} = -1,071$, $SD = 9,031$, $t[112] = -1,260$, $p = 0,210$). Ook het externaliserende gedrag van kinderen nam volgens leerkrachten significant af van voormeting naar nameting ($M_{T0-T1} = 2,613$, $SD = 7,512$, $t[242] = 5,422$, $p < 0,001$) en stabiliseerde van nameting naar follow-up ($M_{T1-T2} = -1,212$, $SD = 9,104$, $t[112] = -1,416$, $p = 0,160$). Voor het door leerkrachten gerapporteerde internaliserende gedrag was de verandering van voormeting naar nameting ($M_{T0-T1} =$

0,897, $SD = 8,927$, $t[242] = 1,567$, $p = 0,119$) evenals de verandering van nameting naar follow-up ($M_{T1-T2} = -0,823$, $SD = 9,645$, $t[112] = -0,907$, $p = 0,366$) niet significant.

De door leerkrachten gerapporteerde veranderingen van zowel het totale probleemgedrag van kinderen als het externaliserende probleemgedrag waren niet afhankelijk van geslacht of leeftijd van de kinderen, of de locatie van de interventie (zie Tabel 3). Met andere woorden, de door leerkrachten gerapporteerde afname was even sterk voor jongens als voor meisjes, voor oudere en jongere kinderen, en was even sterk voor zowel Vlaamse kinderen als Nederlandse kinderen. De verandering in het door leerkrachten gerapporteerde totale probleemgedrag was wel afhankelijk van het niveau van problematiek voorafgaand aan de interventie. Kinderen die volgens leerkrachten op voorhand op een niet-klinisch niveau scoorden, veranderden niet significant in het totale probleemgedrag van voormeting naar nameting ($M_{T0-T1} = -0,443$, $SD = 7,250$, $t[96] = -0,602$, $p = 0,548$), terwijl dit voor kinderen die op voorhand in de klinische range scoorden wel het geval was ($M_{T0-T1} = 4,055$, $SD = 6,756$, $t[145] = 7,252$, $p < 0,001$). Leerkrachten rapporteerden voor kinderen die volgens hen op baseline in de niet-klinische range scoorden geen significante afname van het totale probleemgedrag van nameting naar follow-up ($M_{T1-T2} = -0,872$, $SD = 7,237$, $t[38] = -0,752$, $p = 0,457$). Ook voor kinderen die op baseline volgens leerkrachten in de klinische range van probleemgedrag scoorden, rapporteerden leerkrachten geen significante afname van probleemgedrag van nameting naar follow-up ($M_{T1-T2} = 0,367$, $SD = 9,984$, $t[48] = -0,339$, $p = 0,736$).

Het door leerkrachten gerapporteerde externaliserende probleemgedrag van kinderen die volgens leerkrachten op voorhand in de niet-klinische range scoorden, veranderde niet significant van voormeting naar nameting, terwijl dit voor kinderen die volgens leerkrachten op voorhand in de klinische range scoorden wel het geval was ($M_{T0-T1} = 0,062$, $SD = 7,343$, $t[96] = 0,083$, $p = 0,934$ en $M_{T0-T1} = 4,308$, $SD = 7,158$, $t[145] = 7,273$, $p < 0,001$, respectievelijk). Van nameting tot follow-up nam het door leerkrachten gerapporteerde externaliserende gedrag van kinderen die volgens leerkrachten op voorhand in de niet-klinische range scoorden toe ($M_{T1-T2} = -2,359$, $SD = 7,02$, $t[38] = -2,074$, $p < 0,05$), terwijl dit stabiel bleef voor kinderen die volgens leerkrachten op voorhand in de klinische range van probleemgedrag scoorden ($M_{T1-T2} = 0,857$, $SD = 9,923$, $t[48] = 0,605$, $p = 0,548$).

Het door leerkrachten gerapporteerde internaliserende probleemgedrag van kinderen die volgens leerkrachten op voorhand in de niet-klinische range scoorden nam toe van voormeting tot nameting, terwijl dit afnam voor kinderen die volgens leerkrachten op voorhand in de klinische range scoorden ($M_{T0-T1} = -2,103$, $SD = 9,247$, $t[96] = -2,240$, $p < 0,05$ en $M_{T0-T1} = 2,890$, $SD = 8,147$, $t[145] = 4,223$, $p < 0,001$, respectievelijk). Voor beide groepen was de door leerkrachten gerapporteerde verandering van internaliserend probleemgedrag van nameting tot follow-up niet-significant ($M_{T1-T2} = 0,026$, $SD = 8,552$, $t[38] = 0,019$, $p = 0,985$ en $M_{T1-T2} = 0,286$, $SD = 10,766$, $t[48] = 0,186$, $p = 0,853$). De verandering op het door leerkrachten gerapporteerde internaliserende probleemgedrag van kinderen bleek ook afhankelijk van het geslacht van kinderen. Voor meisjes was de

Tabel 4 Uitkomsten Repeated Measures ANOVA: opvoedstress (OBVL) van voormeting naar follow-up

	Totale opvoedstress			Problemen in ouder-kindrelatie			Problemen in opvoeding			Depressieve stemmingen		
	<i>F</i>	<i>p</i>	η^2	<i>F</i>	<i>p</i>	η^2	<i>F</i>	<i>p</i>	η^2	<i>F</i>	<i>p</i>	η^2
Hoofdeffect	27,597	,000	,305	17,062	,000	,213	46,773	,000	,430	11,756	,000	,157
- Geslacht kind	,008	,992	,000	,570	,567	,009	,307	,736	,005	,572	,566	,009
- Leeftijd kind	,172	,843	,004	,024	,976	,001	1,206	,304	,028	,264	,768	,006
- Baseline opvoedstress	2,511	,085	,039	6,183	,003	,091	,883	,416	,014	2,156	,120	,034
- Implementatie locatie	1,107	,334	,018	1,098	,337	,017	,881	,417	,014	,193	,824	,003

Noot. Dikgedrukte resultaten zijn statistisch significant.

door leerkrachten gerapporteerde afname van voormeting naar nameting significant ($M_{T0-T1} = 3,309$, $SD = 10,189$, $t[41] = 2,105$, $p < 0,05$), maar voor jongens niet ($M_{T0-T1} = 0,393$, $SD = 8,583$, $t[200] = 0,649$, $p = 0,517$). Voor zowel meisjes als jongens was de door leerkrachten gerapporteerde verandering in internaliserend probleemgedrag van nameting naar follow-up niet significant ($M_{T1-T2} = -0,312$, $SD = 15,326$, $t[15] = -0,082$, $p = 0,936$ en $M_{T1-T2} = -0,842$, $SD = 8,548$, $t[94] = -0,960$, $p = 0,339$, respectievelijk). De door leerkrachten gerapporteerde verandering in internaliserend probleemgedrag van kinderen was niet afhankelijk van leeftijd of de locatie van de interventie.

3.4 Veranderingen in Opvoedstress (OBVL)

De door ouders gerapporteerde totale opvoedstress nam significant af van voormeting tot follow-up ($\eta_p^2 = 0,30$; zie Tabel 4). Dit is een groot effect. Deze afname was zichtbaar van voormeting naar nameting ($M_{T0-T1} = 5,294$, $SD = 8,288$, $t[281] = 10,727$, $p < 0,001$) en stabiliseerde van nameting naar follow-up ($M_{T1-T2} = 0,271$, $SD = 8,689$, $t[69] = 0,261$, $p = 0,795$). Wanneer naar de subschalen van de OBVL werd gekeken, was een vergelijkbaar patroon zichtbaar. Gerapporteerde opvoedstress door problemen in de ouder-kindrelatie (OKR; $\eta_p^2 = 0,21$), problemen in de opvoeding (PO; $\eta_p^2 = 0,43$) en depressieve stemmingen (DEP; $\eta_p^2 = 0,16$) namen significant af van voormeting naar follow-up. Ook hierbij was in alle gevallen sprake van een groot effect. De afname van de gerapporteerde opvoedbelasting was steeds significant van voormeting naar nameting (OKR: $M_{T0-T1} = 3,921$, $SD = 8,061$, $t[281] = 8,170$, $p < 0,001$; PO: $M_{T0-T1} = 6,237$, $SD = 8,261$, $t[281] = 12,679$, $p < 0,001$; DEP: $M_{T0-T1} = 2,868$, $SD = 6,857$, $t[280] = 7,012$, $p < 0,001$) en stabiliseerde van nameting naar follow-up (OKR: $M_{T1-T2} = -1,000$, $SD = 7,940$, $t[69] = -1,054$, $p = 0,296$; PO: $M_{T1-T2} = 0,754$, $SD = 7,558$, $t[68] = 0,828$, $p = 0,410$; DEP: $M_{T1-T2} = -3,000$, $SD = 7,913$, $t[69] = -0,317$, $p = 0,752$).

De gerapporteerde veranderingen van problemen in de opvoeding, depressieve stemmingen en totale opvoedstress was niet afhankelijk van het geslacht van het kind, de

Tabel 5 Verdeling in niet-klinische (CBCL T-score ≤ 62 ; OBVL T-score ≤ 62) en borderline tot klinische (CBCL T-score ≥ 63 ; OBVL T-score ≥ 63) range van door ouders gerapporteerd probleemgedrag en opvoedstress

	Voormeting			Nameting			Follow-up		
	<i>N</i>	Niet-klinisch <i>n</i> (%)	Klinisch <i>n</i> (%)	<i>N</i>	Niet-klinisch <i>n</i> (%)	Klinisch <i>n</i> (%)	<i>N</i>	Niet-klinisch <i>n</i> (%)	Klinisch <i>n</i> (%)
Totaal probleemgedrag	632	268 (42,4)	364 (57,6)	325	210 (64,6)	115 (35,4)	193	116 (60,1)	77 (39,9)
Internaliserend gedrag		374 (59,3)	258 (40,7)		238 (73,8)	87 (26,2)		135 (69,9)	58 (30,1)
Externaliserend gedrag		185 (29,3)	447 (70,7)		193 (59,4)	132 (40,6)		108 (56,0)	85 (44,0)
Opvoedstress	533	182 (34,1)	351 (65,9)	337	176 (52,2)	161 (47,8)	90	45 (50,0)	45 (50,0)

leeftijd van het kind, de mate van opvoedstress bij aanvang of locatie van implementatie van STOP4-7 (zie Tabel 4). Met andere woorden, ouders die op de voormeting meer opvoedstress rapporteren ten opzichte van andere ouders, rapporteerden op de nameting en follow-up doorgaans ook veel opvoedstress. De gerapporteerde afname in opvoedstress was even sterk voor ouders van jongens en meisjes, voor ouders van oudere en jongere kinderen, en voor ouders van Vlaamse en Nederlandse kinderen.

Wanneer ouders op voorhand in de klinische range van opvoedstress scoorden, namen problemen in de ouder-kindrelatie significant af van voormeting naar nameting ($M_{T0-T1} = 5,289$, $SD = 7,902$, $t[189] = 9,226$, $p < 0,001$), maar dit was niet het geval wanneer ouders op voorhand in de niet-klinische range van opvoedstress scoorden ($M_{T0-T1} = 1,098$, $SD = 7,680$, $t[91] = 1,371$, $p = 0,174$). Voor beide groepen was de verandering van nameting tot follow-up niet-significant ($M_{T0-T1} = -2,250$, $SD = 10,199$, $t[11] = -0,764$, $p = 0,461$ en $M_{T0-T1} = -1,109$, $SD = 7,596$, $t[51] = -0,968$, $p = 0,338$, respectievelijk).

3.5 Klinische Relevantie

Tabel 5 geeft het aantal kinderen weer dat volgens ouders op de verschillende meetmomenten in de klinische range van probleemgedrag scoorden. Dit aantal nam over de meetmomenten significant af. Voor de door ouders gerapporteerde totale gedragsproblemen (i.e., zowel internaliserend als externaliserend) gold een significant verschil in het aantal kinderen in de klinische range tussen voormeting en nameting ($\chi^2 [1, N = 269] = 59,066$, $p < 0,001$), en tussen voormeting en follow-up ($\chi^2 [1, N = 147] = 18,654$, $p < 0,001$). Van de 150 kinderen die op voormeting volgens ouders in de klinische range van probleemgedrag scoorden, scoorden op nameting 46,0% ($n = 69$) volgens ouders in de niet-klinische range. Op follow-up scoorde 49,5% ($n = 48$) volgens ouders in de niet-klinische range van de kinderen die op voormeting volgens ouders nog in de klinische range scoorden ($n = 97$).

Dit patroon was ook zichtbaar wanneer afzonderlijk naar het door ouders gerapporteerde internaliserende en externaliserende probleemgedrag werd gekeken. Van de 107

kinderen die volgens ouders op voormeting in de klinische range scoorden op internaliserend probleemgedrag, scoorden op nameting 46,7% ($n = 50$) volgens ouders in de niet-klinische range ($\chi^2 [1, N = 269] = 63,678, p < 0,001$). Op follow-up scoorde 42,6% ($n = 29$) van de kinderen volgens ouders in de niet-klinische range die op voormeting volgens ouders wel in de klinische range scoorden ($n = 68; \chi^2 [1, N = 147] = 42,596, p < 0,001$). Van de 192 kinderen die op voormeting volgens ouders in de klinische range scoorden op externaliserend probleemgedrag, scoorde 46,9% ($n = 90$) volgens ouders in de niet-klinische range op nameting ($\chi^2 [1, N = 269] = 47,002, p < 0,001$). Op follow-up scoorde 48,7% ($n = 56$) van de kinderen volgens ouders in de niet-klinische range die op voormeting volgens ouders wel in de klinische range scoorden ($n = 115; \chi^2 [1, N = 147] = 18,046, p < 0,001$).

Tenslotte was ook de afname van ervaren opvoedstress klinisch relevant. Van de 190 ouders die voormeting klinische niveaus van opvoedstress rapporteerden, scoorde 33,7% ($n = 64$) op nameting in de niet-klinische range ($\chi^2 [1, N = 282] = 64,742, p < 0,001$). Op follow-up scoorde 41,0% ($n = 25$) van de ouders in de niet-klinische range van opvoedstress die op voormeting in de klinische range scoorden ($n = 61; \chi^2 [1, N = 79] = 3,682, p = 0,055$).

4 Discussie

Zoals verwacht was na deelname aan de multimodale interventie STOP4-7 zowel externaliserend als internaliserend probleemgedrag bij kinderen, en gerapporteerde opvoedstress bij ouders afgenomen. De bevindingen laten ook zien dat deze afname duurzaam was; de afname in probleemgedrag en opvoedstress die ontstond tijdens de interventie, bleef behouden tot aan de follow-up-meting zes maanden later. De duurzame afname in probleemgedrag was even sterk voor jongens als meisjes en voor jongere als oudere kinderen. Bovendien toonden de analyses aan dat deze veranderingen klinisch relevant waren: een aanzienlijk deel – meer dan 45% – van de kinderen die voor deelname aan STOP4-7 in de klinische range van probleemgedrag scoorden, keerden terug naar een “normaal”, gezond functioneren. Ook de afname van opvoedstress bij ouders was klinisch relevant.

De gevonden veranderingen in door ouders gerapporteerd probleemgedrag en opvoedstress zijn groot te noemen ($\eta_p^2 = 0,15-0,43$). Dit is een relatief gunstige uitkomst als we deze vergelijken met bevindingen uit eerdere meta-analyses, waaruit bleek dat oudertrainingen in het kader van geïndiceerde preventie of bij klinische groepen gemiddeld een klein (Cohen's $d = 0,27$; Menting et al., 2013) tot middelgroot ($d = 0,42$; Lundahl et al., 2006) effect hadden. Echter, ondanks de aangetoonde veranderingen in gerapporteerd probleemgedrag en opvoedstress kan het huidige onderzoek geen definitieve uitspraken doen over de effectiviteit van STOP4-7. Dit heeft te maken met het ontbreken van een controlegroep in het design van de studie. De vraag is of de gevonden effecten door

dit programma werden veroorzaakt, of dat er alternatieve verklaringen denkbaar zijn. Mogelijk groeien kinderen over hun problemen heen – er is dan sprake van spontaan herstel – of hebben gunstige veranderingen in de omgeving (een andere school, een verhuizing) een positieve invloed op het gedrag van kinderen en de stress van ouders. Zelfs als deze alternatieve verklaringen als minder geloofwaardig worden gezien, geldt nog dat we niet met zekerheid kunnen vaststellen dat het gevonden effect specifiek is voor STOP4-7, en niet, bijvoorbeeld, een algemeen effect is van positieve aandacht en tijd die voor het gezin worden vrijgemaakt.

In lijn met resultaten uit eerder onderzoek laat onze studie zien dat de veranderingen van STOP4-7 niet alleen statistisch significant, maar ook klinisch relevant zijn (Kazdin, 1999). Het bleek dat meer dan 40% van alle kinderen volgens hun ouders zodanig vooruitgingen, dat er niet langer sprake meer was van een klinisch probleem. Gezien het negatieve ontwikkelingsperspectief dat gepaard kan gaan met gedragsproblemen op jonge leeftijd, is dit een hoopvol resultaat. Mogelijk kan het op jonge leeftijd – wanneer gedrag mogelijk nog kneedbaar is – inzetten van een multimodaal programma zoals STOP4-7 het percentage volwassenen met klinische problematiek terugdringen. Om een dergelijke hypothese te bevestigen zouden kinderen in de huidige sample echter tot in de volwassenheid gevolgd moeten worden.

Ook in lijn met eerdere bevindingen bleek dat niet alleen externaliserend maar ook internaliserend probleemgedrag significant afnam (Sandler et al., 2011). Een verklaring hiervoor zou kunnen liggen in het feit dat als externaliserend probleemgedrag afneemt, ook veel van de faalervaringen en ervaren marginalisatie in de omgang met autoriteitsfiguren buiten het gezin en met leeftijdgenoten wegvallen. Hierdoor zal het welzijn van kinderen mogelijk kunnen toenemen, en depressieve stemmingen en angstigheid verminderen (Hukkelberg et al., 2019; Overbeek et al., 2021). Opvallend was dat de significante afname van internaliserend gedrag niet werd gerapporteerd door leerkrachten. Mogelijk is dit te verklaren vanuit het feit dat leerkrachten internaliserend gedrag niet zozeer waarnemen, of zelfs als minder serieus beschouwen dan externaliserend gedrag (Splett et al., 2019). Deze verklaring wordt ondersteund door de bevinding dat leerkrachten wel een significante afname van internaliserend gedrag rapporteerden wanneer leerlingen bij de voormeting in de klinische range scoorden.

Ook voor het totale probleemgedrag en externaliserende probleemgedrag bleek dat de afname groter was wanneer kinderen bij aanvang in de klinische range scoorden, wat in lijn is met eerdere bevindingen van Leijten et al. (2018) en Van Aar et al. (2017). Opvallend was dat dit effect alleen naar voren kwam uit de leerkrachtrapportage. Ouders rapporteerden wel een significante afname van het probleemgedrag, maar dit was niet afhankelijk van het niveau van problematiek bij aanvang. Onderzoek in 21 landen naar de gebruikte instrumenten voor probleemgedrag gerapporteerd door ouders en leerkrachten (i.e., de CBCL en TRF) toonde aan dat de overeenstemming tussen ouder- en leerkrachtrapportage van probleemgedrag matig is ($r = 0,29$), en dat ouders vaak meer gedragsproblemen rapporteren dan leerkrachten (Rescorla et al., 2014). Ook in dit onderzoek werd een der-

gelijk patroon gevonden. De correlaties tussen ouder- en leerkrachtrapportages bleken overwegend niet significant, wat aantoont dat ouders en leerkrachten anders oordelen over het probleemgedrag van kinderen, een gegeven dat ook in eerder onderzoek werd gevonden (Winsler & Wallace, 2002). Omdat de gevonden veranderingen in het door leerkrachten gerapporteerde probleemgedrag klein waren ($\eta_p^2 = 0,03-0,05$), is een mogelijke verklaring dat disruptief gedrag in de kindertijd zich anders (en wellicht minder intens) manifesteert in de school- dan in de gezinscontext (De Los Reyes et al., 2013). Dit resultaat wijst in ieder geval ook op de noodzaak om meer context-sensitieve instrumenten te ontwikkelen om probleemgedrag bij kinderen te meten (Winsler & Wallace, 2002).

Opmerkelijk is dat uit de ouder rapportage van probleemgedrag geen verschil blijkt in de gedragsveranderingen na STOP4-7 voor jongens of meisjes. Deze bevinding komt overeen met onderzoek naar een vergelijkbaar multimodaal programma *Incredible Years* (Beauchaine et al., 2005). Ander onderzoek naar de effecten van *Incredible Years* vond echter wel verschillen in effectiviteit voor jongens en meisjes (Gardner et al., 2006). Mogelijk wijzen de huidige bevindingen erop dat multimodale programma's, waarbij zowel kinderen, ouders als leerkrachten betrokken zijn, even goed werken voor jongens als voor meisjes. Anderzijds zijn de huidige bevindingen mogelijk te verklaren vanuit het feit dat de huidige steekproef substantieel meer jongens dan meisjes bevatte. Wel bleek dat leerkrachten bij meisjes wel een afname in internaliserend gedrag rapporteerden, maar niet bij jongens. Leerkrachten hebben mogelijk meer ervaring met internaliserend gedrag bij meisjes omdat dit vaker voorkomt, en herkennen internaliserende symptomen zodoende eerder als problematisch bij meisjes in vergelijking met jongens (Loades & Mastroyannopoulou, 2010). De huidige bevindingen en tegenstrijdigheden in de literatuur onderschrijven de noodzaak aandacht te blijven besteden aan mogelijke sekseverschillen in de effecten van (multimodale) interventies voor probleemgedrag.

Ten slotte vonden we dat de problemen in de ouder-kindrelatie alleen significant afnamen wanneer ouders voor deelname aan STOP4-7 in de klinische range van opvoedbelasting scoorden. Deze bevinding is in lijn met eerdere bevindingen (Van Brussel et al., 2009; De Mey, 2010, De Mey & Braet, 2010) en het model van Belsky (1984). Mogelijk leidt STOP4-7 tot een vermindering van opvoedbelasting bij ouders doordat zij meer controle over hun leven ervaren en ontstaat er hierdoor meer ruimte voor positieve interacties tussen ouder en kind (e.g., Farmer & Lee, 2011). Dit zou met name het geval kunnen zijn bij gezinnen waar de initiële opvoedbelasting hoog is, omdat in die gezinnen de meeste winst te behalen is.

Dit onderzoek geeft als eerste onderzoek inzicht in de mogelijke werkzaamheid van STOP4-7, in zowel Nederland als in Vlaanderen, waarbij ook werd gekeken naar de werkzaamheid van het programma voor verschillende subgroepen en over een langere termijn, aan de hand van rapportages van ouders én leerkrachten. Ondanks het feit dat de studie in een klinische populatie werd uitgevoerd, is de steekproef groot. Bovendien vergroot de multi-informant-benadering van het huidige onderzoek de betrouwbaarheid van de

bevindingen. Tegelijkertijd kent de studie ook enkele beperkingen. Om de bewijskracht voor STOP4-7 te versterken zou in vervolgonderzoek, eventueel middels een wachtlijst-constructie, een actieve (care as usual) controlegroep meegenomen moeten worden. Ook is er in deze studie niet gekeken naar veranderingen in de veronderstelde gezinsmechanismen die ten grondslag zouden moeten liggen aan afnames in probleemgedrag. Het verdient met name aanbeveling om bijvoorbeeld observatiedata te verzamelen met betrekking tot dwingende interactiepatronen waarbij ouders en kinderen elkaar negatief bekrachtigen. Tevens kunnen veranderingen in veronderstelde werkzame opvoedtechnieken, zoals het positief bekrachtigen van gewenst gedrag en effectief stellen van grenzen, middels observaties nauwkeuriger in kaart gebracht kunnen worden.

Verder zouden meer gegevens moeten worden verzameld over de implementatiegraad en -kwaliteit van STOP4-7, zodat kan worden geanalyseerd of gezinnen met een hulpvraag mogelijkserwijs nog effectiever geholpen zouden kunnen worden. Voor het huidige onderzoek waren geen gegevens over de uitval tijdens het programma bekend en zodoende moeten de conclusies van het huidige onderzoek met voorzichtigheid geïnterpreteerd worden. Echter, de ervaring leert dat slechts een kleine groep ouders die starten met STOP4-7 het programma niet geheel doorlopen (i.e., ongeveer één op tien ouders valt uit; De Mey, persoonlijke communicatie) en zodoende verwachten wij niet dat de huidige bevindingen sterk beïnvloed worden doordat ouders er niet in slagen het programma af te maken. Tegelijkertijd is het van belang dat vervolgonderzoek eventuele uitval in kaart brengt om deze verwachting te bevestigen.

Concluderend bleek uit deze multi-informant-studie dat STOP4-7 leidde tot significante, klinisch relevante, en blijvende afnames van zowel externaliserend als internaliserend probleemgedrag van kinderen, en een significante afname van opvoedstress bij ouders. De gevonden afnames waren even sterk voor ouders van jongere en oudere kinderen, en voor jongens en meisjes. STOP4-7 lijkt verder het beste te werken voor kinderen die bij aanvang klinische gedragsproblemen vertonen. Over het algemeen blijkt uit deze studie dat helder geprotocolleerde, op maat afgestemde, en multimodale programma's een aanwinst zijn voor de Nederlandse en Vlaamse jeugdhulp. De huidige veelbelovende resultaten dragen bij aan de *evidence base* voor STOP4-7 en de implementatie daarvan in lokale of regionale context.

Korte Verklaring over de Data en Analyses

Analyse codes en output: Alle data die beschikbaar waren voor de secundaire analyse werden beschreven in dit artikel. De analysecodes en -output zijn vrij toegankelijk via Open Science Framework: <https://osf.io/bz23j/>.

Data: Wegens gevoeligheid van de data zijn deze niet vrij toegankelijk. Informatie/vragen over de data is op te vragen bij/kunnen gesteld worden aan Wim De Mey (Vlaamse data) en Laura Beurskens-Claessens (Nederlandse data).

Literatuur

- Achenbach, T.M. (1991a). *Manual for the Child Behavior Checklist/4-18 and 1991 Child Profile*. Burlington: University of Vermont, Department of Psychiatry.
- Achenbach, T.M. (1991b). *Manual for the Teacher Report Form*. Burlington: University of Vermont Press.
- American Psychiatric Association (2013). *Diagnostic and statistical manual of mental disorders (5th edition)*. American Psychiatric Association.
- Beauchaine, T.P., Webster-Stratton, C., & Reid, M.J. (2005). Mediators, moderators, and predictors of 1-year outcomes among children treated for early-onset conduct problems: a latent growth curve analysis. *Journal of Consulting and Clinical Psychology*, 73(3), 371. <https://psycnet.apa.org/doi/10.1037/0022-006X.73.3.371>
- Belsky, J. (1984). The determinants of parenting: A process model. *Child Development*, 55, 83-96. <https://doi.org/10.2307/1129836>.
- Bridgett, D.J., Burt, N.M., Edwards, E.S., & Deater-Deckard, K. (2015). Intergenerational transmission of self-regulation: A multidisciplinary review and integrative conceptual framework. *Psychological Bulletin*, 141(3), 602-654. doi: 10.1037/a0038662
- Capaldi, D. (1992). Co-occurrence of conduct problems and depressive symptoms in early adolescent boys: II. A 2-year follow-up at Grade 8. *Development and Psychopathology*, 4(1), 125-144. <https://doi.org/10.1017/S0954579400005605>
- De Mey, W. (2010). *Begeleiding van ouders van jonge kinderen met gedragsproblemen vanuit een sociaalpedagogische en ontwikkelingspsychopathologisch perspectief*. [Doctoraat, Universiteit Gent]. RUG Proefschriften: <https://lib.ugent.be/catalog/rug01:001388401>
- De Mey, W., & Braet, C. (2011). STOP4-7, an early intervention for children with behavioural disorders: Client characteristics and preliminary findings. *European Journal of Developmental Psychology*, 8(2), 203-214. <http://dx.doi.org/10.1080/17405620903513943>
- De Mey, W., & Merlevede, E. (2013) STOP4-7, een protocol op maat.
- Delsing, M., Bastiaanssen, I., & Veerman, J.W. (2010, September). Long-term effectiveness and mechanisms of change of a multimodal program for treating behavioral problems of young children. Gepresenteerd op EURSARF, Groningen.
- De Los Reyes, A., Thomas, S.A., Goodman, K.L., & Kundery, S.M. (2013). Principles underlying the use of multiple informants' reports. *Annual Review of Clinical Psychology*, 9, 123-149. <https://doi.org/10.1146/annurev-clinpsy-050212-185617>
- Eddy, J.M., Reid, J.B., & Fetrow, R.A. (2000). An elementary school-based prevention program targeting modifiable antecedents of youth delinquency and violence: Linking the Interests of Families and Teachers (LIFT). *Journal of Emotional and Behavioral Disorders*, 8, 165-176. <https://doi.org/10.1177/10634266000800304>
- Eisenberg, N., Valiente, C., Spinrad, T.L., Cumberland, A., Liew, J., Reiser, M., ... & Losoya, S.H. (2009). Longitudinal relations of children's effortful control, impulsivity, and negative emotionality to their externalizing, internalizing, and co-occurring behavior problems. *Developmental Psychology*, 45(4), 988-1008. <https://psycnet.apa.org/doi/10.1037/a0016213>

- Evers, A., van Vliet-Mulder, J.C., & Groot, C.J. (2001). *Documentatie van tests en testresearch in Nederland*. Assen: Van Gorcum.
- Farmer, A.Y. & Lee, S.K. (2011). The effects of parenting stress, perceived mastery, and maternal depression on parent–child interaction. *Journal of Social Service Research*, 37(5), 516–525. <https://doi.org/10.1080/01488376.2011.607367>
- Field, A. (2013). *Discovering statistics using IBM SPSS statistics*. SAGE.
- Gardner, F., Hutchings, J., Bywater, T., & Whitaker, C. (2010). Who benefits and how does it work? Moderators and mediators of outcome in an effectiveness trial of a parenting intervention. *Journal of Clinical Child & Adolescent Psychology*, 39(4), 568–580. <https://doi.org/10.1080/15374416.2010.486315>
- Hukkelberg, A., Tømmerås, T., & Ogden, T. (2019). Parent training: Effects beyond conduct problems. *Children and Youth Services Review*, 100, 405–414. <https://doi.org/10.1016/j.childyouth.2019.03.009>
- Kazdin, A.E. (1999). The meanings and measurement of clinical significance. *Journal of Consulting and Clinical Psychology*, 67(3), 332–339. <https://psycnet.apa.org/doi/10.1037/0022-006X.67.3.332>
- Kazdin, A.E., & Weisz, J.R. (1998). Identifying and developing empirically supported child and adolescent treatments. *Journal of Consulting and Clinical Psychology*, 66, 19–36.
- Kim-Cohen, J., Caspi, A., Moffitt, T.E., Harrington, H., Milne, B.J., & Poulton, R. (2003). Prior juvenile diagnoses in adults with mental disorder: developmental follow-back of a prospective-longitudinal cohort. *Archives of General Psychiatry*, 60(7), 709–717. doi:10.1001/archpsyc.60.7.709
- La Greca, A.M., Silverman, W.K., & Lochman, J.E. (2009). Moving beyond efficacy and effectiveness in child and adolescent intervention research. *Journal of Consulting and Clinical Psychology*, 77(3), 373–382. <https://psycnet.apa.org/doi/10.1037/a0015954>
- Lahey, B.B., Van Hulle, C.A., Waldman, I.D., Rodgers, J.L., D'Onofrio, B.M., Pedlow, S., Rathouz, P., & Keenan, K. (2006). Testing descriptive hypotheses regarding sex differences in the development of conduct problems and delinquency. *Journal of Abnormal Child Psychology*, 34(5), 730–748.
- Leijten, P., Gardner, F., Landau, S., Harris, V., Mann, J., Hutchings, J., ... & Scott, S. (2018). Research Review: Harnessing the power of individual participant data in a meta-analysis of the benefits and harms of the Incredible Years parenting program. *Journal of Child Psychology and Psychiatry*, 59(2), 99–109. <https://doi.org/10.1111/jcpp.12781>
- Loades, M.E., & Mastroyannopoulou, K. (2010). Teachers' recognition of children's mental health problems. *Child and Adolescent Mental Health*, 15(3), 150–156. <https://doi.org/10.1111/j.1475-3588.2009.00551.x>
- Lundahl, B., Risser, H.J., & Lovejoy, M.C. (2006). A meta-analysis of parent training: Moderators and follow-up effects. *Clinical Psychology Review*, 26(1), 86–104. <https://doi.org/10.1016/j.cpr.2005.07.004>
- Menting, A.T., de Castro, B.O., & 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(8), 901–913. <https://doi.org/10.1016/j.cpr.2013.07.006>
- Mingebach, T., Kamp-Becker, I., Christiansen, H., & Weber, L. (2018). Meta-meta-analysis on the

- effectiveness of parent-based interventions for the treatment of child externalizing behavior problems. *PLoS ONE*, 13(9), e0202855. <https://doi.org/10.1371/journal.pone.0202855>
- Moffitt, T.E., Arseneault, L., Belsky, D., Dickson, N., Hancox, R.J., Harrington, H., Houts, R., Poulton, R., Robert, B.W., Ross, S., Sears, M.R., Thomson, W.M., & Caspi, A. (2011). A gradient of childhood self-control predicts health, wealth, and public safety. *Proceedings of the National Academy of Sciences*, 108(7), 2693–2698. <https://doi.org/10.1073/pnas.1010076108>
- Nelson, J.R., Benner, G.J., Reid, R.C., Epstein, M.H., & Currin, D. (2002). The convergent validity of office discipline referrals with the CBCL-TRF. *Journal of Emotional and Behavioral Disorders*, 10(3), 181–188. <https://doi.org/10.1177/02F1063426602010030601>
- Orobio de Castro, B., Veerman, J.W., Koops, W., Bosch, J.D., & Monshouwer, H.J. (2002). Hostile attribution of intent and aggressive behaviour: A meta-analysis. *Child Development*, 73, 916–934. <https://doi.org/10.1111/1467-8624.00447>
- Overbeek, G., Vollebergh, W., Meeus, W., Engels, R., & Luijpers, E. (2001). Course, co-occurrence, and longitudinal associations of emotional disturbance and delinquency from adolescence to young adulthood: A three-wave six-year study. *Journal of Youth and Adolescence*, 30, 401–427. <https://doi.org/10.1023/A:1010441131941>
- Overbeek, G., Van Aar, J., Orobio de Castro, B., Matthys, W., Weeland, J., Chhangur, R.R., & Leijten, P. (2021). Longer-Term Outcomes of the Incredible Years Parenting Intervention. *Prevention Science*, 21, 419–431. <https://doi.org/10.1007/s1121-020-01176-6>
- Patterson, G.R. (1982). *Coercive family process*. Eugene, or: Castalia.
- Patterson, G.R. (2002). The early development of coercive process. In J.B. Reid, G.R. Patterson & J. Snyder (Eds.), *Antisocial behaviour in children and adolescents. A developmental analysis and model for intervention* (pp. 25–44). Washington, DC: American Psychological Association.
- Patterson, G.R., & Capaldi, D.M. (1990). A mediational model for boys' depressed mood. In J. Rolf, A.S. Masten, D. Cicchetti, K.H. Nuechterlein, & S. Weintraub (Eds.), *Risk and protective factors in the development of psychopathology* (pp. 141–163). New York: Cambridge University Press.
- Patterson, G.R., Reid, J.B., & Dishion, J.B. (1992). *Antisocial boys*. Eugene, or: Castalia.
- Sandler, I.N., Schoenfelder, E.N., Wolchik, S.A., & MacKinnon, D.P. (2011). Long-term impact of prevention programs to promote effective parenting: Lasting effects but uncertain processes. *Annual Review of Psychology*, 62, 299–329. <https://doi.org/10.1146/annurev.psych.121208.131619>
- Shelleby, E.C., & Shaw, D.S. (2014). Outcomes of parenting interventions for child conduct problems: A review of differential effectiveness. *Child Psychiatry & Human Development*, 45(5), 628–645. <https://doi.org/10.1007/s10578-013-0431-5>
- Smith, J.D., Dishion, T.J., Shaw, D.S., Wilson, M.N., Winter, C.C., & Patterson, G.R. (2014). Coercive family process and early-onset conduct problems from age 2 to school entry. *Development and Psychopathology*, 26(4pt1), 917–932. <https://doi.org/10.1017/S0954579414000169>
- Splett, J.W., Garzona, M., Gibson, N., Wojtalewicz, D., Raborn, A., & Reinke, W.M. (2019). Teacher recognition, concern, and referral of children's internalizing and externalizing behavior problems. *School Mental Health*, 11(2), 228–239. <https://doi.org/10.1007/s12310-018-09303-z>
- Reid, J.B., Patterson G.R., & Snyder, J. (2002). *Antisocial behavior in children and adolescents. A*

- developmental analysis and model for intervention*. Washington, DC: American Psychological Association.
- Rescorla, L.A., Bochicchio, L., Achenbach, T.M., Ivanova, M.Y., Almqvist, F., Begovac, I., Bilenberg, N., Bird, H., Dobrea, A., Erol, N., Fombonne, E., Fonseca, A., Frigeiro, A., Fung, D.S.S., Lambert, M.C., Leung, P.W.L., Liu, X., Marković, I., Markovic, J. ... & Verhulst, F.C. (2014). Parent-teacher agreement on children's problems in 21 societies. *Journal of Clinical Child & Adolescent Psychology*, 43(4), 627–642. <https://doi.org/10.1080/15374416.2014.900719>
- Van Aar, J., Leijten, P., de Castro, B. O., & Overbeek, G. (2017). Sustained, fade-out or sleeper effects? A systematic review and meta-analysis of parenting interventions for disruptive child behavior. *Clinical Psychology Review*, 51, 153–163. <https://doi.org/10.1016/j.cpr.2016.11.006>
- Van Brussel, A., De Meyer, R., Veerman, J.W., & De Mey, W. (2009). Evaluatie van een multimodale behandeling voor kleuters met ernstige gedragsproblemen. *Kind en Adolescent*, 30(2), 66–81. <https://doi.org/10.1007/BF03087937>
- Van Dam, C., & De Meyer, R. (2010). *Praktijkgestuurd veranderingsonderzoek Orthopedagogisch Centrum Brabant. Resultaten september 2008-september 2009*. Praktikon.
- Van der Giessen, D., Branje, S., Overbeek, G., Frijns, T., Van Lier, P.A.C., Koot, H.M., & Meeus, W. (2013). Co-occurrence of aggressive behavior and depressive symptoms in early adolescence: A longitudinal multi-informant study. *European Review of Applied Psychology*, 63(4), 193–201. <https://doi.org/10.1016/j.erap.2013.03.001>
- Verhulst, F.C., Van der Ende, J., & Koot, H.M. (1996). *Handleiding voor de CBCL/4–18*. Rotterdam: Sophia Kinderziekenhuis/Academisch ziekenhuis Rotterdam (Afdeling Kinder- en Jeugdpsychiatrie).
- Verhulst, F.C., Van der Ende, J., & Koot, H.M. (1997). *Handleiding voor de Teacher's Report Form (TRF)*. Rotterdam: Sophia Kinderziekenhuis/Academisch ziekenhuis Rotterdam (Afdeling Kinder- en Jeugdpsychiatrie).
- Vermulst, A., Kroes, G., De Meyer, R., Nguyen, L., & Veerman, J. (2015). *Handleiding OBVL*. Praktikon.
- Von Stumm, S., Deary, I.J., Kivimäki, J., Jokela, M., Clark, H., & Batty, G.D. (2011). Childhood behavior problems and health at midlife: 35-year follow-up of a Scottish birth cohort. *Journal of Child Psychology and Psychiatry*, 52, 992–1001. <https://doi.org/10.1111/j.1469-7610.2011.02373>
- Webster-Stratton, C., & Taylor, T. (2001). Nipping early risk factors in the bud: Preventing substance abuse, delinquency, and violence in adolescence through interventions targeted at young children (0–8 years). *Prevention Science*, 2(3), 165–192. <https://doi.org/10.1023/A:1011510923900>
- Winsler, A., & Wallace, G.L. (2002). Behavior problems and social skills in preschool children: Parent-teacher agreement and relations with classroom observations. *Early Education and Development*, 13(1), 41–58. https://doi.org/10.1207/s15566935eed1301_3