

Peer-reviewverslag Evaluatie van Samen sterker Terug Op Pad (STOP4-7)

Originele manuscript en reviews

30-8-2021	Begeleidende brief originele manuscript	zie pagina 2
30-8-2021	Originele manuscript	zie pagina 3
24-11-2021	Reactie reviewer 1	zie pagina 38
24-11-2021	Reactie reviewer 2	zie pagina 40

Tweede versie

21-12-2021	Reactiebrief auteurs	zie pagina 42
21-12-2021	Tweede versie van manuscript	zie pagina 49
24-1-2022	Acceptatiebrief	zie pagina 92

Publicatie

Ontvangst: 30-8-2021

Acceptatie: 14-1-2022

Online: 3-3-2022

Begeleidende brief originele manuscript – 30-8-2021

Beste collega's,

Bijgevoegd vindt u het paper getiteld “Evaluatie van Samen sterker Terug Op Pad (STOP4-7): Een multimodale interventie voor kinderen met gedragsproblemen”.

STOP4-7 is een multimodale preventieve interventie gericht op het verminderen van probleemgedrag bij kinderen tussen 4 en 7 jaar. In de bijgevoegde paper bekeken we de effecten van STOP4-7 middels data die werd verzameld onder 770 kinderen in de leeftijd van 3 tot en met 8 jaar in Nederland en Vlaanderen. We bekeken hierbij niet enkel in hoeverre er veranderingen plaatsvonden in het internaliserend- en externaliserend probleemgedrag van kinderen en opvoedstress bij ouders, maar ook naar de klinische relevantie van deze veranderingen. Bovendien bekeken voor wie het programma mogelijk beter of minder goed werkt middels moderatie analyses.

Het paper draagt bij aan de wetenschappelijke kennis over STOP4-7 en onderschrijft de implementatie daarvan. We zijn van mening dat deze paper interessant zou kunnen zijn voor uw lezers, en zodoende bieden wij het paper graag aan ter publicatie in Jeugd in Ontwikkeling. Er is op dit moment geen andere paper gebaseerd op deze dataset in behandeling bij een ander tijdschrift. Alle auteurs hebben een betekenisvolle bijdrage aan het paper geleverd en zijn akkoord met de inhoud daarvan. Wim de Mey is de hoofdontwikkelaar van STOP4-7 maar was niet betrokken bij de analyses voor de huidige studie, en zodoende is er geen sprake van belangenverstrengeling.

We hopen dat u onze paper in overweging neemt.
<geanonimiseerd>

Originele manuscript – 30-8-2021

Evaluatie van Samen sterker Terug Op Pad (STOP4-7):

Een Multimodale Interventie voor Kinderen met Gedragsproblemen

<geanonimiseerd>

Abstract

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. Hiervoor gebruikten wij voormeting, nameting, en follow-up (6 maanden) gegevens van ouders (*Child Behavior Checklist* en Opvoedingsbelasting Vragenlijst) en leerkrachten (*Teacher Report Form*) over 770 kinderen in de leeftijd van 3 t/m 8 jaar. Deze kinderen waren aangemeld bij entrea lindenhout in Nederland en bij verscheidene Vlaamse jeugdhulp instellingen. 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. Deze 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.

Keywords: STOP4-7; interventie; externaliserend; internaliserend; opvoedstress

Evaluatie van Samen sterker Terug Op Pad (STOP4-7):

Een Multimodale Interventie voor Kinderen met Gedragsproblemen

Disruptief gedrag wordt gekenmerkt door ongehoorzaamheid, opstandigheid, prikkelbare stemmingen, en door verbale of fysieke agressie richting anderen (DSM-V; APA, 2013).

Kinderen die op jonge leeftijd al gedragsproblemen 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 storend 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 disruptief gedrag bij kinderen.

Coërcieve Cycli en Ontwikkeling van Probleemgedrag

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 gedrag tussen ouder en kind, ofwel *coercive process*. Wanneer ouders en kinderen elkaar wederzijds negatief bekrachtigen, kan er een neerwaartse spiraal van coërcief gedrag ontstaan. Deze negatieve cyclus kan worden verklaard vanuit de sociale leertheorie zoals uitgewerkt door Patterson (Patterson, 1982; Patterson, 2002; Patterson, Reid & Dishion, 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 toefelijke gedrag van hun ouders door te kalmeren nadat zij hun zin hebben gekregen. Als gevolg van deze coërcieve interactiepatronen in de gezinscontext missen kinderen bij aanvang van hun schoolloopbaan belangrijke vaardigheden, zoals het reguleren van emoties, delen van aandacht, rekening houden met anderen, samen spelen, maar ook volhouden en iets afmaken. Bovendien blijkt dat kinderen die storend gedrag vertonen vaak over gebrekkige mogelijkheden beschikken om sociale signalen op te vangen en te interpreteren (Orobio de Castro et al., 2002). 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.

Ondanks het multifactorieel bepaalde karakter van disruptief gedrag, lieten verschillende studies zien dat oudertrainingen het meest effectieve middel zijn om dergelijk gedrag bij kinderen te verminderen of voorkomen. Uit de 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 (zowel ouders, kind en leerkracht) effectiever zijn dan wanneer slechts een schakel uit het systeem word aangepakt (Kazdin & Weisz, 1998; Lundahl et al., 2006; Webster-Stratton & Taylor, 2001). Multimodale programma's, waarbij naast ouders ook het kind en de leerkracht betrokken worden, blijken ook effectiever te zijn dan programma's die zich uitsluitend richten op ofwel ouders, kinderen, of leerkrachten.

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 4- tot 7-jarige 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, 2000). 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 Vlaamse 201 kinderen tussen 4 en 8 jaar met klinische gedragsproblemen bleek dat het externaliserend gedrag zowel thuis als op school was verbeterd na deelname aan STOP4-7. Ook waren ouders minder inconsequent in hun disciplineren (De Mey & Braet, 2010). 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 wachtlijst controlegroep (De Mey, 2010). Ook in Nederland werden al veelbelovende resultaten voor STOP4-7 gevonden. Uit onderzoek (Van Brussel et al., 2009) onder 69 kinderen tussen de 4 en 8 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 4 tot en met 8 jaar (aangemeld bij vier ouder- en kind centra in Nederland) liet zien dat de aan STOP4-7 verbonden afnames in probleemgedrag stabiel bleven tot aan een 6-maanden follow-up. Daarnaast bleek dat gezinnen konden worden ingedeeld in 3 subgroepen, die vooral verschilden in de mate van de ernst van de problematiek bij aanvang van de interventie.

Comorbiditeit, Lange Termijn Effecten, en Klinische Relevantie

De eerdere bevindingen impliceren dat STOP4-7 mogelijk leidt tot afname van externaliserend probleemgedrag bij kinderen, en tot een afname van opvoedbelasting bij ouders. Het is echter van belang om deze resultaten te repliceren met recente gegevens uit een grotere steekproef van gezinnen afkomstig uit zowel Nederland als Vlaanderen (over alle jeugdhulp locaties heen die STOP4-7 aanbieden). Daarnaast blijft het van belang om de bredere effecten van de interventie te onderzoeken—niet alleen op externaliserend maar ook op internaliserend probleemgedrag. Zeker voor kinderen met ernstige gedragsproblemen geldt dat comorbiditeit van externaliserend gedrag met internaliserend gedrag eerder de norm is dan de uitzondering (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 op zijn beurt 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 positief neveneffect, namelijk een afname van internaliserend probleemgedrag.

Een recente meta-analyse naar—al dan niet multimodale—oudertrainingen (Van Aar et al., 2016) liet zien dat deze over het algemeen significante langere termijn effecten sorteren, tot aan een gemiddelde follow-up periode van 2.5 jaar. ‘Bredere’ effecten van zulke oudertrainingen zijn 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). 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 de meeste interventie studies 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).

Een andere belangrijke vraag betreft 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 kinderen en gezinnen. Een recente ‘review van reviews’ (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 niveaus van risico. Zo liet een recente *individual participant data* meta-analyse (op basis van gepoolde datasets) zien dat 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).

Onderzoeksvragen en Hypothesen

Op basis van eerder onderzoek naar STOP4-7 verwachtten wij dat deze multimodale interventie zou leiden tot een afname van zowel externaliserend als internaliserend probleemgedrag, en de door ouders gerapporteerde opvoedstress, van voor- naar nameting. Andere onderzoeksvragen werden exploratief onderzocht, omdat er voor STOP4-7 nog geen of slechts beperkte eerdere resultaten op dit vlak beschikbaar waren. Het ging hierbij om vragen over i) mogelijk langere termijn effecten van STOP4-7 (tot aan een 6-maanden follow-up), ii) de mogelijke klinische relevantie van afnames in probleemgedrag, en iii) subgroepen voor wie de effecten van STOP4-7 relatief sterk, of juist minder sterk zijn—afhankelijk van het geslacht en de leeftijd van kinderen, en de ernst van hun problemen bij aanvang van de interventie. Omdat onze onderzoeksgegevens uit Nederland en Vlaanderen kwamen, hebben we daarnaast gecontroleerd of de effecten van STOP4-7 vergelijkbaar waren in deze twee landen.

Methode

Design

Om het effect van STOP4-7 te analyseren, werd data verzameld in Nederland en Vlaanderen gecombineerd. Data was beschikbaar voor 3 meetmomenten: voormeting, nameting en follow-up. De voormetingen vonden gemiddeld 1 maand voor implementatie van STOP4-7 plaats. De nametingen vonden plaats op het eind van de groepstraining (gemiddeld 3 maanden na

de start van de interventie) plaats, en de follow-up vond zes maanden na implementatie van STOP4-7 plaats. Dit onderzoek bevatte geen controlegroep.

Participanten

De huidige studie analyseerde de gegevens van $N = 770$ ouders met kinderen in de leeftijd van 3 tot en met 8 jaar die de interventie STOP4-7 hebben doorlopen. Van de totale dataset werd $n = 241$ (31%) verzameld in Nederland en $n = 529$ (68%) in Vlaanderen, in het kader van *routine outcome monitoring*. Alle ouders gaven passieve toestemming voor het gebruik van hun geanonimiseerde gegevens voor wetenschappelijk onderzoek. In Vlaanderen werd eenzelfde toestemmingsprocedure gebruikt, goedgekeurd door het Ethisch Comité van de Universiteit Gent. Het merendeel van de kinderen was jongen ($n = 614$, 80%). De gemiddelde leeftijd van de kinderen was 5,50 jaar ($SD = 1.11$)—precieze data over de leeftijd van kinderen was echter voor slechts 220 kinderen (28%) bekend, wel is bekend dat de leeftijd van alle kinderen varieerde van 4 tot 8 jaar. Ouders van ongeveer de helft van de kinderen in de steekproef ($n = 368$, 47.8%) rapporteerden bij aanvang van de studie klinische niveaus van probleemgedrag (CBCL T-score totaal probleemgedrag ≥ 63).

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 coërcief interactiepatroon tussen ouder en kind, waarbij ouders en kinderen in een wederzijds dwingend gedragspatroon verstrikte zijn geraakt en er geen duidelijke gezagsrelatie meer is binnen het gezin. Ouders rapporteren hierbij vaak onvoldoende opvoedvaardigheden en ervaren vaak 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, Patterson & Snyder, 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, een 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), modelling, 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.

Meetinstrumenten

Gedragsproblemen. Gedragsproblemen bij kinderen werden gemeten middels de Nederlandse versie van de *Child Behavior Checklist* (CBCL; Achenbach, 1991a; Verhulst et al., 1996). 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 drie-puntschaal (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: $\alpha = .82$; Internaliserend probleemgedrag: $\alpha = .87$; Externaliserend probleemgedrag: $\alpha = .78$).

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 (35 items) en Externaliserend probleemgedrag (34 items) meegenomen. Items worden beantwoord op een drie-puntschaal (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: $\alpha = .76$; Internaliserende probleemgedrag: $\alpha = .74$; Externaliserend probleemgedrag: $\alpha = .77$).

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 de totale score. Een voorbeeld item is "Ik voel me gelukkig met mijn kind". Items worden beantwoord op een 4-puntschaal (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: $\alpha = .79$; Problemen in de opvoeding,: $\alpha = .79$; Depressieve stemmingen: $\alpha = .75$; Totale opvoedstress: $\alpha = .766$).

Statistische Analyses

Per kind werd één informant meegenomen in de huidige analyses. Wanneer er voor een kind rapportage van beide ouders beschikbaar was, werd de ouder geselecteerd voor wie de data het meest compleet was (i.e., voor wie er op meer meetmomenten data beschikbaar waren). 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.

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 de CBCL en de TRF. 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 significant afname van probleemgedrag of opvoedstress van voormeting naar follow-up) voerden we *paired sample t-tests* uit om vast te stellen in welke tijdsperiode (i.e., van voor- naar nameting, of van nameting naar follow-up) de significante verandering plaatsvond. Daarna werd bekeken in hoeverre de resultaten afhankelijk waren van de leeftijd van kinderen, 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-kwadraat toetsen 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-kwadraat toetsen bekeken. Een T-score ≥ 63 werd beschouwd als indicatief voor klinische niveaus van opvoedstress (Vermulst et al., 2015).

Resultaten

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 (i.e., CBCL) voor jongens en meisjes voor aanvang van STOP4-7 ($t[636] = .212, p = .832$). Wel rapporteerden ouders minder gedragsproblemen voor jongere kinderen dan voor oudere kinderen (≥ 6 jaar; $t[217] = -2.681, p < .01$). Leerkrachten rapporteerden evenveel gedragsproblemen voor jongens als voor meisjes ($t[441] = -1.06, p = .915$), en voor jongere en oudere kinderen ($t[175] = .502, p = .616$) voor aanvang van STOP4-7. Er waren geen verschillen voor aanvang van STOP4-7 in ervaren opvoedbelasting voor ouders van jongens en meisjes ($t[531] = .034, p = .973$), of ouders van jongere en oudere kinderen ($t[111] = 1.707, p = .091$).

Zowel de samenhang tussen de subschalen internaliserend gedrag en externaliserend gedrag als de samenhang tussen meetmomenten werd bekeken middels Pearson's r correlaties (zie Tabel 2). De correlaties tussen de meetmomenten voor internaliserend gedrag van kinderen zoals gerapporteerd door ouders (i.e., CBCL) waren matig tot sterk positief (r s van .660 tot .720, $p < .001$). Hetzelfde gold voor de autocorrelaties voor externaliserend gedrag (r s van .476 tot .647, $p < .001$). Een vergelijkbaar patroon was zichtbaar voor het door leerkrachten

gerapporteerd internaliserend en externaliserend gedrag van kinderen (r_s van .454 tot .534, $p < .001$, en r_s van .454 tot .653, $p < .001$, respectievelijk).

Analyse van de correlaties tussen de schalen internaliserend gedrag en externaliserend gedrag zoals gerapporteerd door ouders (i.e., CBCL) liet een matig sterke, positieve correlatie zien op alle meetmomenten (r_s van .606 tot .642, $p < .001$). De correlatie tussen internaliserend gedrag en externaliserend gedrag zoals gerapporteerd door leerkrachten (i.e., TRF) was eveneens matig sterk en positief (r_s van .454 tot .483, $p < .001$). Deze resultaten geven aan dat er op alle meetmomenten sprake was van comorbiditeit tussen internaliserend en externaliserend gedrag van kinderen. Deze correlatie nam niet af na deelname aan STOP4-7.

Veranderingen in Probleemgedrag (CBCL)

Ouders rapporteerden een significante afname van totaal probleemgedrag, internaliserend en externaliserend probleemgedrag van kinderen van voormeting tot follow-up (Tabel 3). Het totale probleem gedrag van kinderen nam significant af van voormeting naar nameting ($M_{T0-T1} = 5.95$, $SD = 7.67$, $t[244] = 12.14$, $p < .001$) en stabiliseerde van nameting naar follow-up ($M_{T1-T2} = .76$, $SD = 7.96$, $t[128] = 1.08$, $p = .280$). Wanneer specifieker naar het type probleemgedrag werd gekeken, was een vergelijkbaar patroon zichtbaar. Het internaliserend probleemgedrag van kinderen nam significant af van voormeting naar nameting ($M_{T0-T1} = 4.15$, $SD = 8.59$, $t[244] = 7.55$, $p < .001$), en bleef stabiel van nameting naar follow-up ($M_{T1-T2} = .78$, $SD = 8.29$, $t[128] = 1.07$, $p = .286$). Dit was eveneens het geval voor externaliserend probleem gedrag van kinderen ($M_{T0-T1} = 6.91$, $SD = 7.90$, $t[244] = 13.69$, $p < .001$; $M_{T1-T2} = .46$, $SD = 8.53$, $t[128] = .61$, $p = .544$).

De veranderingen op zowel het totale probleemgedrag als internaliserend probleemgedrag 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, kinderen die op baseline meer probleemgedrag hadden dan hun leeftijdgenoten, hadden dit op de nameting en follow-up doorgaans ook. De gerapporteerde afname van probleemgedrag was even sterk voor jongens als voor meisjes, en was ook even sterk in Vlaamse als in Nederlandse jeugdhulp locaties.

De verandering in externaliserend gedrag bleek afhankelijk van de locatie van implementatie van STOP4-7. In zowel Nederland als Vlaanderen was de afname van voormeting naar nameting significant ($M_{T0-T1} = 6.41$, $SD = 7.44$, $t[151] = 10.62$, $p < .001$, en $M_{T0-T1} = 7.708$, $SD = 8.56$, $t[92] = 8.681$, $p < .001$), echter was deze afname groter in Vlaanderen. Voor zowel Nederland als Vlaanderen was de verandering van nameting naar follow-up niet significant ($M_{T1-T2} = -.33$, $SD = 8.11$, $t[65] = -.33$, $p = .740$ en $M_{T1-T2} = 1.28$, $SD = 8.94$, $t[62] = 1.142$, $p = .258$, respectievelijk). De verandering in externaliserend gedrag van kinderen was niet afhankelijk van leeftijd, mate van problemen bij aanvang of de locatie van de interventie.

Veranderingen in Probleemgedrag (TRF)

Leerkrachten rapporteerden een significante afname in het totale probleemgedrag en externaliserend probleemgedrag van kinderen van voormeting naar follow-up, maar niet in het internaliserend gedrag (zie Tabel 3). Het totale probleemgedrag van kinderen nam volgens leerkrachten af van voormeting naar nameting ($M_{T0-T1} = 2.26$, $SD = 7.28$, $t[242] = 4.83$, $p < .001$), en stabiliseerde van nameting naar follow-up ($M_{T1-T2} = -1.07$, $SD = 9.03$, $t[112] = -1.260$, $p = .210$). Ook externaliserend gedrag van kinderen nam volgens leerkrachten significant af van voormeting naar nameting ($M_{T0-T1} = 2.61$, $SD = 7.51$, $t[242] = 5.422$, $p < .001$) en stabiliseerde van nameting naar follow-up ($M_{T1-T2} = -1.212$, $SD = 9.10$, $t[112] = -1.416$, $p = .160$). Voor internaliserend gedrag was de verandering van voormeting naar nameting ($M_{T0-T1} = .90$, $SD =$

8.93, $t[242] = 1.57$, $p = .119$) evenals de verandering van nameting naar follow-up ($M_{T1-T2} = -.82$, $SD = 9.64$, $t[112] = -.91$, $p = .366$) niet significant.

Zowel de veranderingen op het totale probleemgedrag gerapporteerd door leerkrachten als het externaliserend 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 is even sterk voor Vlaamse kinderen als voor Nederlandse kinderen. De verandering in het 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 totaal probleemgedrag van voormeting naar nameting ($M_{T0-T1} = .126$, $SD = 6.846$, $t[126] = .207$, $p = .836$), terwijl dit voor kinderen die op voorhand in de klinische range scoorden wel het geval was ($M_{T0-T1} = 4.594$, $SD = 7.056$, $t[115] = 7.013$, $p < .001$). Beide groepen veranderden niet significant van nameting naar follow-up ($M_{T1-T2} = .113$, $SD = 7.223$, $t[52] = .114$, $p = .910$ en $M_{T1-T2} = -.629$, $SD = 10.960$, $t[34] = -.339$, $p = .736$, respectievelijk).

Het externaliserend probleemgedrag van kinderen die op voorhand volgens leerkrachten in de niet-klinische range scoorden, veranderde niet significant van voormeting naar nameting, terwijl dit voor kinderen die op voorhand in de klinische range scoorden wel het geval was ($M_{T0-T1} = .0617$, $SD = 7.343$, $t[96] = .083$, $p = .934$ en $M_{T0-T1} = 4.308$, $SD = 7.158$, $t[145] = 7.273$, $p < .001$, respectievelijk). Van nameting tot follow-up nam het externaliserend gedrag van kinderen die op voorhand in de niet-klinische range scoorden toe ($M_{T1-T2} = -2.359$, $SD = 7.02$, $t[38] = -2.074$, $p < .05$), terwijl dit stabiel bleef voor kinderen die op voorhand in de klinische range van probleemgedrag scoorden ($M_{T1-T2} = -.857$, $SD = 9.923$, $t[48] = .605$, $p = .548$). Het

internaliserend probleemgedrag van kinderen die op voorhand volgens leerkrachten in de niet-klinische range scoorden nam toe van voormeting tot nameting, terwijl dit afnam voor kinderen die op voorhand in de klinische range scoorden ($M_{T0-T1} = -2.103$, $SD = 9.247$, $t[96] = -2.240$, $p < .05$ en $M_{T0-T1} = 2.890$, $SD = 8.147$, $t[145] = 4.223$, $p < .001$, respectievelijk). Voor beide groepen was de verandering van nameting tot follow-up niet-significant ($M_{T1-T2} = 0.026$, $SD = 8.552$, $t[38] = .019$, $p = .985$ en $M_{T1-T2} = .286$, $SD = 10.766$, $t[48] = .186$, $p = .853$).

De verandering op het internaliserend gedrag van kinderen bleek ook afhankelijk van het geslacht van kinderen. Voor meisjes was de afname van voormeting naar nameting significant ($M_{T0-T1} = 4.03$, $SD = 9.29$, $t[32] = 2.49$, $p < .05$), maar voor jongens niet ($M_{T0-T1} = .030$, $SD = 8.48$, $t[166] = .046$, $p = .964$). Voor zowel meisjes als jongens was de verandering van nameting naar follow-up niet significant ($M_{T1-T2} = -.31$, $SD = 15.32$, $t[15] = -.08$, $p = .936$ en $M_{T1-T2} = -.26$, $SD = 8.55$, $t[86] = -.288$, $p = .774$, respectievelijk). De verandering in internaliserend gedrag van kinderen was niet afhankelijk van leeftijd of de locatie van de interventie.

Veranderingen in Opvoedstress (OBVL)

De door ouders gerapporteerde totale opvoedstress nam significant af van voormeting tot follow-up (zie Tabel 4). Deze afname was zichtbaar van voormeting naar nameting ($M_{T0-T1} = 5.29$, $SD = 8.29$, $t[281] = 10.73$, $p < .001$) en stabiliseerde van nameting naar follow-up ($M_{T1-T2} = .27$, $SD = 8.69$, $t[69] = .261$, $p = .795$). Wanneer naar de subschalen van de OBVL werd gekeken, was een vergelijkbaar patroon zichtbaar. Gerapporteerde opvoedstress door problemen in de ouder-kind relatie (OKR), problemen in de opvoeding (PO) en depressieve stemmingen (DEP) namen significant af van voormeting naar follow-up. De afname van gerapporteerde belasting was steeds significant van voormeting naar nameting (OKR: $M_{T0-T1} = 3.92$, $SD = 8.06$, $t[281] = 8.170$, $p < .001$; PO: $M_{T0-T1} = 6.24$, $SD = 8.26$, $t[281] = 12.68$, $p < .001$; DEP: $M_{T0-T1} =$

2.87, $SD = 6.86$, $t[280] = 7.01$, $p < .001$) en stabiliseerde van nameting naar follow-up (OKR: $M_{T1-T2} = -1.000$, $SD = 7.94$, $t[69] = -1.054$, $p = .296$; PO: $M_{T1-T2} = .75$, $SD = 7.56$, $t[68] = .828$, $p = .410$; DEP: $M_{T1-T2} = -3.00$, $SD = 7.91$, $t[69] = -.317$, $p = .752$).

De gerapporteerde veranderingen van problemen in de opvoeding, depressieve stemmingen en totale opvoedstress was niet afhankelijk van het geslacht van het kind, de 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 voormeting meer opvoedstress rapporteren ten opzichte van andere ouders, hebben dit op de nameting en follow-up doorgaans ook. De gerapporteerde afname in opvoedstress was even sterk voor ouders van jongens als meisjes, voor oudere en jongere kinderen, en voor Vlaamse en Nederlandse kinderen.

Wanneer ouders op voorhand in de klinische range van opvoedstress scoorden namen problemen in de ouder-kind relatie significant af van voormeting naar nameting ($M_{T0-T1} = 5.289$, $SD = 7.902$, $t[189] = 9.226$, $p < .001$), maar dit was niet het geval wanneer ouders op voorhand in de niet-klinische range van opvoedstress scoorden ($M_{T0-T1} = 1.098$, $SD = 7.680$, $t[91] = 1.371$, $p = .174$). Voor beide groepen was de verandering van nameting tot follow-up niet-significant ($M_{T0-T1} = -2.250$, $SD = 10.199$, $t[11] = -.764$, $p = .461$ en $M_{T0-T1} = -1.109$, $SD = 7.596$, $t[51] = -.968$, $p = .338$, respectievelijk).

Klinische Relevantie

Tabel 5 geeft het aantal kinderen weer dat op de verschillende meetmomenten in de klinische range scoorden op de CBCL. Dit aantal nam over de meetmomenten significant af. Voor totale gedragsproblemen (i.e., zowel internaliserend als externaliserend) gold een significant verschil in het aantal kinderen in de klinische range tussen voormeting en nameting ($\chi^2 [1, N = 245] = 48.815$, $p < .001$), en tussen voormeting en follow-up ($\chi^2 [1, N = 119] =$

12.241, $p < .01$). Van de 141 kinderen die op voormeting in de klinische range van probleemgedrag scoorden, scoorden op nameting 47.5% ($n = 67$) in de niet-klinische range. Op follow-up scoorde 50.6% ($n = 43$) in de niet-klinische range van de kinderen die op voormeting nog in de klinische range scoorden ($n = 85$).

Dit patroon was ook zichtbaar wanneer afzonderlijk naar internaliserend en externaliserend probleemgedrag werd gekeken. Van de 101 kinderen die op voormeting in de klinische range scoorden op internaliserend probleemgedrag, scoorden op nameting 47.5% ($n = 48$) in de niet-klinische range ($\chi^2 [1, N = 245] = 59.343, p < .001$). Op follow-up scoorde 42.1% ($n = 24$) van de kinderen in de niet-klinische range die op voormeting in de klinische range scoorden ($n = 57; \chi^2 [1, N = 119] = 33.926, p < .001$). Van de 180 kinderen die op voormeting in de klinische range scoorden op externaliserend probleemgedrag, scoorde 46.7% ($n = 84$) in de niet-klinische range op nameting ($\chi^2 [1, N = 245] = 44.004, p < .001$). Op follow-up scoorde 46.9% ($n = 46$) van de kinderen in de niet-klinische range die op voormeting in de klinische range scoorden ($n = 98; \chi^2 [1, N = 119] = 13.225, p < .001$).

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 ($\chi^2 [1, N = 282] = 64.742, p < .001$). 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 ($n = 61; \chi^2 [1, N = 79] = 5.314, p = .055$).

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 zijn voor het merendeel middelgroot tot groot te noemen (partial η^2 .15-.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 een gemiddelde Cohen's *d* van .27 (Menting et al., 2013) en .42 (Lundahl et al., 2006) hadden. Echter, ondanks de aangetoonde veranderingen in gerapporteerd probleemgedrag 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—is er 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, Tømmerås, & Ogden, 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 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).

Ook voor het totale probleemgedrag en externaliserend 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 naar de CBCL en TRF in 21 landen toonde aan dat de overeenstemming tussen ouder- en leerkrachtrapportage van probleemgedrag matig is ($r = .29$), en dat ouders vaak meer gedragsproblemen rapporteren dan leerkrachten (Rescorla et al., 2014). Mogelijk verklaart deze discrepantie in het oordeel tussen ouders en leerkrachten onze bevinding—leerkrachten beschouwen gedrag mogelijk eerder als verbeterd ten opzichte van ouders, en dit is mogelijk met name het geval wanneer klinische gedragsproblemen erg aanwezig zijn. Een alternatieve verklaring zou kunnen zijn dat disruptief gedrag zich anders voordoet in de school- en gezinscontext (De Los Reyes et al., 2013).

Tenslotte vonden we dat de problemen in de ouder-kind relatie 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 (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 in Nederland resultaten over 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. Tegelijkertijd kent de studie ook enkele beperkingen. Om de bewijskracht voor STOP4-7 te versterken zou in vervolgonderzoek, eventueel middels een wachtlijstconstructie, 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 coërcieve cycli tussen ouders en kinderen. 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. Bovendien zouden meer gegevens moeten worden verzameld over de implementatiegraad en -kwaliteit van STOP4-7, zodat kan worden geanalyseerd of gezinnen met een hulpvraag mogelijksterwijs nog effectiever geholpen zouden kunnen worden.

Concluderend bleek uit deze 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.

Referenties

- 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.
- 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., & Kunder, 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/10.1177/106342660000800304>
- 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>
- Hukkelberg, A., Tømmerås, T., & Ogden, . (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>
- 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>
- Moffitt, T. E., Arseneault, L., Belsky, D., Dickson, N., Hancox, R. J., Harrington, H., ... & 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%2F10634266020100030601>
- 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/s11121-020-01176-6>

- 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>
- 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., ... & 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 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>

<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>

Tabel 1*Gemiddelden en Standaarddeviaties.*

	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>
CBCL									
Totaal		63.69	9.72		58.21	10.55		59.30	10.73
Internaliserend	63	59.19	10.33	328	54.90	10.73	195	55.33	11.41
Externaliserend	9	66.98	9.88		60.46	9.98		61.31	10.13
TRF									
Totaal		57,71	9,43		61.44	8.36		61.98	9.25
Internaliserend	58	66,73	9,55	322	55.52	8.36	196	56.13	10.61
Externaliserend	3	64,47	8,94		63.43	8.33		63.62	8.79
OBVL									
Totaal		65.15	10.15	337	60.28	10.51	90	60.48	10.70
Ouder-kind relatie		63.94	9.93	337	60.71	9.59	90	62.48	9.98
Problemen in opvoeding	53 3	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

Tabel 2 *Pearson Correlaties.*

	1.	2.	3.	4.	5.	6.	7.	8.	9.
1. INT-T0	--	.457***	.454***	.454***	.291***	.203**	.704***	.341***	.292***
2. INT-T1	.660***	--	.534***	.225***	.447***	.143	.329***	.655***	.301***
3. INT-T2	.675***	.720***	--	.199*	.178	.483***	.337***	.368***	.728***
4. EXT-T0	.606***	.393***	.342***	--	.653***	.529***	.895***	.568***	.474***
5. EXT-T1	.364***	.624***	.438***	.647***	--	.454***	.650***	.911***	.437***
6. EXT-T2	.309***	.392***	.642***	.476***	.606***	--	.485***	.403***	.896***
7. TOT-T0	.836***	.596***	.543***	.871***	.609***	.443***	--	.636***	.513***
8. TOT-T1	.533***	.831***	.605***	.568***	.900***	.559***	.691***	--	.493***
9. TOT-T2	.482***	.586***	.833***	.450***	.608***	.900***	.569***	.701***	--

Noot. Correlaties voor CBCL onder de diagonaal, voor TRF boven de diagonaal; Grijs gekleurde vakken geven de correlatie tussen subschalen aan. T0 = voormeting, T1 = nameting, T2 = follow-up meting; * $p < .05$, *** $p < .001$.

Tabel 3

Uitkomsten Repeated Measures ANOVA: Van Voormeting (T0) tot Follow-Up (T2) op CBCL en TRF.

	CBCL			TRF		
	<i>F</i>	<i>p</i>	<i>Par. Eta²</i>	<i>F</i>	<i>p</i>	<i>Par. Eta²</i>
Totaal probleemgedrag	48.673	.000	.327	3.13	.050*	.035
* Geslacht kind	.897	.467	.018	1.45	.238*	.017
* Leeftijd kind	.101	.904	.002	.712	.472*	.018
* Baseline gedragsproblemen	2.071	.129	.020	9.618	.000*	.101
* Implementatie locatie	2.734	.067	.027	.340	.695*	.004
Internaliserend gedrag	23.819	.000	.192	.033	.968	.000
* Geslacht kind	1.326	.262	.026	4.216	.016	.047
* Leeftijd kind	1.749	.178	.028	.321	.686*	.008
* Baseline gedragsproblemen	2.671	.072	.026	4.799	.011	.053
* Implementatie locatie	.334	.713	.003	.582	.560	.007
Externaliserend gedrag	53.994	.000	.351	5.022	.010*	.055
* Geslacht kind	.972	.424	.019	.539	.567*	.006
* Leeftijd kind	.075	.928	.001	.579	.447*	.019
* Baseline gedragsproblemen	1.679	.189	.017	9.779	.000*	.102
* Implementatie locatie	3.606	.029	.035	.051	.938	.001

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

Tabel 4.*Uitkomsten Repeated Measures ANOVA: Voormeting naar Follow-Up op de OBVL.*

	Totale Opvoedstress			Ouder-Kind Relatie			Problemen in Opvoeding			Depressieve Stemmingen		
	<i>Pa</i>			<i>Pa</i>			<i>Pa</i>			<i>Pa</i>		
	<i>F</i>	<i>p</i>	<i>r. Eta 2</i>	<i>F</i>	<i>p</i>	<i>r. Eta 2</i>	<i>F</i>	<i>p</i>	<i>r. Eta 2</i>	<i>F</i>	<i>p</i>	<i>r. Eta 2</i>
Hoofdeffe	27.5	.00	.30	17.0	.00	.21	46.7	.00	.43	11.7	.00	.15
ct	97	0	5	62	0	3	73	0	0	56	0	7
*	.008	.99	.00	.570	.56	.00	.884	.42	.04	.572	.56	.00
Geslacht		2	0		7	9		2	9		6	9
kind												
*	.227	.78	.00	.429	.65	.01	1.39	.25	.04	.583	.56	.01
Leeftijd		9	7		3	4	1	6	3		1	8
kind												
*	2.51	.08	.03	6.18	.00	.09	.883	.41	.01	2.15	.12	.03
Baseline	1	8	9	3	3	1		6	4	6	0	4
opvoedstre												
ss												
*	1.10	.33	.01	1.09	.33	.01	.881	.41	.01	.193	.82	.00
Implement	7	4	8	8	7	7		7	4		4	3
atie locatie												

Noot. Dikgedrukte resultaten zijn statistisch significant.

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 Probleemgedrag en Opvoedstress.

	Voormeting			Nameting			Follow-up		
	<i>N</i>	Niet- klinisc h <i>n</i> (%)	Klinisc h <i>n</i> (%)	<i>N</i>	Niet- klinisc h <i>n</i> (%)	Klinisc h <i>n</i> (%)	<i>N</i>	Niet- klinisc h <i>n</i> (%)	Klinisc h <i>n</i> (%)
Totaal		271	368	212	116		118	77	
probleemgedrag		(42.4)	(57.6)	(64.6)	(35.4)		(60.5)	(39.5)	
Internaliserend gedrag	63	379	260	32	242	86	19	137	58
	9	(59.3)	(40.7)	8	(73.8)	(26.2)	5	(70.3)	(29.7)
Externaliserend gedrag		188	451		196	132		110	85
		(29.4)	(70.6)		(59.8)	(40.2)		(56.4)	(43.6)
Opvoedstress	53	182	351	33	176	161	90	45	45 (50)
	3	(34.1)	(65.9)	7	(52.2)	(47.8)		(50)	

Reactie reviewer 1 – 24-11-2021

Evaluatie van Samen sterker Terug op Pad (STOP4-7): Een multimodale interventie voor kinderen met gedragsproblemen

In deze paper wordt de effectiviteit van het STOP4-7 programma nagegaan op probleemgedrag bij kinderen en opvoedstress bij ouders aan de hand van een pre- post, en follow-up design. De studie heeft veel belangrijke sterktes. Zo is STOP4-7 een intensief programma dat gebaseerd is op een duidelijk theoretisch kader. Doorheen de jaren werd met dit programma een unieke, grote, en diverse steekproef verzameld. De huidige studie biedt wellicht één van de grootste en meest volledige toetsen van de effectiviteit van het programma, ook al is het ontbreken van een controlegroep (zoals ook aangestipt door de auteurs) een zeer belangrijke beperking. Een andere sterkte is het gebruik van een multi-informant benadering om de uitkomsten te meten, waarbij niet enkel ouders maar ook leerkrachten deelnamen aan het onderzoek. De data werden adequaat geanalyseerd en de resultaten worden helder weergegeven. Alles bij elkaar is dit een waardevolle paper die voor veel lezers van Jeugd in Ontwikkeling relevant zal zijn. Hieronder geef ik nog enkele grotere en kleinere suggesties mee om de bijdrage van de paper verder te versterken.

De innovatieve bijdrage van deze studie aan de literatuur zou meer uit de verf mogen komen. Op p. 5 lijkt het alsof deze studie vooral een replicatie is van eerder onderzoek, met een wat grotere steekproef en bijkomende aandacht voor internaliserende problemen als een uitkomst. Echter, uit de paragraaf daarvoor blijkt dat een paar eerdere studies ook al effecten vonden op internaliserende problemen en dat sommige studies ook al beroep deden op eerder grote steekproeven. Om het vernieuwende karakter van internaliserende problemen als afhankelijke variabele meer te beklemtonen zouden de auteurs kunnen wachten met het vermelden van deze variabele tot in de paragraaf waarin ze de vernieuwende bijdrage van hun eigen studie bespreken. Daar kunnen ze dan zeggen dat tot nu slechts één (juist?) eerdere studie het effect van STOP4-7 naging op internaliserende problemen.

Nog een manier om de vernieuwende aspecten van de studie te verhelderen is door vroeger en meer uitgebreid het belang te bespreken van onderzoek naar moderatoren van de effectiviteit van het programma. Dit lijkt me een meer wezenlijke bijdrage van deze studie, maar de auteurs staan er slechts kort bij stil helemaal op het einde van de inleiding. In dit opzicht vond ik het interessant dat er ook analyses werden gedaan om na te gaan of de effectiviteit van het programma ook afhangt van de mate van problemen die kinderen bij aanvang vertonen. Deze analyses komen echter wat als een verrassing en zouden beter geïntegreerd en aangekondigd kunnen worden in de Inleiding.

Wat betreft het effect van STOP4-7 op internaliserende problemen lijken de auteurs er sterk vanuit te gaan dat dit effect een secundair gevolg is van het effect van het programma op externaliserende problemen. Ze lijken hiermee opnieuw wat afbreuk te doen aan dit potentieel innovatieve aspect van hun eigen studie. Als dit veronderstelde effect het gevolg is van een eerder reeds robuust aangetoond effect, wat is dan de grote meerwaarde ervan? Zou het niet kunnen dat STOP4-7 ook een direct en uniek effect heeft op internaliserende problemen, bijvoorbeeld omdat kinderen beter met hun gevoelens weten om te gaan en omdat ze zich beter begrepen voelen door hun ouders nadat zij de interventie hebben doorlopen?

Is een vernieuwend aspect van deze studie ook dat er gebruik gemaakt wordt van leerkrachtrapportage? Of is dit standaard in dit type onderzoek? Sowieso zou iets meer aandacht mogen geschonken worden aan het belang en de sterkte van deze multi-informant benadering. Vraageffecten en sociale wenselijkheid spelen vermoedelijk minder een rol bij leerkrachten dan

bij ouders, en dus vormt de leerkrachtrapportage een sterkere en meer conservatieve toets van de hypothesen.

Zijn er ook cijfers beschikbaar over uitval tijdens het programma? Ongeveer hoeveel ouders (van alle ouders die starten met het programma) doorlopen het programma tot op het einde? En wat betekent dit voor de implicaties van deze studie? Moet de vrij optimistische conclusie van het artikel niet wat getemperd worden indien blijkt dat veel ouders er niet in slagen om het programma af te ronden?

Is het niet uitzonderlijk om vast te stellen dat er geen geslachtsverschillen waren inzake internaliserende en externaliserende gedragsproblemen (p. 12)? Hoe verhoudt deze bevinding zich tot eerdere literatuur of wat kan hiervoor een verklaring zijn? Speelt hierbij een selectie-effect waarbij alle deelnemers aan het onderzoek (jongens en meisjes) al verhoogde probleemscores vertoonden?

Hoe sterk was het verband tussen ouder-gerapporteerde en door de leerkracht gerapporteerde probleemgedragingen? Deze correlaties zouden eventueel op de diagonaal van Tabel 2 kunnen toegevoegd worden.

Gelieve voor elk van de repeated measures ANOVAs de effectgrootte weer te geven van de gevonden effecten (bijvoorbeeld eta squared) en de effectgrootte ook expliciet te bespreken. Dit gebeurt nu kort in de Discussie maar kan best ook al in de Resultaten gebeuren.

Kleinere suggesties en vragen:

Het is niet nodig om de namen van de gebruikte vragenlijsten te vermelden in de Abstract. De term ‘disruptief gedrag’ in de eerste zin zou ik vervangen door ‘storend gedrag’.

Zou de moeilijke term ‘coërcieve cyclus’ ook kunnen vervangen worden door ‘dwingende cyclus’?

Op p. 3, in de paragraaf die begint met ‘Ondanks het multifactorieel bepaalde karakter ...’, komen verschillende zinnen voor met sterke uitspraken, die niet worden ondersteund door een referentie.

Op p. 7 begreep ik niet waarom de lange-termijn effecten van het STOP4-7 programma slechts op een exploratieve manier worden onderzocht? De verwachting is toch dat de veranderingen kort na het programma behouden blijven op de langere termijn? Bovendien toonden een paar eerdere studies dit al aan, dus waarom niet gewoon de gerichte hypothese vooropstellen dat er follow-up effecten zullen zijn?

Op p. 11 is het enigszins verwarrend dat er ‘per kind één informant werd meegenomen in de analyses’. Betreft deze uitspraak enkel de moeders en vaders als informanten of ook de leerkrachten? Hierbij aansluitend: kunnen de auteurs ook preciseren van hoeveel moeders en hoeveel vaders uiteindelijk data werden gebruikt in de analyses? En bij hoeveel kinderen precies waren er ook leerkrachtdata beschikbaar?

Gelieve op p. 13 en 14 meer betekenisvolle titels te gebruiken, die verwijzen naar ouder- en leerkrachtrapportage, respectievelijk. De afkortingen CBCL en TRF zullen voor sommige lezers niet betekenisvol zijn.

Reactie reviewer 2 – 24-11-2021

Er wordt in dit manuscript een interessant en relevant onderzoek voorgesteld. Ik heb nog enkele vragen ter verduidelijking voor de auteurs.

Abstract

- De zin ‘hiervoor gebruikten wij...’ dient herschreven te worden. Misschien goed om het design van de interventiestudie te benoemen en vervolgens aan te geven welke data werden verzameld.

Inleiding

- De auteurs starten het manuscript met een definitie van ‘disruptief gedrag’, maar spreken zowel in de titel als verder in het manuscript over gedragsproblemen. Ik vraag me af of dit volledig hetzelfde is. Disruptief gedrag wordt volgens mij voornamelijk genoemd in het kader van gedragsstoornissen en minder met betrekking tot gedragsproblemen. Dit alles hangt natuurlijk erg nauw samen, maar het lijkt me goed hier nog een keer kritisch naar te kijken zodat er zeker naar de juiste terminologie en definitie wordt verwezen en dit ook door te trekken doorheen het manuscript.
- Bij de zinnen die beginnen met ‘zij gaan in sociale situaties weinig...’ moet nog een referentie toegevoegd worden.
- Er wordt gesproken over het multifactorieel karakter van disruptief gedrag. Echter, uit het voorgaande gaat men enkel in op het coërcieve interactiepatroon. Misschien goed om dit bij het begin wat verder open te trekken en kort in te gaan op het multifactoriele karakter zodat er hier later op kan teruggegaan worden.
- Er wordt verwezen naar de effectiviteit van bepaalde oudertrainingen. Echter, de referenties staan enkel samen op het einde van deze alinea. Het zou goed zijn moest er hier meer precies gerefereerd kunnen worden. Welke conclusies worden getrokken o.b.v. welke onderzoeken. Rond multimodale programma’s moet zeker nog een referentie toegevoegd worden.
- Fout in zin: Uit eerder onderzoek naar STOP4-7 onder Vlaamse 201 kinderen tussen 4 en 8 jaar met klinische gedragsproblemen bleek dat het externaliserend gedrag zowel thuis als op school was verbeterd na deelname aan STOP4-7.
- In de alinea die start met ‘Een recente meta- analyse...’ beschrijven de auteurs verschillende mogelijke aspecten die onderzocht kunnen worden. Het is echter niet helemaal duidelijk hoe dit linkt aan het voorgestelde onderzoek en hoe dit meegenomen wordt. Op basis van de volgende alinea die start met ‘een andere belangrijke vraag’ ga ik ervan uit dat in voorgaande alinea ook verschillende vragen beschreven worden. Deze kunnen volgens mij meer expliciet beschreven worden.
- Een review-van-reviews kan volgens mij omschreven worden als een meta-review.
- Ik vraag me af waarom bepaalde vragen exploratief bekeken worden. Hier kunnen op basis van eerdere onderzoeken (m.b.t. andere programma’s), waar de onderzoekers ook naar verwijzers, toch ook hypothesen worden geformuleerd.

Methode

- Zin 1 van design, het woord ‘gecombineerd’ schrappen.

- De ‘gemiddelde’ momenten waarop de voor-, nameting en follow up plaatsvond wordt beschreven. Het lijkt me zinvol hier de range aan toe te voegen waarbinnen op elk moment de data werden verzameld.
- In de voorlaatste zin van het design staat een keer ‘plaats’ te veel
- APA: cijfers ten en met tien moeten voluit geschreven worden.
- Misschien goed om nog toe te lichten waarom van de Opvoedbelasting Vragenlijst enkel de drie eerste subschalen werden gebruikt.
- Er wordt weinig informatie gegeven over de respondentengroep. Over de kinderen is de informatie beperkt, maar is als ik het goed heb niet meer informatie beschikbaar. Over de ouders en leerkrachten wordt geen informatie gegeven. Geslacht, leeftijd, aantal kinderen, aantal jaren ervaring,... Zij deze gegevens beschikbaar. Het lijkt me interessant deze toe te voegen. Daarnaast vraag ik me af of er voor elk meetmoment van alle informatie voor de 770 kinderen beschikbaar was. Ook deze info dient volgens mij toegevoegd te worden.

Resultaten

- Het lijkt me goed om de verwoording in de resultaten nog een keer goed na te kijken. Bijvoorbeeld p14. ‘kinderen die op baseline meer probleemgedrag hadden dan hun leeftijdgenoten, hadden dit op de nameting en follow-up doorgaans ook.’ Moet volgens mij omschreven worden als: ‘kinderen waarvoor ouders tijdens de baselinemeting meer probleemgedrag rapporteerden dan hun leeftijdgenoten..’
- P. 15 spreken de auteurs meermaals over ‘de kinderen veranderden niet’. Ik begrijp niet helemaal wat hiermee wordt bedoeld. Bedoelen de auteurs hiermee veranderen van klinisch naar niet-klinisch niveau? Ik denk dat hier wat toelichting of een andere beschrijving nodig is.

Discussie

- P. 22 zin herbekijken: ‘Dit onderzoek geeft al eerste onderzoek in Nederland...’

Algemeen

- In het hele document moeten de referenties nagekeken worden. APA 7: vanaf 3 auteurs in de tekst enkel de eerste noemen.

Reactiebrief auteurs – 21-12-2021

Amsterdam, 21 december 2021

Beste redactieleden van Jeugd in Ontwikkeling,

Hartelijk dank voor de mogelijkheid om onze paper getiteld “*Evaluatie van Samen sterker Terug Op Pad (STOP4-7): Een Multimodale Interventie voor Kinderen met Gedragsproblemen*” te herzien. Hartelijk dank ook aan de reviewers voor hun aardige, opbouwende feedback op ons manuscript. Hieronder vindt u onze respons op de door de reviewers aangedragen punten en een overzicht van de gedane aanpassingen. We hopen dat de paper met deze aanpassingen voldoet aan de standaard van Jeugd in Ontwikkeling. We zien uit naar uw reactie.
<geanonimiseerd>

Reviewer #1

- De innovatieve bijdrage van deze studie aan de literatuur zou meer uit de verf mogen komen. Op p. 5 lijkt het alsof deze studie vooral een replicatie is van eerder onderzoek, met een wat grotere steekproef en bijkomende aandacht voor internaliserende problemen als een uitkomst. Echter, uit de paragraaf daarvoor blijkt dat een paar eerdere studies ook al effecten vonden op internaliserende problemen en dat sommige studies ook al beroep deden op eerder grote steekproeven. Om het vernieuwende karakter van internaliserende problemen als afhankelijke variabele meer te beklemtonen zouden de auteurs kunnen wachten met het vermelden van deze variabele tot in de paragraaf waarin ze de vernieuwende bijdrage van hun eigen studie bespreken. Daar kunnen ze dan zeggen dat tot nu slechts één (juist?) eerdere studie het effect van STOP4-7 naging op internaliserende problemen.
- Nog een manier om de vernieuwende aspecten van de studie te verhelderen is door vroeger en meer uitgebreid het belang te bespreken van onderzoek naar moderatoren van de effectiviteit van het programma. Dit lijkt me een meer wezenlijke bijdrage van deze studie, maar de auteurs staan er slechts kort bij stil helemaal op het einde van de inleiding. In dit opzicht vond ik het interessant dat er ook analyses werden gedaan om na te gaan of de effectiviteit van het programma ook afhangt van de mate van problemen die kinderen bij aanvang vertonen. Deze analyses komen echter wat als een verrassing en zouden beter geïntegreerd en aangekondigd kunnen worden in de Inleiding.
 - Hartelijk dank voor de suggesties om de bijdrage van het huidige onderzoek aan de literatuur krachtiger weer te geven. We hebben het vernieuwende karakter van het onderzoek duidelijker in de inleiding beschreven.
 - Om het vernieuwende karakter van ons onderzoek aan te zetten, hebben we aan de inleiding een korte alinea toegevoegd waarin dit expliciet wordt benoemd (pagina 3, regel 15-22).
 - Op pagina 7 (vanaf regel 10) en pagina 8 hebben we duidelijker verwoord waarom het van belang is om zowel naar de veranderingen van externaliserend probleemgedrag als internaliserend probleemgedrag te kijken.
 - Daarnaast hebben we op pagina 9 (vanaf regel 1) het belang van onderzoek naar moderatoren van de effectiviteit van STOP4-7 duidelijker verwoord.

- Wat betreft het effect van STOP4-7 op internaliserende problemen lijken de auteurs er sterk vanuit te gaan dat dit effect een secundair gevolg is van het effect van het programma op externaliserende problemen. Ze lijken hiermee opnieuw wat afbreuk te doen aan dit potentieel innovatieve aspect van hun eigen studie. Als dit veronderstelde effect het gevolg is van een eerder reeds robuust aangetoond effect, wat is dan de grote meerwaarde ervan? Zou het niet kunnen dat STOP4-7 ook een direct en uniek effect heeft op internaliserende problemen, bijvoorbeeld omdat kinderen beter met hun gevoelens weten om te gaan en omdat ze zich beter begrepen voelen door hun ouders nadat zij de interventie hebben doorlopen?
 - Dank u wel voor deze waardevolle suggestie. We hebben dit aangepast op pagina 8 (regel 14-16).

- Is een vernieuwend aspect van deze studie ook dat er gebruik gemaakt wordt van leerkrachtrapportage? Of is dit standaard in dit type onderzoek? Sowieso zou iets meer aandacht mogen geschonken worden aan het belang en de sterkte van deze multi-informant benadering. Vraageffecten en sociale wenselijkheid spelen vermoedelijk minder een rol bij leerkrachten dan bij ouders, en dus vormt de leerkrachtrapportage een sterkere en meer conservatieve toets van de hypothesen.
 - We hebben het gebruik van onze multi-informant benadering duidelijker besproken in de inleiding (pagina 3, regel 19-20; pagina 7, regel 3-5) en het voordeel hiervan expliciet genoemd in de discussie (pagina 28, regel 8-9).

- Zijn er ook cijfers beschikbaar over uitval tijdens het programma? Ongeveer hoeveel ouders (van alle ouders die starten met het programma) doorlopen het programma tot op het einde? En wat betekent dit voor de implicaties van deze studie? Moet de vrij optimistische conclusie van het artikel niet wat getemperd worden indien blijkt dat veel ouders er niet in slagen om het programma af te ronden?
 - Hartelijk dank voor deze terechte vraag. Helaas zijn er geen gepubliceerde cijfers bekend over de uitval bij STOP4-7. Wel weten we uit persoonlijke communicatie met uitvoerders van het programma dat de uitval laag is. Wanneer zij aan het programma zijn begonnen, stopt slechts één op de tien tot zestien gezinnen vroegtijdig met STOP4-7. Dit hebben we opgenomen in de methoden op pagina 11 (regel 18-20). Ook bespreken we dit in de discussie op pagina 28 (vanaf regel 21).

- Is het niet uitzonderlijk om vast te stellen dat er geen geslachtsverschillen waren inzake internaliserende en externaliserende gedragsproblemen (p. 12)? Hoe verhoudt deze bevinding zich tot eerdere literatuur of wat kan hiervoor een verklaring zijn? Speelt hierbij een selectie-effect waarbij alle deelnemers aan het onderzoek (jongens en meisjes) al verhoogde probleemscores vertoonden?
 - Hartelijk dank voor uw kritische vragen en suggesties omtrent de vergelijkbare effectiviteit van STOP4-7 voor jongens en meisjes. We hebben deze bevinding duidelijker toegelicht in de discussie op pagina 27 (regel 4-18).

- Hoe sterk was het verband tussen ouder-gerapporteerde en door de leerkracht gerapporteerde probleemgedragingen? Deze correlaties zouden eventueel op de diagonaal van Tabel 2 kunnen toegevoegd worden.
 - Bedankt voor deze waardevolle suggestie. De correlaties tussen leerkracht- en ouder rapportage zijn toegevoegd aan Tabel 2 (pagina 40). De correlaties worden in de resultaten sectie besproken op pagina 17 (vanaf regel 19), evenals in de discussie op pagina 26 (vanaf regel 18).
- Gelieve voor elk van de repeated measures ANOVAs de effectgrootte weer te geven van de gevonden effecten (bijvoorbeeld eta squared) en de effectgrootte ook expliciet te bespreken. Dit gebeurt nu kort in de Discussie maar kan best ook al in de Resultaten gebeuren.
 - De partial eta squared voor de gevonden effecten zijn in de resultaten sectie toegevoegd.
- Het is niet nodig om de namen van de gebruikte vragenlijsten te vermelden in de Abstract.
 - Dank voor deze suggestie. We hebben de namen van de gebruikte vragenlijsten verwijderd uit de Abstract.
- De term ‘disruptief gedrag’ in de eerste zin zou ik vervangen door ‘storend gedrag’.
 - Hartelijk dank voor het aandragen van dit punt, waar ook Reviewer #2 feedback op gaf. We hebben ervoor gekozen de termen “gedragsproblemen” en “ongewenst gedrag” te gebruiken in de gehele manuscript. We hopen dat het hierdoor duidelijk is welke constructen in onze paper worden bekeken, zonder teveel verschillende termen door elkaar te gebruiken.
- Zou de moeilijke term ‘coërcieve cyclus’ ook kunnen vervangen worden door ‘dwingende cyclus’?
 - Hartelijk dank voor uw suggestie. We hebben de term “coërcieve cyclus” vervangen voor “dwingende interactiepatronen”. Wel hebben we op pagina 4 (regel 4) de term “*coercive process*” laten staan. Dit is een bekende vakterm en zo hopen wij duidelijk te maken dat hier wordt verwezen naar de wederzijds bekrachtigende, dwingende interactiepatronen zoals beschreven door Patterson.
- Op p. 3, in de paragraaf die begint met ‘Ondanks het multifactorieel bepaalde karakter ...’, komen verschillende zinnen voor met sterke uitspraken, die niet worden ondersteund door een referentie.
 - Dank dat u en Reviewer #2 ons op deze tekortkoming hebben gewezen. Het mogelijk multifactoriële karakter van disruptief gedrag was niet de hoofdfocus van het huidige onderzoek. Om verwarring te voorkomen, is in deze zin op pagina 5 (regel 1) aangepast.
- Op p. 7 begreep ik niet waarom de lange-termijn effecten van het STOP4-7 programma slechts op een exploratieve manier worden onderzocht? De verwachting is toch dat de veranderingen kort na het programma behouden blijven op de langere termijn?

Bovendien toonden een paar eerdere studies dit al aan, dus waarom niet gewoon de gerichte hypothese vooropstellen dat er follow-up effecten zullen zijn?

- Hartelijk dank voor dit punt. We zijn het met u en Reviewer #2 eens dat op basis van de literatuur gerichte hypothesen gesteld kunnen worden. We hebben dit aangepast op pagina 10 (regel 7-13).
- Op p. 11 is het enigszins verwarrend dat er ‘per kind één informant werd meegenomen in de analyses’. Betreft deze uitspraak enkel de moeders en vaders als informanten of ook de leerkrachten? Hierbij aansluitend: kunnen de auteurs ook preciseren van hoeveel moeders en hoeveel vaders uiteindelijk data werden gebruikt in de analyses? En bij hoeveel kinderen precies waren er ook leerkrachtdata beschikbaar?
 - Dank dat u ons heeft gewezen op deze onduidelijkheid. We hebben dit duidelijker verwoord op pagina 13 (vanaf regel 17) en pagina 14 (tot regel 4). Tevens hebben we bij alle uitkomstmaten aangegeven hoeveel data beschikbaar was en hebben we in Figuur 1 aangegeven hoeveel data beschikbaar was op alle meetmomenten.
- Gelieve op p. 13 en 14 meer betekenisvolle titels te gebruiken, die verwijzen naar ouder- en leerkrachtrapportage, respectievelijk. De afkortingen CBCL en TRF zullen voor sommige lezers niet betekenisvol zijn.
 - Hartelijk dank voor deze suggestie. We hebben dit aangepast.

Reviewer #2

Abstract

- De zin ‘hiervoor gebruikten wij...’ dient herschreven te worden. Misschien goed om het design van de interventiestudie te benoemen en vervolgens aan te geven welke data werden verzameld.
 - Dank u wel voor uw kritische blik. We hebben de Abstract aangepast en het design van het huidige onderzoek duidelijker benoemd (pagina 2, regel 5-8).

Inleiding

- De auteurs starten het manuscript met een definitie van ‘disruptief gedrag’, maar spreken zowel in de titel als verder in het manuscript over gedragsproblemen. Ik vraag me af of dit volledig hetzelfde is. Disruptief gedrag wordt volgens mij voornamelijk genoemd in het kader van gedragsstoornissen en minder met betrekking tot gedragsproblemen. Dit alles hangt natuurlijk erg nauw samen, maar het lijkt me goed hier nog een keer kritisch naar te kijken zodat er zeker naar de juiste terminologie en definitie wordt verwezen en dit ook door te trekken doorheen het manuscript.
 - Hartelijk dank voor uw opmerking. We zijn het met u en Reviewer #1 eens dat het van belang is de juiste terminologie consistent toe te passen. We hebben ervoor gekozen in het gehele manuscript gebruik te maken van de termen “gedragsproblemen” en “ongewenst gedrag”.
- Bij de zinnen die beginnen met ‘zij gaan in sociale situaties weinig...’ moet nog een referentie toegevoegd worden.
 - Hartelijk dank voor uw kritische blik. We hebben op pagina 4 (regels 14-22) duidelijker verwezen naar de gebruikte literatuur.

- Er wordt gesproken over het multifactorieel karakter van disruptief gedrag. Echter, uit het voorgaande gaat men enkel in op het coërcieve interactiepatroon. Misschien goed om dit bij het begin wat verder open te trekken en kort in te gaan op het multifactorieel karakter zodat er hier later op kan teruggegaan worden.
 - Dank u wel voor dit belangrijke punt. Ook Reviewer #1 kaartte dit punt aan. We hebben ervoor gekozen de tekst op pagina 5 (regel 1) te herschrijven, waarbij we het woord “multifactorieel” hebben weggelaten. Het multifactoriële karakter van disruptief gedrag was niet de focus van het huidige onderzoek en zodoende is het niet relevant om dit in de inleiding te bespreken.

- Er wordt verwezen naar de effectiviteit van bepaalde oudertrainingen. Echter, de referenties staan enkel samen op het einde van deze alinea. Het zou goed zijn moest er hier meer precies gerefereerd kunnen worden. Welke conclusies worden getrokken o.b.v. welke onderzoeken. Rond multimodale programma's moet zeker nog een referentie toegevoegd worden.
 - Hartelijk dank voor uw kritische blik. We hebben de alinea op pagina 5 (regel 1-9) herschreven zodat duidelijk is de beschreven bevindingen gebaseerd zijn op de publicaties van Kazdin & Weisz (1998), Lundahl en collega's (2006) en Webster-Stratton & Taylor (2001). De stelling over de effectiviteit van multimodale programma's is gebaseerd op het onderzoek van Webster-Stratton & Taylor (2001). Dit hebben we duidelijker aangegeven in de tekst.

- Fout in zin: Uit eerder onderzoek naar STOP4-7 onder Vlaamse 201 kinderen tussen 4 en 8 jaar met klinische gedragsproblemen bleek dat het externaliserend gedrag zowel thuis als op school was verbeterd na deelname aan STOP4-7.
 - Deze zin is aangepast.
- In de alinea die start met 'Een recente meta-analyse...' beschrijven de auteurs verschillende mogelijke aspecten die onderzocht kunnen worden. Het is echter niet helemaal duidelijk hoe dit linkt aan het voorgestelde onderzoek en hoe dit meegenomen wordt. Op basis van de volgende alinea die start met 'een andere belangrijke vraag' ga ik ervan uit dat in voorgaande alinea ook verschillende vragen beschreven worden. Deze kunnen volgens mij meer expliciet beschreven worden.
 - Hartelijk dank voor deze suggestie.

- Een review-van-reviews kan volgens mij omschreven worden als een meta-review.
 - Dank u wel. We hebben dit aangepast op pagina 9 (regel 11).

- Ik vraag me af waarom bepaalde vragen exploratief bekeken worden. Hier kunnen op basis van eerdere onderzoeken (m.b.t. andere programma's), waar de onderzoekers ook naar verwijzers, toch ook hypothesen worden geformuleerd.
 - Hartelijk dank voor uw feedback. Ook Reviewer #1 stelde zich deze vraag. In eerdere onderzoeken werden lange termijn effecten van programma's vergelijkbaar met STOP4-7 gevonden, en zodoende hebben we ervoor gekozen de hypothese betreffende lange termijn effecten niet als exploratief te presenteren. Dit geldt ook voor de hypothese betreffende de subgroep gezinnen die bij aanvang van STOP4-7 ernstigere problemen rapporteerden. Voor de hypothesen met

betrekking tot leeftijd en geslacht van kinderen geldt dat er minder richtinggevend bewijs beschikbaar is. We hebben de verwoording op pagina 10 (regel 10-13) echter aangepast, zodat het duidelijk is dat het hier niet om een exploratieve onderzoeksvraag gaat. De mogelijke klinische relevantie van de afnames in gerapporteerd probleemgedrag onderzoeken we wel exploratief, omdat hierover weinig eerdere literatuur beschikbaar is.

Methodes

- Zin 1 van design, het woord ‘gecombineerd’ schrappen.
 - Deze zin is aangepast.

- De ‘gemiddelde’ momenten waarop de voor-, nameting en follow up plaatsvond wordt beschreven. Het lijkt me zinvol hier de range aan toe te voegen waarbinnen op elk moment de data werden verzameld.
 - Dank u wel deze suggestie. We hebben de verwoording aangepast, zodat duidelijker is op welke momenten de vragenlijsten werden ingevuld door ouders en leerkrachten (pagina 10, regel 20-23).

- In de voorlaatste zin van het design staat een keer ‘plaats’ te veel.
 - Dank u wel. We hebben dit aangepast.

- APA: cijfers ten en met tien moeten voluit geschreven worden.
 - Dank u wel. We hebben dit aangepast.

- Misschien goed om nog toe te lichten waarom van de Opvoedbelasting Vragenlijst enkel de drie eerste subschalen werden gebruikt.
 - Hartelijk dank voor deze suggestie. We hebben dit uitgelegd op pagina 14 (vanaf regel 22).

- Er wordt weinig informatie gegeven over de respondentengroep. Over de kinderen is de informatie beperkt, maar is als ik het goed heb niet meer informatie beschikbaar. Over de ouders en leerkrachten wordt geen informatie gegeven. Geslacht, leeftijd, aantal kinderen, aantal jaren ervaring,... Zij deze gegevens beschikbaar. Het lijkt me interessant deze toe te voegen.
 - We zijn het met u eens dat het wenselijk is zoveel mogelijk informatie over de respondentengroep te beschrijven. Helaas was er voor het huidige onderzoek geen verdere demografische gegevens van leerkrachten of ouders beschikbaar. Dit hebben we toegevoegd op pagina 11 (regel 16-17).

- Daarnaast vraag ik me af of er voor elk meetmoment van alle informatie voor de 770 kinderen beschikbaar was. Ook deze info dient volgens mij toegevoegd te worden.
 - We hebben bij alle uitkomstmaten aangegeven welke data beschikbaar was en in Figuur 1 weergegeven hoeveel informatie er per uitkomstmaat beschikbaar was op alle meetmomenten.

Resultaten

- Het lijkt me goed om de verwoording in de resultaten nog een keer goed na te kijken. Bijvoorbeeld p14. ‘kinderen die op baseline meer probleemgedrag hadden dan hun

leeftijdgenoten, hadden dit op de nameting en follow-up doorgaans ook.’ Moet volgens mij omschreven worden als: ‘kinderen waarvoor ouders tijdens de baselinemeting meer probleemgedrag rapporteerden dan hun leeftijdsgenoten..’

- Hartelijk dank voor deze suggestie. We hebben de resultaten sectie nauwkeurig nagelopen en de verwoording aangepast.
- P. 15 spreken de auteurs meermaals over ‘de kinderen veranderden niet’. Ik begrijp niet helemaal wat hiermee wordt bedoeld. Bedoelen de auteurs hiermee veranderen van klinisch naar niet-klinisch niveau? Ik denk dat hier wat toelichting of een andere beschrijving nodig is.
 - Hartelijk dank voor het aankaarten van deze onduidelijkheid. Het ging hier niet om een verandering van klinisch naar niet-klinisch niveau, maar om de gemiddelde afname van externaliserend en internaliserend gedrag zoals gerapporteerd door leerkrachten bij kinderen die volgens leerkracht ten tijde van de voormeting in de klinische range van gedragsproblemen scoorden. We hebben de verwoording van de resultaten sectie (en specifiek op pagina 17) aangepast zodat dit duidelijk is.

Discussie

- P. 22 zin herbekijken: ‘Dit onderzoek geeft al eerste onderzoek in Nederland...’
 - Dank u wel. We hebben dit aangepast.

Algemeen

- In het hele document moeten de referenties nagekeken worden. APA 7: vanaf 3 auteurs in de tekst enkel de eerste noemen.
 - Dank u wel. We hebben dit aangepast.

Tweede versie van manuscript – 21-12-2021

Evaluatie van Samen sterker Terug Op Pad (STOP4-7):

Een Multimodale Interventie voor Kinderen met Gedragsproblemen

<geanonimiseerd>

Abstract

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 jeugdhulp instellingen. 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. Deze 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.

Keywords: STOP4-7; interventie; externaliserend; internaliserend; opvoedstress

Evaluatie van Samen sterker Terug Op Pad (STOP4-7):

Een Multimodale Interventie voor Kinderen met Gedragsproblemen

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 lange termijn effecten 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 leerkracht rapportage om veranderingen in probleemgedrag van kinderen in kaart te brengen. Vernieuwend is bovendien dat wordt gekeken naar de klinische relevantie van waargenomen veranderingen en naar *voor wie* dit programma beter of minder goed werkt.

Dwingende Ouder-Kind Interactiepatronen en de Ontwikkeling van Probleemgedrag

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 standhoudend 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 standhoudende 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 toefelijke gedrag van hun ouders door te kalmeren nadat zij hun zin hebben gekregen. Ook een gebrek aan warmte in de ouder-kind relatie 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 het meest effectieve 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 (zowel ouders, kind en leerkracht) effectiever zijn dan wanneer slechts een schakel uit het systeem wordt aangepakt. Multimodale programma's, waarbij naast ouders ook het kind en de leerkracht betrokken worden, blijken ook effectiever te zijn dan programma's die zich uitsluitend richten op ofwel ouders, kinderen, of leerkrachten (Webster-Stratton & Taylor, 2001).

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 4- tot 7-jarige 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 4 en 8 jaar met klinische gedragsproblemen bleek dat het externaliserend 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 wachtlijst controlegroep (De Mey, 2010). In Nederland werden soortgelijke resultaten voor STOP4-7 gevonden. Uit onderzoek (Van Brussel et al., 2009) onder 69 kinderen tussen de 4 en 8 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 sociaal 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 sociaal 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 kind centra in Nederland) liet zien dat de aan STOP4-7 verbonden afnames in probleemgedrag stabiel bleven tot aan een 6 maanden follow-up. Daarnaast bleek dat gezinnen konden worden ingedeeld in drie subgroepen, die vooral verschilden in de mate van de ernst van de problematiek bij aanvang van de interventie.

Lange Termijn Effecten, Comorbiditeit, en Klinische Relevantie

De eerdere bevindingen impliceren dat STOP4-7 mogelijk leidt tot 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 (over alle jeugdhulp locaties heen die STOP4-7 aanbieden) middels een multi-informant benadering gekeken naar de korte en lange termijn effecten 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 langere termijn effecten sorteren, 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 op zijn beurt 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 en 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 interventie studies 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 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 niveaus van risico. Zo liet een recente *individual participant data* meta-analyse (op basis van gepoolde datasets) zien dat 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 multi-modale 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).

Onderzoeksvragen en Hypothesen

Op basis van eerder onderzoek verwachtten wij dat de multimodale interventie STOP4-7 zou leiden 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 behouden blijven tot aan de 6-maanden follow-up. Verder verwachten wij dat STOP4-7 mogelijk differentiële —relatief sterke of juist minder sterke— effecten heeft voor subgroepen 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 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.

Methode

Design

Om het effect van STOP4-7 te analyseren, werden twee bestaande datasets samengevoegd. Deze datasets bevatte data die in Nederland en in Vlaanderen werd verzameld in het kader van *routine outcome monitoring*. Data was beschikbaar voor 3 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.

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 toestemmingsprocedure gebruikt, goedgekeurd door het Ethisch Comité van de Universiteit Gent. Het merendeel van de kinderen was jongen ($n = 683$, 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 over gemiddeld genomen slechts één op tien tot zestien gezinnen vroegtijdig stopt met STOP4-7 en het programma niet afmaakt (de Mey & Geertse, persoonlijke communicatie).

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 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, een 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.

Meetinstrumenten

Gedragsproblemen – ouder rapportage. 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 drie-puntschaal (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: $\alpha = .84$; Internaliserend probleemgedrag: $\alpha = .87$; Externaliserend probleemgedrag: $\alpha = .80$).

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 moeder gerapporteerd en voor $n = 84$ kinderen werd door 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.

Gedragsproblemen – leerkracht rapportage. 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 (35 items) en Externaliserend probleemgedrag (34 items) meegenomen. Items worden beantwoord op een drie-puntschaal (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: $\alpha = .76$; Internaliserende probleemgedrag: $\alpha = .74$; Externaliserend probleemgedrag: $\alpha = .77$). 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.

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 totaal score 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-kind relatie. Een voorbeeld item van de OBVL is “Ik voel me gelukkig met mijn kind”. Items worden beantwoord op een 4-puntschaal (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: $\alpha = .79$; Problemen in de opvoeding,: $\alpha = .79$; Depressieve stemmingen: $\alpha = .75$; Totale opvoedstress: $\alpha = .77$).

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 voor $n = 148$ ouders rapporteerden over het probleemgedrag van hun kind, maar niet over hun ervaren opvoedstress.

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 gerapporteerd probleemgedrag (i.e., TRF). Ook werd 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 significant afname van probleemgedrag of opvoedstress van voormeting naar follow-up) voerden we *paired sample t-tests* uit om vast te stellen in welke tijdsperiode (i.e., van voor- naar nameting, of van nameting naar follow-up) de significante verandering plaatsvond. De effect

groottes werd bekeken aan de hand van de *partial eta squared* (η_p^2) waarbij $\eta_p^2 = .01$ als klein een klein effect wordt geïnterpreteerd, $\eta_p^2 = .06$ als een middelgroot effect en $\eta_p^2 = .14$ als een groot effect (Field, 2013).

Daarna werd bekeken in hoeverre de resultaten afhankelijk waren van de leeftijd van kinderen, 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-kwadraat toetsen 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-kwadraat toetsen bekeken. Een T-score ≥ 63 werd beschouwd als indicatief voor klinische niveaus van opvoedstress (Vermulst et al., 2015).

Resultaten

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 ($t[629] = .286, p = .775$). Wel rapporteerden ouders minder gedragsproblemen voor jongere kinderen dan voor oudere kinderen (≥ 6 jaar; $t[214] = -2.728, p < .01$). Leerkrachten rapporteerden evenveel gedragsproblemen voor jongens als voor meisjes ($t[580] = .707, p = .480$), en voor jongere en oudere kinderen ($t[175] = .881, p = .379$) voor aanvang van STOP4-7. Er waren geen verschillen

voor aanvang van STOP4-7 in ervaren opvoedbelasting voor ouders van jongens en meisjes ($t[531] = .136, p = .892$), of ouders van jongere en oudere kinderen ($t[171] = 1.298, p = .196$).

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 door ouder gerapporteerd internaliserend probleemgedrag van kinderen waren matig tot sterk positief (r s van .648 tot .682, $p < .001$). Hetzelfde gold voor de autocorrelaties voor door ouders gerapporteerd externaliserend probleemgedrag (r s van .488 tot .652, $p < .001$). Een vergelijkbaar patroon was zichtbaar voor het door leerkrachten gerapporteerd internaliserend en externaliserend probleemgedrag van kinderen (r s van .454 tot .534, $p < .001$, en r s van .454 tot .653, $p < .001$, 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 .604 tot .642, $p < .001$). De correlatie tussen internaliserend en externaliserend probleemgedrag zoals gerapporteerd door leerkrachten was eveneens matig sterk en positief (r s van .454 tot .483, $p < .001$). 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 ($r = .253, p < .05$). De correlatie tussen door ouders en leerkracht gerapporteerd internaliserend probleemgedrag op de voormeting en nameting was niet

significant (r_s .073 en .141, $p > .05$). Ook de correlatie tussen door ouders en leerkrachten gerapporteerd externaliserend probleemgedrag was niet significant (r_s van -.003 tot .142, $p > .05$).

Veranderingen in door Ouders Gerapporteerd Probleemgedrag

Ouders rapporteerden een grote, significante afname van totaal probleemgedrag ($\eta_p^2 = .31$), internaliserend probleemgedrag ($\eta_p^2 = .20$) en externaliserend probleemgedrag ($\eta_p^2 = .38$) van kinderen van voormeting tot follow-up (Tabel 3). Het door ouders gerapporteerde totale probleemgedrag van kinderen nam significant af van voormeting naar nameting ($M_{T0-T1} = 5.736$, $SD = 7.702$, $t[268] = 12.214$, $p < .001$) en stabiliseerde van nameting naar follow-up ($M_{T1-T2} = -.273$, $SD = 8.842$, $t[142] = -.369$, $p = .731$). Wanneer specifieker naar het type probleemgedrag werd gekeken, was een vergelijkbaar patroon zichtbaar. Het door ouders gerapporteerde internaliserend probleemgedrag van kinderen nam significant af van voormeting naar nameting ($M_{T0-T1} = 4.145$, $SD = 8.753$, $t[268] = 7.67$, $p < .001$), en bleef stabiel van nameting naar follow-up ($M_{T1-T2} = .195$, $SD = 9.066$, $t[142] = .258$, $p = .797$). Dit was eveneens het geval voor het door ouders gerapporteerde externaliserend probleemgedrag van kinderen ($M_{T0-T1} = 6.907$, $SD = 8.001$, $t[268] = 14.158$, $p < .001$; $M_{T1-T2} = -.440$, $SD = 9.377$, $t[142] = -.562$, $p = .575$).

De door ouders gerapporteerde veranderingen op zowel het totale probleemgedrag als internaliserend probleemgedrag als het externaliserend 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 jeugdhulp locaties.

Veranderingen in door Leerkrachten Gerapporteerd Probleemgedrag

Leerkrachten rapporteerden een kleine, significante afname van het totale probleemgedrag ($\eta_p^2 = .03$) en externaliserend probleemgedrag ($\eta_p^2 = .05$) van kinderen van voormeting naar follow-up, maar rapporteerden geen significante afname van het internaliserend gedrag (zie Tabel 3). Het totale probleemgedrag van kinderen nam volgens leerkrachten af van voormeting naar nameting ($M_{T0-T1} = 2.259$, $SD = 7.285$, $t[242] = 4.834$, $p < .001$), en stabiliseerde van nameting naar follow-up ($M_{T1-T2} = -1.071$, $SD = 9.031$, $t[112] = -1.260$, $p = .210$). Ook externaliserend gedrag van kinderen nam volgens leerkrachten significant af van voormeting naar nameting ($M_{T0-T1} = 2.613$, $SD = 7.512$, $t[242] = 5.422$, $p < .001$) en stabiliseerde van nameting naar follow-up ($M_{T1-T2} = -1.212$, $SD = 9.104$, $t[112] = -1.416$, $p = .160$). Voor door leerkrachten gerapporteerd internaliserend gedrag was de verandering van voormeting naar nameting ($M_{T0-T1} = .897$, $SD = 8.927$, $t[242] = 1.567$, $p = .119$) evenals de verandering van nameting naar follow-up ($M_{T1-T2} = -.823$, $SD = 9.645$, $t[112] = -.907$, $p = .366$) niet significant.

Zowel de door leerkrachten gerapporteerde veranderingen van het totale probleemgedrag van kinderen als het externaliserend 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 Vlaamse kinderen als voor 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 totaal probleemgedrag van voormeting naar nameting ($M_{T0-T1} = -.443$, $SD = 7.250$, $t[96] = -.602$, $p = .548$), terwijl dit voor kinderen die op voorhand in de klinische range scoorden wel het geval was ($M_{T0-T1} = 4.055$, $SD = 6.756$, $t[145] = 7.252$, $p < .001$). 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 ($M_{T1-T2} = -.872$, $SD = 7.237$, $t[38] = -.752$, $p = .457$). 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 ($M_{T1-T2} = .367$, $SD = 9.984$, $t[48] = -.339$, $p = .736$).

Het door leerkrachten gerapporteerde externaliserend 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 ($M_{T0-T1} = .062$, $SD = 7.343$, $t[96] = .083$, $p = .934$ en $M_{T0-T1} = 4.308$, $SD = 7.158$, $t[145] = 7.273$, $p < .001$, respectievelijk). Van nameting tot follow-up nam het door leerkrachten gerapporteerde externaliserend gedrag van kinderen die volgens leerkrachten op voorhand in de niet-klinische range scoorden toe ($M_{T1-T2} = -2.359$, $SD = 7.02$, $t[38] = -2.074$, $p < .05$), terwijl dit stabiel bleef voor kinderen die volgens leerkrachten op voorhand in de klinische range van probleemgedrag scoorden ($M_{T1-T2} = .857$, $SD = 9.923$, $t[48] = .605$, $p = .548$).

Het door leerkrachten gerapporteerde internaliserend 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 ($M_{T0-T1} = -2.103$, $SD = 9.247$, $t[96] = -2.240$, $p < .05$ en $M_{T0-T1} = 2.890$, $SD = 8.147$, $t[145] = 4.223$, $p < .001$, respectievelijk). Voor beide groepen was de door leerkrachten gerapporteerde verandering van internaliserend probleemgedrag van nameting tot follow-up niet-significant ($M_{T1-T2} = 0.026$, $SD = 8.552$, $t[38] = .019$, $p = .985$ en $M_{T1-T2} = .286$, $SD = 10.766$, $t[48] = .186$, $p = .853$). De verandering op het door leerkrachten gerapporteerde internaliserend probleemgedrag van kinderen bleek ook afhankelijk van het geslacht van kinderen. Voor meisjes was de door leerkrachten gerapporteerde afname van voormeting naar nameting significant ($M_{T0-T1} = 3.309$, $SD = 10.189$, $t[41] = 2.105$, $p < .05$), maar voor jongens niet ($M_{T0-T1} = .393$, $SD = 8.583$, $t[200] = .649$, $p = .517$). Voor zowel meisjes als jongens was de door leerkrachten gerapporteerde verandering in internaliserend probleemgedrag van nameting naar follow-up niet significant ($M_{T1-T2} = -.312$, $SD = 15.326$, $t[15] = -.082$, $p = .936$ en $M_{T1-T2} = -.842$, $SD = 8.548$, $t[94] = -.960$, $p = .339$, respectievelijk). De door leerkrachten gerapporteerde verandering in internaliserend probleemgedrag van kinderen was niet afhankelijk van leeftijd of de locatie van de interventie.

Veranderingen in Opvoedstress (OBVL)

De door ouders gerapporteerde totale opvoedstress nam significant af van voormeting tot follow-up ($\eta_p^2 = .30$; zie Tabel 4). Dit is een groot effect. Deze afname was zichtbaar van voormeting naar nameting ($M_{T0-T1} = 5.294$, $SD = 8.288$, $t[281] = 10.727$, $p < .001$) en stabiliseerde van nameting naar follow-up ($M_{T1-T2} = .271$, $SD = 8.689$, $t[69] = .261$, $p = .795$). Wanneer naar de subschalen van de OBVL werd gekeken, was een vergelijkbaar patroon zichtbaar. Gerapporteerde opvoedstress door problemen in de ouder-kind relatie (OKR; $\eta_p^2 = .21$), problemen in de opvoeding (PO; $\eta_p^2 = .43$) en depressieve stemmingen (DEP; $\eta_p^2 = .16$) namen significant af van voormeting naar follow-up. Ook hierbij was in alle gevallen sprake van

een groot effect. De afname van gerapporteerde opvoedbelasting was steeds significant van voormeting naar nameting (OKR: $M_{T0-T1} = 3.921$, $SD = 8.061$, $t[281] = 8.170$, $p < .001$; PO: $M_{T0-T1} = 6.237$, $SD = 8.261$, $t[281] = 12.679$, $p < .001$; DEP: $M_{T0-T1} = 2.868$, $SD = 6.857$, $t[280] = 7.012$, $p < .001$) en stabiliseerde van nameting naar follow-up (OKR: $M_{T1-T2} = -1.000$, $SD = 7.940$, $t[69] = -1.054$, $p = .296$; PO: $M_{T1-T2} = .754$, $SD = 7.558$, $t[68] = .828$, $p = .410$; DEP: $M_{T1-T2} = -3.000$, $SD = 7.913$, $t[69] = -.317$, $p = .752$).

De gerapporteerde veranderingen van problemen in de opvoeding, depressieve stemmingen en totale opvoedstress was niet afhankelijk van het geslacht van het kind, de 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 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-kind relatie significant af van voormeting naar nameting ($M_{T0-T1} = 5.289$, $SD = 7.902$, $t[189] = 9.226$, $p < .001$), maar dit was niet het geval wanneer ouders op voorhand in de niet-klinische range van opvoedstress scoorden ($M_{T0-T1} = 1.098$, $SD = 7.680$, $t[91] = 1.371$, $p = .174$). Voor beide groepen was de verandering van nameting tot follow-up niet-significant ($M_{T0-T1} = -2.250$, $SD = 10.199$, $t[11] = -.764$, $p = .461$ en $M_{T0-T1} = -1.109$, $SD = 7.596$, $t[51] = -.968$, $p = .338$, respectievelijk).

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 ($\chi^2 [1, N = 269] = 59.066, p < .001$), en tussen voormeting en follow-up ($\chi^2 [1, N = 147] = 18.654, p < .001$). 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 door ouders gerapporteerd internaliserend en externaliserend 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 ($\chi^2 [1, N = 269] = 63.678, p < .001$). 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 ($n = 68$; $\chi^2 [1, N = 147] = 42.596, p < .001$). 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 ($\chi^2 [1, N = 269] = 47.002, p < .001$). 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 ($n = 115$; $\chi^2 [1, N = 147] = 18.046, p < .001$).

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 ($\chi^2 [1, N = 282] = 64.742, p < .001$). 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 ($n = 61; \chi^2 [1, N = 79] = 3.682, p = .055$).

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 ($\eta_p^2 = .15-.43$). 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 $d = .27$; Menting et al., 2013) tot middelgroot ($d = .42$; 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 gerag (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 externaliserend 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 ouder gerapporteerd en leerkracht gerapporteerd probleemgedrag (i.e., de CBCL en TRF) toonde aan dat de overeenstemming tussen ouder- en leerkrachtrapportage van probleemgedrag matig is ($r = .29$), en dat ouders vaak meer gedragsproblemen rapporteren dan leerkrachten (Rescorla et al., 2014). Ook in dit onderzoek werd een dergelijk patroon gevonden. De correlaties tussen ouder en leerkracht rapportages 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 gerapporteerd probleemgedrag klein waren ($\eta_p^2 = .03-.05$), 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 ouderrapportage van probleemgedrag geen verschil blijkt in de gedragsveranderingen na STOP4-7 voor jongens of meisjes. Deze bevinding komt overeen met onderzoek naar een vergelijkbaar multi-modaal 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 multi-modale 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 sekse verschillen in de effecten van (multi-modale) interventies voor probleemgedrag.

Tenslotte vonden we dat de problemen in de ouder-kind relatie 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 wachtlijstconstructie, 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 mogelijksterwijs 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,

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.

Referenties

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., & Kunder, 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/10.1177/106342660000800304>
- 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, . (2019). Parent training: Effects beyond conduct problems. *Children and Youth Services Review, 100*, 405-414.
<https://doi.org/10.1016/j.chidyouth.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., ... & 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%2F10634266020100030601>

- 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/s11121-020-01176-6>
- 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., ... & 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.

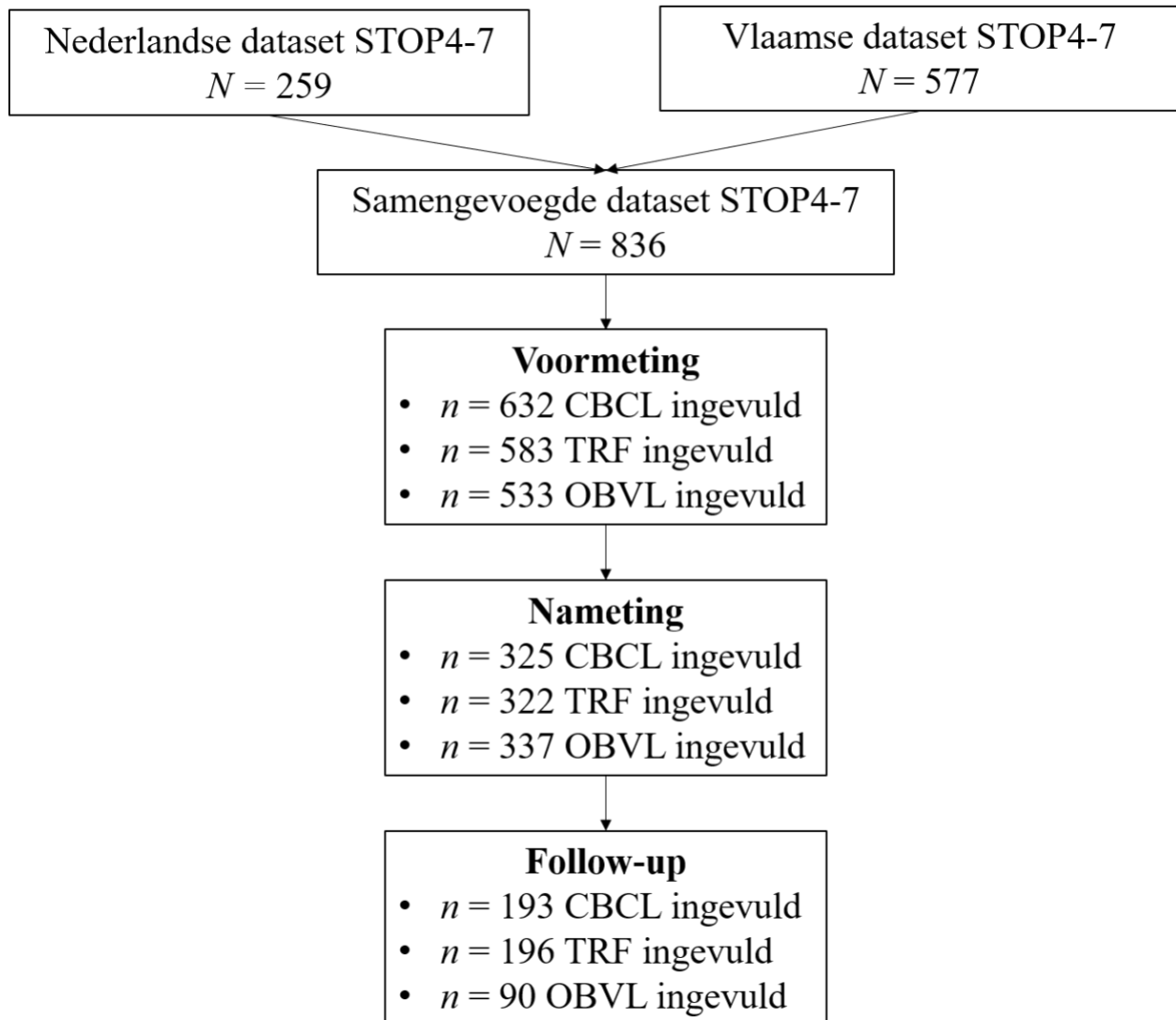
<https://doi.org/10.1111/j.1469-7610.2011.02373>

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

Figuur 1. Flowchart van beschikbare data op de drie meetmomenten.



Tabel 1.

Gemiddelden en Standaarddeviaties: Ouder- en Leerkracht Gerapporteerde Gedragsproblemen en Opvoedstress van Voormeting naar Follow-up.

	Voormeting			Nameting			<i>n</i>	<i>M</i>
	<i>n</i>	<i>M</i>	<i>SD</i>	<i>n</i>	<i>M</i>	<i>SD</i>		
Gedragsproblemen- Ouder rapportage								
Totaal		63.73	9.68		58.21	10.58		
Internaliserend	632	59.24	10.31	325	54.94	10.77	193	
Externaliserend		67.03	9.87		60.46	10.02		
Gedragsproblemen- Leerkracht rapportage								
Totaal		64.47	8.94		61.44	8.36		
Internaliserend	583	57.71	9.43	322	55.52	8.36	196	
Externaliserend		66.73	9.55		63.43	8.33		
Opvoedstress								
Totaal		65.15	10.15	337	60.28	10.51	90	
Ouder-kind relatie		63.94	9.93	337	60.71	9.59	90	
Problemen in opvoeding	533	66.18	9.70	337	60.96	9.53	89	
Depressieve stemming		61.80	9.04	336	59.19	8.83	90	

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.
1. INT-T0	.073	.457***	.454***	.454***	.291***	.203**	.704**	.3
2. INT-T1	.648**	.141	.534***	.225***	.447***	.143	.329**	.6
3. INT-T2	.681**	.682**	.253*	.199*	.178	.483**	.337**	.3
4. EXT-T0	.604**	.380**	.425**	.053	.653**	.529**	.895**	.5
5. EXT-T1	.391**	.625**	.421**	.652**	-.003	.454**	.650**	.9
6. EXT-T2	.349**	.382**	.642**	.488**	.550**	.142	.485**	.4
7. TOT-T0	.834**	.572**	.582**	.870**	.627**	.460**	-.016	.6
8. TOT-T1	.552**	.830**	.566**	.581**	.901**	.514**	.696**	.0
9. TOT-T2	.521***	.568**	.832**	.512**	.569**	.901**	.590**	.6

Noot. Correlaties voor CBCL onder de diagonaal, voor TRF boven de diagonaal; Licht grijs gekleurde vakken geven de correlatie tussen subschalen weer. Donker grijs gekleurde vakken op de diagonaal geven de correlatie tussen ouder- en leerkrachtrapportage weer. T0 = voormeting, T1 = nameting, T2 = follow-up meting; * $p < .05$, ** $p < .01$.

Tabel 3

Uitkomsten Repeated Measures ANOVA: Van Voormeting (T0) 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.

Tabel 4.

Uitkomsten Repeated Measures ANOVA: Opvoedstress (OBVL) van Voormeting naar Follow-Up.

	Totale Opvoedstress			Problemen in Ouder-Kind Relatie			Problemen in Opvoeding			<i>F</i>
	<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	11.1
* Geslacht kind	.008	.992	.000	.570	.567	.009	.307	.736	.005	.57
* Leeftijd kind	.172	.843	.004	.024	.976	.001	1.206	.304	.028	.20
* Baseline opvoedstress	2.511	.085	.039	6.183	.003	.091	.883	.416	.014	2.1
* Implementatie locatie	1.107	.334	.018	1.098	.337	.017	.881	.417	.014	.19

Noot. Dikgedrukte resultaten zijn statistisch significant.

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 Ouder Gerapporteerd Probleemgedrag en Opvoedstress.

	Voormeting			Nameting			N
	N	Niet-klinisch n (%)	Klinisch n (%)	N	Niet-klinisch n (%)	Klinisch n (%)	
Totaal probleemgedrag		268 (42.4)	364 (57.6)		210 (64.6)	115 (35.4)	
Internaliserend gedrag	632	374 (59.3)	258 (40.7)	325	238 (73.8)	87 (26.2)	19
Externaliserend gedrag		185 (29.3)	447 (70.7)		193 (59.4)	132 (40.6)	
Opvoedstress	533	182 (34.1)	351 (65.9)	337	176 (52.2)	161 (47.8)	90

Acceptatiebrief – 24-1-2022

Met veel plezier laten we weten dat we het artikel in de huidige versie accepteren!

Bij de reviewers waren geen inhoudelijke vragen, wel had een reviewer een aantal kleine opmerkingen. Zie de bijlage.

Verder schrijven ze:

Reviewer 1

Ik heb met veel plezier de herziene versie van dit artikel gelezen. Het is een erg interessant en relevant artikel geworden en door de aanpassingen die door de auteurs werden gedaan heb ik geen verdere vragen of opmerkingen meer.

Reviewer 2

Gelieve in bijlage mijn laatste (kleine suggesties) te vinden bij deze herwerkte paper. Ik vond het een sterk artikel en de auteurs hebben ook echt werk gemaakt van de revisie.

Bijlage

De auteurs kwamen zeer goed tegemoet aan mijn suggesties. De innovatieve meerwaarde van de studie komt nu meer uit de verf. Mijn complimenten voor deze hoogstaande bijdrage aan de literatuur en aan Jeugd in Ontwikkeling.

Hieronder nog een paar kleinigheden voor de eindredactie:

- P. 5: ‘Verschillende studies lieten zien dat oudertrainingen het meest effectieve middel zijn om dergelijk gedrag bij kinderen te verminderen.’: Meest effectief in vergelijking met welke andere types trainingen?
- P. 5, 6^e lijn: word moet wordt zijn.
- Onderaan p. 6: de volgende zin lijkt wat verloren te staan en mag misschien weg: ‘Daarnaast bleek dat gezinnen konden worden ingedeeld in drie subgroepen, die vooral verschilden in de mate van de ernst van de problematiek bij aanvang van de interventie.’
- Op p. 10 ontbreekt volgens mij het woord ‘wordt’ in de volgende zin: ‘De mogelijke klinische relevantie van afnames in probleemgedrag na STOP4-7 exploratief onderzocht.’
- In de zin ‘Wel is bekend dat over gemiddeld genomen slechts één op tien tot zestien gezinnen vroegtijdig stopt’ mag het woord ‘over’ weg.
- Op p. 15 ontbreekt het woord ‘de’ in de zin ‘Ook werd cross-sectionele relatie tussen ouder- en leerkrachtrapportage van gedragsproblemen bekeken.’