

Peer-reviewverslag ‘Jeugdbescherming op maat: Een pilotonderzoek naar ARIJ-Needs’

Originele manuscript en reviews

07-11-2022	Originele manuscript	Pagina 2
23-11-2022	Reactie reviewer 1	Pagina 25
09-11-2022	Reactie reviewer 2	Pagina 27

Tweede versie

31-09-2023	Reactiebrief van auteur	Pagina 28
31-09-2023	Tweede versie van het manuscript	Pagina 37
04-10-2023	Tweede reactie reviewer 1	Pagina 64
04-10-2023	Tweede reactie reviewer 2	Pagina 65

Publicatie

Acceptatie: 04-10-2023

Online: 06-12-2023

Originele manuscript – 07-11-2022

Jeugdbescherming op maat: Een pilotonderzoek naar [NAAM VAN TOOL GEMASKEERD VOOR REVIEW]

Samenvatting

Recentelijk is [NAAM VAN TOOL GEMASKEERD VOOR REVIEW, VERVANGEN DOOR ‘TOOL GEMASKEERD’ IN MANUSCRIPT] ontwikkeld om professionals te ondersteunen in zowel behoeftetaxatie als besluitvorming over passende zorg voor gezinnen met risico’s op (herhaling van) kindermishandeling. De huidige studie betrof een kwalitatief onderzoek naar de klinische waarde en bruikbaarheid van [TOOL GEMASKEERD]. Hiervoor werden professionals ($N = 15$) geïnterviewd waarbij een fictieve gezinscasus werd voorgelegd om verschillen te onderzoeken tussen ongestructureerde klinische behoeftetaxatie zonder gebruik van [TOOL GEMASKEERD] en gestructureerde behoeftetaxatie met gebruik van [TOOL GEMASKEERD]. Volgens de resultaten signaleren professionals significant meer zorgbehoeften als zij [TOOL GEMASKEERD] gebruiken dan wanneer zij zorgbehoeften taxeren volgens hun klinisch en ongestructureerd oordeel zonder gebruik van [TOOL GEMASKEERD]. Ook blijken professionals meer opvoedings- en ouderfactoren te signaleren bij gebruik van [TOOL GEMASKEERD], wat sterke voorspellers zijn voor kindermishandeling en daarmee belangrijke aanknopingspunten voor behandeling. De zorgkeuzemodule van [TOOL GEMASKEERD] werd beoordeeld als een ‘verbreding van de horizon’ in het voortdurend groeiende aanbod van zorg en interventies in de jeugdbeschermingspraktijk. In conclusie ondersteunen de resultaten in deze studie de klinische waarde en bruikbaarheid van [TOOL GEMASKEERD].

Abstract

The assessment instrument “[TOOL GEMASKEERD]” has recently been developed to support practitioners in assessing treatment needs and selecting appropriate interventions for families at risk for child maltreatment. This qualitative study examined the clinical value and usability of the [TOOL GEMASKEERD] assessment instrument. Practitioners ($N = 15$) were interviewed in which a case vignette was used to examine differences between unstructured clinical needs assessment without [TOOL GEMASKEERD] and structured needs assessment with [TOOL GEMASKEERD]. The results showed that practitioners identify significantly more treatment needs when they use [TOOL GEMASKEERD] than when they assess health

care needs according to their clinical and unstructured judgment. Professionals specifically identified more parent- and parenting factors when they use [TOOL GEMASKEERD], which are strong predictors of child abuse and therefore important treatment targets. The results also showed that [TOOL GEMASKEERD] supports practitioners in selecting interventions from a vastly growing range of child welfare interventions that directly target the assessed factors. In conclusion, this study supports the clinical value and usability of the [TOOL GEMASKEERD] assessment instrument.

Trefwoorden: *jeugdbescherming, [TOOL GEMASKEERD], zorgbehoefte, interventiekeuze*

Aantal woorden: 4910

1. Inleiding

1.1. Aanleiding van het Onderzoek

Kindermishandeling is een omvangrijk probleem met ernstige gevolgen voor zowel slachtoffers als de maatschappij (Stoltenborgh e.a., 2015). Een effectieve aanpak gericht op het voorkomen van mishandeling is dan ook essentieel. Hoewel de preventie van kindermishandeling steeds meer aandacht krijgt, laten diverse meta-analyses zien dat preventieve interventies ter voorkoming van kindermishandeling slechts in beperkte mate effectief zijn (e.g., Euser e.a., 2015; [AUTEURS GEMASKEERD VOOR REVIEW]). Een mogelijke oorzaak van deze matige effectiviteit is dat interventies niet of onvoldoende zijn toegespitst op zorgbehoeften en kenmerken die per risicogezin kunnen verschillen (Ng & Weisz, 2016; Weisz, 2014), terwijl het personaliseren van behandeling de effectiviteit van interventies juist kan vergroten, zoals onder meer het bewezen effectieve *Risk-Need-Responsivity*-model beschrijft (Andrews e.a., 1990; Andrews e.a., 2011; Bonta and Andrews, 2007; Bonta and Andrews, 2016). Dit model is ontwikkeld voor het effectief inrichten van strafrechtelijke zorg en schrijft in een aantal principes voor hoe interventies vormgegeven moeten worden om de effectiviteit van zorg te vergroten. In diverse studies is inmiddels betoogd dat dit model ook toepasbaar is in jeugdbescherming en dat het werken volgens de RNR-principes de effectiviteit van preventieve interventies voor kindermishandeling kan vergroten ([AUTEURS GEMASKEERD VOOR REVIEW]). Om de RNR-principes toepasbaar te maken in de context van jeugdbescherming zijn recentelijk de [TOOL GEMASKEERD] ([AUTEURS GEMASKEERD VOOR REVIEW]) en [TOOL GEMASKEERD] ([AUTEURS GEMASKEERD VOOR REVIEW]) ontwikkeld. [TOOL GEMASKEERD] heeft als doel professionals te ondersteunen in het taxeren van zorgbehoeften en het kiezen van passende zorg voor risicogezinnen in de jeugdbescherming ([AUTEURS GEMASKEERD VOOR REVIEW]). Het doel van het huidige onderzoek was het verkrijgen van meer inzicht in (1) de wijze waarop het besluitvormingsproces over passende interventies plaatsvindt in de klinische jeugdbeschermingspraktijk, en (2) de praktische bruikbaarheid van [TOOL GEMASKEERD] ter ondersteuning van het besluitvormingsproces over passende interventies.

1.2. RNR-Principes

In strafrechtelijke zorg worden de Risk-Need-Responsivity principes veelvuldig toegepast om recidive op effectieve wijze te voorkomen (Andrews e.a., 1990). Het RNR-model beschrijft

op basis van drie principes hoe behandelingen gepersonaliseerd moeten worden voor individuele cliënten om de behandel-effectiviteit te vergroten. Zo stelt het risicoprincipe dat de intensiteit van een interventie moet aansluiten bij het recidiverisico en het behoefteprincipe dat een interventie gericht moet zijn op criminogene behoeften (i.e. de veranderbare risicofactoren die samenhangen met recidive). Tot slot schrijft het responsiviteitsprincipe voor dat interventies moeten aansluiten bij onder andere het cognitieve niveau, de motivatie en demografische eigenschappen (waaronder geslacht en etnische afkomst) van de cliënt. Interventies die zijn vormgegeven volgens deze principes blijken effectiever in het voorkomen van recidive dan interventies die niet voldoen aan deze principes (Andrews e.a., 1990; Dowden & Andrews, 1996; 2003; 2006; Hanson e.a., 2009).

Hoewel het RNR-model is ontwikkeld ter voorkoming van criminele recidive, is het model ook veelbelovend voor de preventie van (herhaling van) kindermishandeling ([AUTEURS GEMASKEERD VOOR REVIEW]). Zo kunnen zowel delinquent gedrag als kindermishandeling worden verklaard door de balans tussen risicofactoren (denk aan armoede, stress en psychopathologie) en beschermende factoren (zoals sociale steun) in verschillende sociale systemen rondom een cliënt of risicogezin, zoals de familie, een school, en de buurt waarin een cliënt of gezin woont (Belsky, 1993; [AUTEURS GEMASKEERD VOOR REVIEW]). Desondanks worden de RNR-principes tot op heden in beperkte mate toegepast in de jeugdbescherming. Om beter te kunnen werken volgens de RNR-principes is onlangs [TOOL GEMASKEERD] ontwikkeld ([AUTEURS GEMASKEERD VOOR REVIEW]), waarmee zorgbehoeften van risicogezinnen op gestructureerde wijze geïnventariseerd kunnen worden (i.e., de veranderbare risicofactoren die samenhangen met kindermishandeling). Daarnaast wordt er met deze tool een koppeling gemaakt tussen getaxeerde zorgbehoeften en het beschikbare zorgaanbod, waarmee [TOOL GEMASKEERD] de toepassing van het behoefteprincipe van het RNR-model faciliteert.

1.3. [TOOL GEMASKEERD]

Professionals kunnen met [TOOL GEMASKEERD] inventariseren welke dynamische (veranderbare) risicofactoren voor kindermishandeling, oftewel zorgbehoeften, aanwezig zijn in een gezin en dus aanknopingspunten bieden voor behandeling. In dit onderdeel voor behoeftetaxatie gaat het om factoren die kindermishandeling veroorzaken en/of in stand houden ([AUTEURS GEMASKEERD VOOR REVIEW]). Voor de ontwikkeling van het behoeftetaxatie-onderdeel van [TOOL GEMASKEERD] werden zulke dynamische risicofactoren geselecteerd op basis van (overzicht)studies naar significante voorspellers van

kindermishandeling (o.a., [AUTEURS GEMASKEERD VOOR REVIEW]; Cash, 2001; Hindley e.a., 2006; [AUTEURS GEMASKEERD VOOR REVIEW]; Stith e.a., 2009). De geselecteerde dynamische risicofactoren werden vervolgens onderverdeeld in *opvoedingsfactoren* (o.a., problemen in de ouder-kind interactie, onvoldoende bescherming/toezicht/veiligheid), *gezinsfactoren* (o.a., weinig sociale steun, financiële problemen), *ouderfactoren* (o.a., opvoeder ervaart stress, problemen met agressie-regulatie), *kindfactoren* (o.a., internaliserende problematiek, leerproblemen) en *doelgroepfactoren* (o.a., kinderen of ouders met een cognitieve beperking, gezinnen met een niet-westerse achtergrond).

Voor de ontwikkeling van de module om professionals te ondersteunen in het kiezen van passende zorg en/of interventies is een database gemaakt van in totaal 116 interventies voor (de preventie van) kindermishandeling die in Nederland beschikbaar zijn. Elk van deze interventies grijpt tenminste aan op één risicofactor voor kindermishandeling. Om interventies te identificeren zijn meerdere bronnen geraadpleegd, namelijk de databank Effectieve Jeugdinterventies (Nederlands Jeugdinstituut), de databank Effectieve interventies huiselijk geweld en seksueel geweld (Ministerie van Volksgezondheid, Welzijn en Sport & Movisie) en een overzicht van Jeugdbescherming Regio Amsterdam met daarin de interventies en andere zorgvormen die worden toegepast bij cliënten. Van elke geïdentificeerde interventie werd de handleiding of het protocol gescreend op de dynamische factoren waarop de interventie aangrijpt. Zo was het mogelijk de interventies die werden opgenomen in de database van [TOOL GEMASKEERD] te koppelen aan de veranderbare risicofactoren (ofwel de mogelijke zorgbehoeften) die gemeten worden met het behoefte-taxatie-onderdeel van [TOOL GEMASKEERD] ([AUTEURS GEMASKEERD VOOR REVIEW]). De verwachting is dat [TOOL GEMASKEERD] de besluitvorming over het inzetten van passende hulpverlening in de jeugdbescherming effectiever, efficiënter en minder subjectief maakt, omdat deze tool het mogelijk maakt om het behoefteprincipe van het RNR model op systematische wijze in de klinische praktijk van jeugdbescherming toe te passen.

2. Methode

2.1. Participanten

In totaal werden bij 15 professionals uit de jeugdzorg, waarvan 12 vrouwen en 3 mannen, semigestructureerde interviews afgenomen. Deze professionals hadden de volgende functies: gezinscoach, medewerker Crisis Interventie Team, klinisch/GZ-(kinder- en jeugd)psycholoog,

jeugdmaatschappelijk werker, gedragswetenschapper, schoolmaatschappelijk werker en jeugdbeschermer. Zij waren werkzaam bij GGD Hollands-Midden, Opvoedpoli, Levvel, Bureau Jeugdzorg (Limburg) en Bascule.

2.2. Procedure

De professionals werden geworven via de instellingen die betrokken waren bij het [PROJECT GEMASKEERD VOOR REVIEW] en via het netwerk van de onderzoekers in de periode september 2020 Tot en met januari 2021. De participanten ontvingen informatie over het doel en de procedure van het onderzoek, waarna zij een informed-consent-formulier ondertekenden. De interviews duurden gemiddeld 37 minuten en werden op de werkplek van de participant of via een videoverbinding afgenomen. Er werd toestemming gevraagd voor het opnemen van het interview, met de expliciete vermelding dat de data vertrouwelijk en geanonimiseerd zouden worden verwerkt.

2.3. Instrumenten

2.3.1 Interviews

In de semigestructureerde interviews werd een vooraf opgestelde vragenlijst gebruikt. De interviews bestonden uit de volgende onderdelen:

1. De Jeugdbeschermingspraktijk zonder [TOOL GEMASKEERD]

1.1. Vragen over de huidige behoeftetaxatie-procedure en het besluitvormingsproces over passende zorg in de jeugdbeschermingspraktijk;

1.2. Vragen over eventuele knelpunten in de toeleiding naar passende zorg;

1.3. Vragen over mogelijke oplossingen ter verbetering van de zorg;

1.4. Taxatie van zorgbehoeften bij een fictieve gezinscasus zonder [TOOL GEMASKEERD] (Bijlage 1);

1.5. Vragen over welke zorg past bij de casus;

2. Behoeftetaxatie en keuze passende zorg met [TOOL GEMASKEERD]

2.1. Introductie van [TOOL GEMASKEERD]: taxatie van de zorgbehoeften in de casus

2.2. Evaluatie van de taxatieresultaten met [TOOL GEMASKEERD] en vragen over de gebruikerservaringen met [TOOL GEMASKEERD].

2.3.2 [TOOL GEMASKEERD]

[TOOL GEMASKEERD] is een computerapplicatie die is ontwikkeld om zorgprofessionals te ondersteunen in enerzijds het taxeren van zorgbehoeften en anderzijds het kiezen van zorg

die aangrijpt op de getaxeerde zorgbehoeften ([AUTEURS GEMASKEERD VOOR REVIEW]). Deze zorg bestaat uit werkwijzen, behandelingen en interventies die volgens handleidingen en protocollen aangrijpen op tenminste één risicofactor voor kindermishandeling. In Figuur 1 is een schermafbeelding zichtbaar van de applicatie. Links in het scherm is het behoeftetaxatie-onderdeel zichtbaar waarin mogelijke zorgbehoeften in de vorm van dynamische risicofactoren voor kindermishandeling staan. Deze factoren zijn onderverdeeld in opvoedings-, gezins-, ouder-, en kindfactoren. In Figuur 1 worden de factoren getoond voor het tabblad opvoedingsfactoren. Rechts in het scherm worden de interventies getoond die aangrijpen op de geselecteerde factoren.

Hulpverleners kunnen meerdere factoren per tabblad selecteren. Met een aparte knop ('Toelichting risicofactoren') kan een toelichting voor elke factor in [TOOL GEMASKEERD] worden opgevraagd, zodat gebruikers de factoren zoveel mogelijk hetzelfde interpreteren. Nadat een gebruiker alle aanwezige zorgbehoeften voor een casus heeft geselecteerd, kunnen interventies worden gezocht die aangrijpen op deze getaxeerde zorgbehoeften (zie knop 'Zoek interventies/Profielen' in Figuur 1). Zo zijn in Figuur 1 drie opvoedingsfactoren geselecteerd, waarbij in de database van [TOOL GEMASKEERD] vier interventies zitten die elk aangrijpen op (tenminste) deze drie opvoedingsfactoren (zie de interventies rechts in het scherm; Figuur 1). Van elke getoonde interventie kan een toelichting worden opgevraagd (knop 'Toelichting interventie'), waarna informatie volgt over de doelstelling van een interventie, de doelgroep waarop een interventie zich richt, beschikbare kennis over de effectiviteit van een interventie en waar in Nederland een interventie wordt aangeboden. Zoals Figuur 1 laat zien worden de interventies die aangrijpen op de selectie van zorgbehoeften getoond in afnemende volgorde van effectiviteitsbewijs waarbij de classificatie van Van Yperen, Veerman en Bijl (2017) is aangehouden. Na elke taxatie van zorgbehoeften en een zoekopdracht naar interventies die aangrijpen op de getaxeerde behoeften kunnen de resultaten van de behoeftetaxatie en de zoekopdracht worden opgeslagen.

Figuur 1

Schermafbeelding van [TOOL GEMASKEERD] met daarin het Behoeftetaxatie-Onderdeel (Links) en Resultaten van de Zoekopdracht naar Interventies die Aangrijpen op de Getaxeerde Zorgbehoeften (Rechts)

[FIGUUR GEMASKEERD VOOR REVIEW]

2.4. Data-Analyse

De audio-opnames van de interviews zijn getranscribeerd en vervolgens gecodeerd met ATLAS.ti (versie 8) volgens de codeerfases uit de richtlijnen van Boeije (2014). Eerst werden relevante fragmenten en quotes voorzien van een code (open coderen), waarna overlappende thema's werden verzameld in code-groepen (axiaal coderen). Ten slotte werden thema's vergeleken en aan elkaar gerelateerd in netwerken (selectief coderen). Het coderen van alle interviews resulteerde in een totaal van 159 codes, verdeeld over 16 codegroepen.

3. Resultaten

3.1. De Jeugdbeschermingspraktijk zonder [TOOL GEMASKEERD]

3.1.1. Huidige Werkwijze Behoefte-taxatie en Zorgkeuze

De meeste professionals ($n = 12$) taxeren zorgbehoeften tijdens een (intake)gesprek met de cliënt/het gezin (*"In eerste instantie gaan wij in gesprek met ouders en de kinderen om te bekijken wat de grootste zorg is in het gezin."*). Zes professionals noemden daarbij dat overleg met collega's belangrijk is in het bepalen van de primaire zorgbehoeften en het opstellen van een behandelplan (*"Ik geef tijdens een vergadering aan waar ik naar op zoek ben en vervolgens geven collega's suggesties over welke zorg passend is."*). Ook gaven sommige professionals ($n = 4$) aan dat ze over zorgbehoeften overleggen met andere betrokken instanties (waaronder zorgverwijzers en gemeenten) (*"Wij hebben een casuïstiekbespreking samen met ouders en het kind, ketenpartners en de gemeente."*). Vijf professionals gaven aan meestal af te gaan op hun eigen ervaring in het taxeren van zorgbehoeften (*"Bij ons is er wel een gestructureerd protocol, maar in de praktijk doet iedere collega het toch op zijn eigen manier."*). De participanten noemden verschillende instrumenten/methodieken in hun antwoord op de vraag hoe zij doorgaans een gestructureerde behoefte-taxatie uitvoeren: ARIJ-Risk, RiS 2.0, het meten van de Big Five middels een persoonlijkheidsvragenlijst, het Dialoog-model, Richtlijnen Jeugdhulp, de TOP-3 methodiek, SIRE-lijst en Methodiek Spoedhulp.

De meeste professionals ($n = 9$) gaven aan in overleg met collega's te bepalen welke zorg/interventies nodig zijn voor de cliënt/het gezin (*"Ik neem contact op met onze gedragswetenschapper en daarnaast werk ik veel samen met een zorgcoördinator. Soms raadpleeg ik een consulent van het gemeentelijke jeugdloket."*). Ook wordt vaak zorg ingezet op basis van eerdere ervaringen met cliënten (*"Ik kies voor zorg of een interventie waarmee ik bekend ben in mijn omgeving, of op grond van mijn eerdere ervaringen met bepaalde zorg of interventies"*). Drie professionals gebruikten de Zorgkeuze Midden-Holland (Gemeente

Gouda, 2021), wat een website omvat met daarop een sociale kaart van zorgaanbieders in de regio. Daarnaast gaven drie professionals aan rekening te houden met het zorgverleden, het cognitief niveau of de identiteit (waaronder geloofsovertuiging) van cliënten bij het vinden van passende zorg. Ook noemde een professional vaak even te ‘Googelen’ naar welke zorg passend kan zijn.

3.1.2. Knelpunten in de Toeleiding naar Passende Zorg

Wachlijsten en Stroeve Samenwerking. De meeste professionals ($n = 12$) noemden de lange wachlijsten in de jeugdzorg als groot knelpunt in de toeleiding van cliënten naar passende zorg (“*De route die een cliënt moet afleggen om een doorverwijzing te krijgen van de gemeente verloopt echt moeizaam. Bij het gemeentelijke loket is de gemiddelde wachttijd zo ’n zes weken. Vervolgens wordt daar de zorgbehoefte nogmaals in kaart gebracht en pas daarna wordt een aanmelding gedaan bij een zorginstantie, waarna eerst weer enkele maanden wachttijd verstrijkt.*”). Daardoor wordt soms ook een volgens de professionals minder passend alternatief aangeboden (“*Soms is een wachttijd te lang, en dan ga ik op zoek naar een alternatief.*”, of “*Ik zit soms te stutten en steunen totdat ik de hulp kan inzetten die passend is.*”). Zes professionals gaven daarnaast aan dat een gebrek aan samenwerking of wisselingen tussen hulpverleners/instanties soms ook leidt tot vertraging in de zorg (“*Je moet net de goede persoon treffen die zegt: ik ga dat gelijk voor je regelen.*”).

Beperkte Inkoop, Ingewikkeld Financieringsbeleid en Lastige Verwijscriteria. Zeven professionals noemden een beperkte inkoop van zorg door gemeenten als knelpunt bij het inzetten van passende zorg (“*Het inkoopbeleid van de gemeente is zó leidend dat je niet altijd de best passende zorg kan inzetten.*”). Twee professionals gaven daarnaast aan dat het verkrijgen van de benodigde financiering van zorg ook vaak leidt tot vertraging (“*Ik wacht vaak op administratieve rompslomp, terwijl het juist zo belangrijk is om aan een gezin te laten zien dat sprake is van een goede samenwerking met andere partijen.*”). Drie professionals noemden ook dat ingewikkelde verwijscriteria een barrière vormen voor aanmelding bij zorginstanties (“*Soms valt een gezin net niet onder de ingewikkelde verwijscriteria van een instantie, en dat is niet altijd even makkelijk.*”).

Onvoldoende zicht op Interventieaanbod en Zorgbehoeften. Drie professionals vertelden niet altijd bekend te zijn met het hulpaanbod (“*Ik vraag me wel eens af: ‘Is er niet een betere oplossing?’ Ik kies toch snel voor een interventie waarmee ik bekend ben, terwijl er misschien een beter alternatief is.*”). Ook gaven twee professionals aan dat ze soms te weinig informatie over problematiek en zorgbehoeften hebben om een inschatting te kunnen

maken van welke zorg of hulp nodig is (“*Het is lastig wanneer je nog weinig informatie hebt en snel beslissingen moet nemen.*”).

Complexe Problematiek en Gebrekkige Motivatie. Vijf professionals gaven aan dat er niet altijd passende hulp beschikbaar is, omdat de problematiek te ingewikkeld is voor het zorgaanbod waaruit gekozen kan worden (“*Bij complexe echtscheidingen vind ik het heel lastig om passende hulp in te zetten, omdat ouders tegelijkertijd met een rechtszaak bezig zijn.*”). Vier professionals benoemden het gebrek aan motivatie bij cliënten als knelpunt (“*Wat doe je als mensen gewoon geen hulp willen?*”).

3.1.3. Suggesties ter Verbetering van de Jeugdbeschermingspraktijk

Professionals gaven meerdere suggesties op de vraag wat er nodig zou zijn om de knelpunten in het vinden van passende zorg in de jeugdbescherming te verhelpen. Zo is er behoefte aan meer kennis over (1) het hulpaanbod (“*Het zou helpend zijn om meer kennis te hebben over het aanbod in de regio.*”), (2) een betere samenwerking tussen betrokken instanties, (3) een specifiek zoekstelsel met mogelijke interventies (“*Thema’s zijn nu zo algemeen, zoals ‘autisme’ of ‘psychiatrie’, ik wil specifiekere kunnen zoeken.*”), (4) het inzetten van wachtlijstcoaches of een casusregisseur, en (5) een aangepast financieringssysteem (“*Het zou mooi zijn als ieder kind een persoonlijk budget krijgt, want dan maakt het niet uit met welke instanties er wel of geen contracten zijn.*”).

3.2. Behoeftetaxatie Met- en Zonder [TOOL GEMASKEERD]

Na het lezen van de fictieve casus bleek dat professionals significant meer zorgbehoeften taxeerden met gebruik van [TOOL GEMASKEERD] ($M = 21.67$, $SD = 5.98$), dan tijdens de eerste taxatieronde die gebaseerd was op alleen de eigen klinische inschatting ($M = 9.07$, $SD = 3.88$) ($t(28) = -6.84$, $p < .001$) (zie Tabel 2). Professionals taxeerden significant meer *familiefactoren* (o.a. huiselijk geweld, een problematische relatie tussen opvoeders en financiële problemen), *opvoedingsfactoren* (o.a. inconsistente opvoeding, verstoorde ouder-kindrelatie en problemen met het stellen van grenzen) en *ouderfactoren* (o.a. ouderlijke stress en agressieregulatie-problematiek).

Tabel 2

Aantal Gesignaleerde Zorgbehoeften (zonder en met [TOOL GEMASKEERD])

Zonder [TOOL GEMASKEERD]	Met [TOOL GEMASKEERD]
--------------------------	-----------------------

Categorie	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>t</i>
Familie	3.53	1.06	5.87	0.83	-6.70***
Kind	2.53	2.48	3.60	1.50	-1.43
Opvoeding	1.73	0.88	6.40	3.38	-5.18***
Ouder	1.27	1.28	5.80	2.11	-7.11***
Totaal	9.07	3.88	21.67	5.98	-6.84***

Noot. Voor elke categorie is een independent samples *t*-test uitgevoerd om het verschil in het gemiddeld aantal getaxeerde zorgbehoeften te toetsen tussen de behoeftetaxaties met en zonder [TOOL GEMASKEERD]. Het gebruik van de tool was een dichotome variabele (0 = zonder [TOOL GEMASKEERD], 1 = met [TOOL GEMASKEERD]), wat betekent dat een negatieve *t*-waarde een hoger gemiddeld aantal getaxeerde zorgbehoeften impliceert met [TOOL GEMASKEERD].

*** $p < .001$.

3.3. Zorgkeuze Met- en Zonder [TOOL GEMASKEERD]

Na de behoeftetaxatie zonder [TOOL GEMASKEERD] gaven professionals verschillende suggesties voor zorg die zij zouden inzetten bij de fictieve gezinscasus, waaronder: sociaal maatschappelijke ondersteuning, systeemgerichte zorg, opvoedondersteuning, ambulante spoedhulp, mediation, stresshantering, veiligheidsafspraken, logopedie voor het kind en een kindercoach. Ook werden verschillende interventies genoemd, waaronder: ‘Kinderen uit de Knel’, ‘En nu ik...!’, ‘Families First’, ‘KIES’, ‘Ouderschap Blijft’, ‘Piep zei de muis’, ‘SCHIP-aanpak’, ‘Signs of Safety’ en ‘Words and Pictures’.

Bij de eerste taxatie met [TOOL GEMASKEERD], waarbij de respondenten alle aanwezige zorgbehoeften bij de casus taxeerden, bleek dat maar twee interventies in [TOOL GEMASKEERD]-database zaten die op *alle* aanwezige zorgbehoeften in de casus aansloten. Multisysteemtherapie – Child Abuse and Neglect (MST-CAN) kwam bijna altijd naar voren (93.3%), en Intensieve Ambulante Gezinsbegeleiding kwam één keer naar voren (6.7%). Bij de tweede taxatie met [TOOL GEMASKEERD], waarbij respondenten de zorgbehoeften met de meeste prioriteit taxeerden, bleken meer interventies in [TOOL GEMASKEERD]-database passend te zijn bij de casus. MST-CAN bleek ook hierbij het vaakst passend (93.3%), gevolgd door Ambulante Spoedhulp (66.7%), Voorwaardelijke Interventies in Gezinnen (60%), Families First en Veilig Sterk en Verder (46.7%) en Parent Child Interaction Therapy en Intensieve Ambulante Gezinsbegeleiding (40%).

3.4 Evaluatie van [TOOL GEMASKEERD]

Interventieaanbevelingen. De meeste professionals (10) konden zich vinden in de interventieaanbevelingen die volgden uit de behoefte-taxatie met [TOOL GEMASKEERD]. Acht professionals gaven aan niet bekend te zijn met één of meerdere interventies. Twee professionals vonden de interventies wel passend, maar pas in een later behandelstadium (*“Ouders moeten eerst een stabiele woonomgeving krijgen.”*).

Module voor Behoeftetaxatie. Bij de boordeling van de bruikbaarheid van het behoefte-taxatie-onderdeel was het merendeel van de feedback positief. Professionals benoemden onder andere dat het overzicht van de verschillende domeinen structuur biedt en handig is tijdens het schrijven van rapportages. Ook gaven ze aan dat de tool makkelijk in teamverband gebruikt kan worden en goed inzetbaar is bij multiprobleemgezinnen. Daarnaast gaven professionals aan dat er overlap zit in sommige factoren (*“Vooral bij opvoedingsfactoren, waarbij veel andere factoren lijken te vallen onder de algemene factor ‘er is onvoldoende stabiliteit in de opvoeding’.”*). Ook benoemden twee professionals het belang van het kunnen aangeven van de ernst van de problematiek, oftewel dat de huidige dichotome antwoordoptie (wel/niet aanwezig van een factor) uitgebreid zou moeten worden met een ernstgradatie.

Module voor Zorgkeuze. Vijf professionals noemden het interventieoverzicht in [TOOL GEMASKEERD] een verbreding van de horizon (*“Het geeft een mooi beeld van de mogelijkheden, ook van interventies waaraan je in eerste instantie niet zou denken. Normaal gesproken kies je toch voor een interventie waar je bekend mee bent, of waar je goede ervaringen mee hebt.”*). Zes professionals vonden het interventieoverzicht in het algemeen nuttig (*“Het is best concreet en ik denk dat het heel helpend kan zijn.”*). Drie professionals benoemden dat het interventieoverzicht moet aansluiten bij het beschikbare aanbod in de regio waarin gewerkt zou worden met [TOOL GEMASKEERD]. Verder gaven professionals aan dat het zinvol zou zijn om het interventieaanbod uit te breiden met diverse methodieken, waaronder Signs of Safety en de MPG-Aanpak (ketenaanpak voor zorg aan multiprobleemgezinnen, regio Amsterdam).

Ontbrekende Informatie. Alle professionals gaven aan dat er geen overbodige informatie in de tool staat. Wel zeiden drie professionals de ruimte voor positieve/beschermende factoren (zoals sociale steun) te missen in de tool. Andere missende factoren waren volgens de professionals: hulpverleningsgeschiedenis, (complexe) scheiding van ouders, specifieke kindfactoren (i.p.v. algemeen externaliserende- en internaliserende problematiek), trauma (kind), leeftijdsindicaties van interventies en het mentaliserend

vermogen van ouders. Daarnaast misten enkele participanten de optie voor prioritering van factoren (ernstgradatie) en de verwerking van veiligheidsafspraken als eventuele bijlage.

Gebruikservaringen en Suggesties. Twaalf professionals vonden [TOOL GEMASKEERD] over het algemeen een gebruiksvriendelijke en overzichtelijke tool (“*Ik vind hem heel overzichtelijk, vooral de tabbladen/categorieën zijn erg helpend.*”). Drie professionals suggereerden dat het makkelijker zou zijn als de Ctrl-toets niet ingedrukt hoeft te worden bij de selectie van behoeftefactoren. Andere suggesties waren: een minder ‘basic’ ontwerp, online beschikbaarheid van de tool, minder gebruik van jargon, ‘onbekend’ toevoegen als antwoordoptie en continue zichtbaarheid van de toelichtingen bij de factoren. Eén professional benoemde dat het ook nuttig zou zijn als ouders de tool zelf kunnen gebruiken om te onderzoeken welke hulp zij nodig zouden hebben.

Gebruik in de Praktijk. Alle professionals gaven aan de tool in de praktijk te willen gebruiken (“*Je kiest toch snel voor een interventie die je vaker hebt ingezet en waar je positieve ervaringen mee hebt in plaats van een minder bekende, maar eventueel beter passende interventie. Ik denk dat je hiermee een heel goed beeld krijgt van de mogelijkheden.*”). Ook zouden de professionals de tool bij collega’s aanraden (“*Zeker voor wijkteams of spoedhulp.*”). Ten slotte benoemden meerdere professionals dat het nuttig zou zijn om [TOOL GEMASKEERD] te koppelen aan de [TOOL GEMASKEERD].

4. Discussie

[TOOL GEMASKEERD] is, voor zover bekend, het eerste instrument dat specifiek ontwikkeld is om professionals in jeugdbescherming te ondersteunen in behoeftetaxatie én besluitvorming over passende zorg. Het doel van deze studie was om te onderzoeken of [TOOL GEMASKEERD] daadwerkelijk ondersteunend is in behoeftetaxatie en het besluitvormingsproces. Om hier uitspraken over te kunnen doen werd in deze studie ook nagegaan wat de huidige werkwijze is van professionals voor het taxeren van zorgbehoeften en het kiezen van passende zorg.

4.1. [TOOL GEMASKEERD] in de Jeugdbeschermingspraktijk

Uit dit onderzoek blijkt dat ongestructureerde besluitvorming over passende zorg nog altijd gebruikelijk is in de jeugdbeschermingspraktijk, ondanks de voordelen van een gestructureerd klinisch besluitvormingsproces (Douglas et al., 2002; [TOOL GEMASKEERD]). De jeugdzorgprofessionals die in dit onderzoek zijn geïnterviewd baseren hun keuze over passende zorg doorgaans op eigen ervaringen en adviezen van collega’s. Valkuilen van

dergelijke intuïtieve besluitvorming zijn het optreden van vooroordelen (bias) in het beoordelen van een casus en het diagnostisch proces (Helm, 2011; Saltiel, 2016). Zo kan een professional door bias in de oordeelsvorming vooral aandacht hebben voor bepaalde herkenbare problematiek door eerdere ervaringen met dezelfde problemen en minder aandacht hebben voor andere (complexe) informatie die niet direct samenhangt met de gesignaleerde problematiek (Saltiel, 2016).

De resultaten van dit onderzoek sluiten aan op eerder onderzoek waaruit blijkt dat gestructureerde besluitvorming (zoals met [TOOL GEMASKEERD]) tunnelvisie kan voorkomen en een holistische benadering van probleemgezinnen in de jeugdbescherming bevordert (Bartelink et al., 2015). Bij een holistisch besluitvormingsproces wordt naast zorgbehoeften van kinderen rekening gehouden met zorgbehoeften van de opvoeders en/of de bredere sociale omgeving. Zo bleek dat professionals tijdens de gestructureerde behoefte-taxatie met [TOOL GEMASKEERD] significant meer zorgbehoeften signaleerden dan tijdens de klinische taxatie zonder [TOOL GEMASKEERD]. Daarnaast viel op dat professionals vooral gezins- en kindfactoren en relatief weinig ouderfactoren taxeerden in hun eigen klinisch oordeel van zorgbehoeften, terwijl zij significant meer gezins-, ouder-, en opvoedingsfactoren signaleerden wanneer gewerkt werd met [TOOL GEMASKEERD]. Hiermee lijkt [TOOL GEMASKEERD] een belangrijke meerwaarde te hebben ten opzichte van ongestructureerde behoefte-taxatie, omdat opvoedingsfactoren (o.a. problematische ouder-kind relatie) en ouderfactoren (o.a. crimineel gedrag en psychopathologie) aanzienlijk sterkere voorspellers zijn voor kindermishandeling dan kindfactoren ([AUTEURS GEMASKEERD VOOR REVIEW]; Stith et al., 2009). Ook uit een calamiteitenonderzoek (Toezicht Sociaal Domein, 2015) bleek dat professionals onvoldoende aandacht hadden voor risicofactoren van ouders in kwetsbare gezinnen waarin de veiligheid van kinderen in gevaar was. [TOOL GEMASKEERD] kan professionals dus ondersteunen in het signaleren van de meest relevante risicofactoren die fungeren als aanknopingspunten voor behandeling.

Over de gebruiksvriendelijkheid van [TOOL GEMASKEERD] zeiden professionals dat het overzicht van risicofactoren in verschillende domeinen structuur biedt bij het in kaart brengen van hulpbehoeften. Daarbij werd benoemd dat zo'n overzicht belangrijke en inzichtelijke aanknopingspunten biedt voor het maken van doeltreffende rapportages of gezinsplannen, wat gezinswerkers vooral beschouwen als een administratieve last (Sekreve et al., 2020). De zorgkeuzemodule van [TOOL GEMASKEERD] werd gezien als een 'verbreding van de horizon' in het voortdurend groeiende aanbod van zorg en interventies, die al dan niet effectief zijn bevonden. Uit recent onderzoek blijkt dat veel gezinswerkers

onbekend zijn met effectieve behandelvormen die bijvoorbeeld zijn opgenomen in de NJI Databank Effectieve Jeugdinterventies (DEJ) (Sekreve et al., 2020). [TOOL GEMASKEERD] vergroot niet alleen de kennis van professionals over het beschikbare aanbod van behandelingen en interventies, maar maakt ook inzichtelijk welke interventies aangrijpen op de risicofactoren die professionals selecteren in een behoeftetaxatie. Bovendien maakt [TOOL GEMASKEERD] inzichtelijk in welke mate de werkzaamheid van interventies wetenschappelijk is onderbouwd ([TOOL GEMASKEERD]), wat een belangrijk aspect is om mee te wegen in besluitvorming over passende zorg.

4.2. Aanbevelingen voor de Doorontwikkeling van [TOOL GEMASKEERD]

De geïnterviewde professionals noemden ook suggesties om de tool te verbeteren en de inzetbaarheid ervan in de jeugdbeschermingspraktijk te vergroten. Ten eerste noemden zij dat niet alle interventies in de database van [TOOL GEMASKEERD] beschikbaar zijn in elke regio in Nederland vanwege het decentrale, gemeentelijk inkoopbeleid van jeugdzorg. Enerzijds belemmert dat de algemene bruikbaarheid van de tool, maar anderzijds biedt de tool op deze manier wel inzicht in lacunes in het regionale hulpaanbod. Zo kunnen jeugdzorginstanties in overleg met gemeenten aangeven welke nog niet beschikbare interventies geschikt zouden zijn voor hun cliënten.

Ten tweede merkten professionals op dat de huidige module voor behoeftetaxatie alleen (veranderbare) risicofactoren meet en geen beschermende factoren. In de doorontwikkeling van [TOOL GEMASKEERD] kan worden overwogen om beschermende factoren op te nemen in de behoeftetaxatie om het draagvlak voor deze tool in de jeugdbescherming te vergroten (Wilkins, 2015). Het blijft daarbij wel belangrijk dat dynamische risicofactoren centraal staan in zowel behoeftetaxatie als de behandeling van gezinnen, zodat de klinische praktijk aansluit bij het bewezen effectieve Needs-principe van het RNR-model (REF). Bovendien blijkt uit onderzoek dat kinderen uit hoog-risico-gezinnen minder profiteren van de eventuele aanwezigheid van beschermende factoren dan laag-risico-gezinnen (Luthar & Goldstein, 2004). Ten derde bleek dat er slechts een beperkt aanbod is van interventies die aansluiten op multidimensionale gezinsproblematiek waarbij veel en diverse behoeftefactoren aanwezig zijn, zoals het geval was in de casus die in dit onderzoek werd voorgelegd aan de respondenten. Naarmate minder risicofactoren worden geselecteerd in [TOOL GEMASKEERD] neemt het aantal interventies dat aangrijpt op alle geselecteerde risicofactoren toe en daardoor toont [TOOL GEMASKEERD] ook meer passende interventies in de zorgkeuze-module. Professionals gaven de suggestie om een ernstgradatie toe te voegen

aan de dichotome antwoordcategorieën (wel/niet aanwezig) bij de behoefte-taxatie. Met deze mogelijkheid kunnen professionals zicht krijgen op behoeftefactoren met de meeste prioriteit en interventiemogelijkheden die aansluiten op deze factoren. Ten slotte is het belangrijk om naast gebruikerservaringen ook de psychometrische kwaliteit van [TOOL GEMASKEERD] te onderzoeken. Hoewel uit eerder onderzoek blijkt dat toepassing van RNR-instrumenten de betrouwbaarheid en predictieve validiteit van het besluitvormingsproces in de jeugdbescherming versterken ([AUTEURS GEMASKEERD VOOR REVIEW]), blijft het een urgente vraag of implementatie van [TOOL GEMASKEERD] werkelijk leidt tot betere besluitvorming en dus tot minder kindermishandeling.

4.3. Klinische Implicaties

In dit onderzoek benoemden professionals knelpunten in de huidige jeugdbeschermingspraktijk die ook in eerdere onderzoeken naar voren zijn gekomen, zoals lange wachtlijsten, beperkingen in het hulpaanbod door het inkoopbeleid van gemeenten en een gebrekkige samenwerking tussen instanties ([AUTEURS GEMASKEERD VOOR REVIEW]; Inspectie Gezondheidszorg en Jeugd, 2020). Een gestructureerd, consistent besluitvormingsproces over passende interventies kan de samenwerking tussen partijen in de jeugdbeschermingspraktijk verbeteren (De Bortoli et al., 2016). Het is belangrijk om te benadrukken dat [TOOL GEMASKEERD] geen vervanging is van het oordeel van professionals, omdat casus-specifieke en tijdgebonden factoren, zoals ernst en spoedeisendheid, meewegen in de keuze voor de best passende interventie ([AUTEURS GEMASKEERD VOOR REVIEW]). Daarnaast is een inschatting van responsiviteitsfactoren van cliënten in de jeugdbescherming (waaronder niveau van motivatie, het cognitief functioneren van opvoeder(s) en de culturele achtergrond van een gezin) belangrijk om zorg op maat te kunnen bieden wat de effectiviteit van de zorg ten goede komt ([AUTEURS GEMASKEERD VOOR REVIEW]). Ook worden ouders en kinderen bij voorkeur betrokken bij het besluitvormingsproces om het vertrouwen en de kans op een positieve samenwerking te vergroten (van Bijleveld et al., 2019; Helm, 2011; [AUTEURS GEMASKEERD VOOR REVIEW]).

4.4. Conclusie

Uit deze pilotstudie blijkt dat [TOOL GEMASKEERD] professionals ondersteunt in behoefte-taxatie en besluitvorming over passende zorg voor het verlagen van het risico op (de herhaling van) kindermishandeling. Professionals signaleerden significant meer

zorgbehoeften bij een multiprobleemgezin wanneer zij zorgbehoeften taxeerden met [TOOL GEMASKEERD] dan wanneer zij taxeerden volgens hun eigen klinische ongestructureerde oordeel. Ook signaleerden professionals meer opvoedings- en ouderfactoren met [TOOL GEMASKEERD], wat de sterkste voorspellers zijn voor kindermishandeling en daarmee belangrijke aanknopingspunten voor behandeling. De zorgkeuzemodule van [TOOL GEMASKEERD] werd beoordeeld als een ‘verbreding van de horizon’ in het voortdurend groeiende aanbod van zorg en interventies in de jeugdbeschermingspraktijk.

Korte Verklaring over de Data en Analyses

Analyse codes en output: De analyse codes en output zijn niet vrij toegankelijk, maar wel opvraagbaar bij de auteur(s).

Data: Wegens gevoeligheid van de (interview)data zijn deze niet vrij toegankelijk.

Informatie/vragen over de data is op te vragen bij/kunnen gesteld worden aan
[GEMASKEERD VOOR REVIEW].

Financiering

Deze studie is gesubsidieerd door ZonMw (projectnummer: [GEMASKEERD VOOR REVIEW]).

Belangen

[GEMASKEERD VOOR REVIEW]

Literatuur

[REFERENTIE GEMASKEERD VOOR REVIEW]

Andrews, D. A., Bonta, J., & Hoge, R. D. (1990). Classification for effective rehabilitation: Rediscovering psychology. *Criminal Justice and Behavior*, 17, 19-52.

<https://doi.org/10.1177/0093854890017001004>

Andrews, D. A., Bonta, J., & Wormith, J. S. (2011). The Risk-Need-Responsivity (RNR) Model: Does adding the Good Lives Model contribute to effective crime prevention? *Criminal Justice and Behavior*, 38(7), 735-755.

<https://doi.org/10.1177/0093854811406356>

Bartelink, C., van Yperen, T. A., & ten Berge, I. J. (2015). Deciding on child maltreatment: A literature review on methods that improve decision-making. *Child Abuse & Neglect*, 49, 142-153. <https://doi.org/10.1016/j.chiabu.2015.07.002>

Belsky, J. (1993). Etiology of child maltreatment: A developmental ecological analysis. *Psychological Bulletin*, 114(3), 413-434.

<http://dx.doi.org/10.1037/0033-2909.114.3.413>

Bijleveld, van, G. G., Bunders-Aelen, J. F. G., Dedding, C. W. M. (2019). Exploring the essence of enabling child participation within child protection services. *Child & Family Social Work*, 25(2), 286-293. <https://doi.org/10.1111/cfs.12684>

[REFERENTIE GEMASKEERD VOOR REVIEW]

Boeije, H. (2014). *Analyseren in kwalitatief onderzoek*. Amsterdam: Boom

Bonta, J., & Andrews, D. A. (2016). *The psychology of criminal conduct (sixth edition)*. New York: Routledge

De Bortoli, L., Ogloff, J., Coles, J., & Dolan, M. (2017). Towards best practice: Combining evidence-based research, structured assessment and professional judgement. *Child & Family Social Work*, 22(2), 660-669. <https://doi.org/10.1111/cfs.12280>

Douglas, K. S., & Kropp, P. R. (2002). A prevention-based paradigm for violence risk assessment: Clinical and research applications. *Criminal Justice and Behavior*, 29(5), 617-658. <https://doi.org/10.1177/009385402236735>

Dowden, C., & Andrews, D. A. (1996). What works for female offenders: A meta-analytic review. *Crime and Delinquency*, 45(4), 438-452.

<https://doi.org/10.1177/0011128799045004002>

Dowden, C., & Andrews, D. A. (2000). Effective correctional treatment and violent reoffending: A meta-analysis. *Canadian Journal of Criminology*, 42, 449-467.

<https://doi.org/10.3138/cjcrim.42.4.449>

- Dowden, C., & Andrews, D. A. (2003). Does family intervention work for delinquents? Results of a meta-analysis. *Canadian Journal of Criminology and Criminal Justice*, 45(3), 327-342. <https://doi.org/10.3138/cjccj.45.3.327>
- Euser, S., Alink, L. R. A., Stoltenborgh, M., & Bakermans-Kranenburg, M. J. (2015). A gloomy picture: A meta-analysis of randomized controlled trials reveals disappointing effectiveness of programs aiming at preventing child maltreatment. *BMC Public Health*, 15, 1068. <https://doi.org/10.1186/s12889-015-2387-9>
- Gemeente Gouda. (2021). *Zorgkeuze Midden Holland*. Verkregen via: <https://www.zorgkeuzemiddenholland.nl>
- [REFERENTIE GEMASKEERD VOOR REVIEW]
- Hanson, R. K., Bourgon, G., Helmus, L., & Hodgson, S. (2009). The principles of effective correctional treatment also apply to sexual offenders: A meta-analysis. *Criminal Justice and Behavior*, 36(9), 865-891. <https://doi.org/10.1177/0093854809338545>
- Helm, D. (2011). Judgments or assumptions? The role of analysis in assessing children and young people's needs. *British Journal of Social Work*, 41, 894-911. <https://doi.org/10.1093/bjsw/bcr096>
- Hindley, N., Ramchandani, P. G., & Jones, D. P. (2006). Risk factors for recurrence of maltreatment: A systematic review. *Archives of Disease in Childhood*, 91(9), 744-752. <http://dx.doi.org/10.1136/adc.2005.085639>
- Inspectie Gezondheidszorg en Jeugd, Ministerie van Volksgezondheid, Welzijn en Sport. (2020). *Voortgangsrapportage kwetsbare kinderen onvoldoende beschermd*. Verkregen via: <https://www.rijksoverheid.nl/documenten/rapporten/2020/10/15/tk-bijlage-voortgangsrapportage-igj-en-ijenv-kwetsbare-kinderen-onvoldoende-beschermd>
- Luthar, S. S., & Goldstein, A. (2004). Children's exposure to community violence: Implications for understanding risk and resilience. *Journal of Clinical Child and Adolescent Psychology*, 33(3), 499-505. https://doi.org/10.1207/s15374424jccp3303_7
- [REFERENTIE GEMASKEERD VOOR REVIEW]
- Ng, M. Y., & Weisz, J. R. (2016). Annual research review: Building a science of personalized intervention for youth mental health. *Journal of Child Psychology and Psychiatry*, 57(3), 216-236. <https://doi.org/10.1111/jcpp.12470>

[REFERENTIE GEMASKEERD VOOR REVIEW]

[REFERENTIE GEMASKEERD VOOR REVIEW]

[REFERENTIE GEMASKEERD VOOR REVIEW]

[REFERENTIE GEMASKEERD VOOR REVIEW]

[REFERENTIE GEMASKEERD VOOR REVIEW]

Saltiel, D. (2016). Observing front line decision making in child protection. *British Journal of Social Work*, 46, 2104-2119. <https://doi.org/10.1093/bjsw/bcv112>

Sekreve, A., Campmans, M., van Bon-Martens, M., & Orobio de Castro, B. (2020).

Doelgericht werken in wijkteams laat nog te wensen over. *Kind & Adolescent Praktijk*, 19, 6-12. <https://doi.org/10.1007/s12454-020-0627-3>

Stoltenborgh, M., Bakermans-Kranenburg, M.J., Alink, L. R. A., & Van IJzendoorn, M.H. (2015). The Prevalence of child maltreatment across the globe: Review of a series of meta-analyses. *Child Abuse Review*, 24(1), 37-50. <https://doi.org/10.1002/car.2353>

Stith, S. M., Ting Liu, L., Davies, C., Boykin, E. L., Alder, M. C., Harris, J. M., Som, A., McPherson, M., & Dees, J. E. M. E. G. (2009). Risk factors in child maltreatment: A meta analytic review of the literature. *Agression and Violent Behavior*, 14(1), 13-29. <https://doi.org/10.1016/j.avb.2006.03.006>

Toezicht Jeugd/Sociaal Domein. (2015). *Calamiteitenonderzoek Amsterdam: Borgen van veiligheid in kwetsbare gezinnen*. Verkregen via: <https://www.toezichtsociaaldomein.nl/documenten/rapporten/2015/09/30/calamiteit-amsterdam-borgen-van-veiligheid-in-kwetsbare-gezinnen>

[REFERENTIE GEMASKEERD VOOR REVIEW]

[REFERENTIE GEMASKEERD VOOR REVIEW]

Weisz, J. R. (2014). Building robust psychotherapies for children and adolescents.

Perspectives on Psychological Science, 9(1), 81-84. <https://doi.org/10.1177/1745691613512658>

Wilkins, D. (2015). Balancing risk and protective factors: How do social workers and social work managers analyse referrals that may indicate children are at risk of significant harm. *British Journal of Social Work*, 45, 395-411. <https://doi.org/10.1093/bjsw/bct114>

Yperen, van, T., Veerman, J. W., & Bijl, B. (2017). *Zicht op effectiviteit. Handboek voor resultaatgerichte ontwikkeling van interventies in de jeugdsector*. Rotterdam: Lemniscaat

Bijlage 2 – Casus

Het gezin bestaat uit vader, moeder en hun drie zoons K. (9), I. (7) en U. (3). Ouders zijn na 12 jaar huwelijk uit elkaar gegaan. Over hun relatie is bekend dat er veel ruzie was. Volgens vader is moeder regelmatig vreemdgegaan, al geeft moeder aan dat op deze momenten de relatie tussen vader en moeder over was. Na het verbreken van de relatie heeft vader nog lang hoop gehouden dat het weer goed zou komen. Dit heeft voor veel onduidelijkheid gezorgd over de relatie, ook bij de kinderen.

Ouders strijden nog altijd met elkaar, waarbij ook af en toe sprake is van fysiek geweld. Zo heeft vader recent moeder in het gezicht geslagen en heeft moeder een week later tijdens een ruzie een brandende kaars naar vader gegooid. Vader geeft aan dat moeder hem meerdere malen fysiek heeft aangevallen. Tijdens een van deze ruzies is de politie betrokken geweest.

Moeder heeft geen eigen woning. Ze woont momenteel in een huis dat haar ouders voor haar en de kinderen hebben gehuurd. Het jongste kind woont bij haar, de middelste af en toe. Vader heeft ook geen eigen woning. Hij woont bij zijn ouders. Het oudste kind woont bij vader en de middelste is ook meestal daar.

Moeder heeft momenteel geen werk, ze heeft schulden en staat onder bewind voering. Vader heeft een fulltimebaan bij een bakkerij. Ook vader heeft schulden en is bezig met een plan voor schuldhulpverlening. Moeder voelt zich nog steeds erg afhankelijk van vader en vader probeert hier zoveel mogelijk aan tegemoet te komen. Zo vraagt moeder veel hulp in de opvoeding en vraagt zij regelmatig om geld bij vader.

De kinderen hebben veel moeite met de scheiding van ouders. De oudste zoon K. geeft moeder de schuld van de scheiding. Hij wil dan ook niet graag naar zijn moeder toe. Als hij gaat is hij tegendraads en luistert hij slecht. Moeder belt dan naar vader om hem weer op te halen. Verder heeft K. veel woede-uitbarstingen (schopt en slaat moeder). De woede-uitbarstingen komen volgens vader niet voor als de oudste bij hem is.

De middelste zoon I. neemt het zijn vader kwalijk dat hij moeder heeft geslagen en vindt het lastig om te kiezen waar hij wil wonen. Ouders maken hier geen duidelijke keuze in voor hem. Hij is getuige geweest van het feit dat vader zijn moeder heeft geslagen en dat de politie hieraan te pas moest komen. Dit heeft hij uitgebreid op school in de kring verteld.

De jongste zoon U. moet veel huilen en klampt zich veel aan moeder op het moment dat zij weggaat. Daarnaast slaapt hij veel bij moeder in bed, omdat hij bang is. Opvallend aan de jongste is, is dat hij niet goed verstaanbaar praat.

In de eerste paar gesprekken komt naar voren dat de situatie in het gezin niet verbetert. Vader en moeder willen verbeteren, maar ze blijven veel ruzie maken. Er zijn sterke twijfels of de kinderen wel thuis kunnen blijven wonen. Moeder is bereid met de hulpverlening mee te werken. Ze geeft aan stress te ervaren en moeite te hebben met de opvoeding. Vader wil echter niet meewerken. Hij schaamt zich voor waar hij woont, zijn ouders roken en hij is bang dat de kinderen uit huis worden geplaatst.

Reactie reviewer 1 – 23-11-2022

Het onderzoek is inhoudelijk interessant en past qua onderwerp bij JiO. Het is prettig geschreven. Verder handelt het over een inhoudelijk relevant onderwerp, namelijk besluitvorming in de jeugdbescherming. Wel zijn er een aantal kritische kanttekeningen te plaatsen bij het artikel en dient er meer abstractieniveau aangebracht kunnen worden en consistentie in het artikel tussen doelstelling, theoretische inbedding en resultaten/discussie. Onderstaande opmerkingen kunnen helpen om het manuscript te verbeteren.

1. In de inleiding mist theoretische inkadering van de factoren die besluitvormingsprocessen binnen jeugdbescherming beïnvloeden. Er zijn verschillende theoretische modellen die naast inhoudelijke factoren gerelateerd aan een casus ook andere factoren betrekken als eigenschappen van een professional of gedeelde besluitvorming. Waar focust dit onderzoek zich op als je inzoomt op factoren die van invloed zijn op besluitvorming?
2. Verder onduidelijk waarom een overzicht van Jeugdbescherming Regio Amsterdam met daarin de interventies en andere zorgvormen die worden toegepast bij cliënten is opgenomen. Waarom geen andere regio's? Waar draagt dit extra aan bij? Zegt dit iets over de interventies die in de database zijn opgenomen?
3. Als ik het goed begrijp zou de analyse van verschillende databases met interventies geleid moeten hebben tot een overzicht van factoren/thema's waarop interventies aangrijpen. Is het zinvol om hier een overzicht van te geven? En in hoeverre sluit dit aan bij het overzicht wat eerder genoemd is (gezinsfactoren etc.)
4. Verder zou de vraagstelling aangescherpt kunnen worden ten aanzien van de kennisvraag die erboven hangt. Aan welk kennistekort draagt dit onderzoek bij? Nu is de doelstelling gericht op: "Het doel van het huidige onderzoek was het verkrijgen van meer inzicht in (1) de wijze waarop het besluitvormingsproces over passende interventies plaatsvindt in de klinische jeugdbeschermingspraktijk, en (2) de praktische bruikbaarheid van [TOOL GEMASKEERD] ter ondersteuning van het besluitvormingsproces over passende interventies." Met name de tweede doelstelling is het onduidelijk wat met praktische bruikbaarheid bedoeld wordt en welke theorie hieraan ten grondslag ligt. Hoe toets je of het praktisch bruikbaar is.
5. Beschrijving van het design ontbreekt in de methode. Het betreft kwalitatief onderzoek en daarom is het goed om in het manuscript aan te geven hoe rekening is gehouden met de kwaliteitscriteria voor kwalitatief onderzoek (triangulatie, saturatie etc.)
6. Ik zou graag meer informatie willen over de samenstelling van de onderzoeksgroep. Wat zijn leeftijden, jaren werkervaring, scholingsachtergrond, ervaring met gebruik van de tool etc. Hier kan je ook een koppeling maken met factoren gelieerd aan professional die besluitvorming in jeugdbescherming beïnvloeden.
7. Het verdient de voorkeur om niet herleidbaar tot personen te gebruiken ipv geanonimiseerd.
8. De figuren zijn voor de reviewer niet zichtbaar waardoor het ingewikkeld is om inzicht te krijgen in het gebruik van de tool
9. Onderzoek lijkt te bestaan uit twee delen (ervaringen besluitvormingsprocedures en praktische bruikbaarheid) maar in de methode zie ik dat niet terug, evenals in de resultaten.

10. Onder procedure wordt het niet helder welke rol de casus speelt bij afname van het interview. Wordt er specifiek gekeken welke stappen zijn gezet? Of wordt na gebruik van de tool de ervaringen uitgevraagd?
11. Onduidelijk blijft hoe er precies gecodeerd is. Is er sprake geweest van deductief coderen? Kijkend naar de structuur van de resultaten lijkt de structuur van het interviewprotocol gehanteerd te zijn. Vergelijking tussen wel en niet gebruik van de tool lijkt niet kwalitatief bekeken te zijn maar kwantitatief. Onduidelijk is hoe deze data verzameld is. Ook vraag ik me af wat de waarde van deze uitkomsten is gezien de toch kleine onderzoeksgroep. Ook weten we weinig over de samenstelling van de onderzoeksgroep daarmee wordt de vraag naar representativiteit opgeroepen.
12. In de data-analyse wordt geen aandacht besteed aan de betrouwbaarheid van de analyse. Kan daar informatie over toegevoegd worden?
13. Het instrument wordt toegelicht er lijkt geen onderzoek te zijn gedaan naar de betrouwbaarheid en validiteit. Dit onderzoek lijkt een eerste onderzoek die hier een bijdrage aan moet leveren (?) en is iets terug te lezen is over inhoudsvaliditeit en ecologische validiteit?
14. In de resultaten komen veel verschillende termen/interventies voor die voor niet-ingewijden moeilijk te volgen zijn. Verder worden de resultaten er operationeel (ctrl-toets) en onduidelijk hoe dit bijdraagt aan het verkrijgen inzicht in...zie doelstelling.
15. Er worden aantallen genoemd hoe vaak een professional iets weergeeft, bij kwalitatief onderzoek gaat het vooral om de ervaringen die naar voren gebracht worden en niet de aantallen (daarom hoeft de onderzoeksgroep niet representatief te zijn).
16. Verder dienen de auteurs goed te kijken in hoeverre de resultaten in lijn liggen met de onderzoeksvraag. De vergelijking tussen wel en niet gebruik van het instrument lijkt wat verder af te liggen van de onderzoeksvraag.
17. Wat betreft de discussie lijken deze niet gebaseerd te zijn op de resultaten. Een voorbeeld: "Daarnaast viel op dat professionals vooral gezins- en kindfactoren en relatief weinig ouderfactoren taxeerden in hun eigen klinisch oordeel van zorgbehoeften, terwijl zij significant meer gezins-, ouder-, en opvoedingsfactoren signaleerden wanneer gewerkt werd met..." Er is geen kwantitatief onderzoek gedaan.
18. Een kritische reflectie op de kwaliteit van het onderzoek ontbreekt.

Reactie reviewer 2 – 09-11-2022

Jeugdbescherming op maat: Een pilotonderzoek

Onderwerp

Het is een interessant en relevant onderwerp voor het tijdschrift en passend. Echter heb ik nog een aantal verzoeken tot aanpassing van het artikel/vragen aan de auteurs alvorens publicabel.

Inhoud: Het artikel bevat in principe de noodzakelijke theoretische en empirische onderbouwing. De auteurs hebben het terecht over de noodzaak van een gepersonaliseerde aanpak. Echter leidt de toepassing van het RNR-model niet per sé tot een gepersonaliseerde aanpak want inventarisatie van zorgbehoefte op basis van risicofactoren hoeft niet aan te sluiten bij wat gezinnen als behoeften zien. Verder wordt de zorgbehoefte gekoppeld aan bestaand aanbod. Dus dan is sprake van meer of minder geprotocolleerde interventies en is mij ook hierbij niet duidelijk in hoeverre recht wordt gedaan aan een gepersonaliseerde aanpak. Dat zou naar mijn idee nog explicieter benoemd mogen worden hier of in de discussie. In de discussie wordt pas in de laatste zin besproken hoe belangrijk het is gezinnen te betrekken bij de besluitvorming. Als we het over evidence based practice hebben dan gaat het om een gezamenlijk besluitvormingsproces waarbij evidentie uit wetenschap, de klinisch expertise van de hulpverlener (en ook deze wordt in het artikel als weinig relevant weggezet) en de wensen, behoeften, ervaringen, normen en waarden van kind/ouder/gezin een evenwichtige inbreng hebben. Dus kortom: enigszins terughoudendheid over het gebruik van gepersonaliseerd en een iets kritischer blik op de mogelijke opbrengsten van de tool in de discussie/conclusie lijken mij op z'n plaats.

Verder staan doelgroepfactoren in de inleiding onder dynamische factoren genoemd en dat is denk ik niet de bedoeling. Geef heir duidelijker aan dat deze factoren mee worden genomen maar niet dynamisch zijn.

Methode: De methode is naar mijn idee onvolledig. Bij de paragraaf procedure wordt alleen het interview maar niet het deel van het casus gedeelte met en zonder tool beschreven. Bij paragraaf data-analyse wordt alleen op het kwalitatieve deel van de studie ingegaan maar niet op de analyses voor wel/niet gebruik van de tool.

Resultaten: zijn correct weergegeven.

Discussie en conclusie: zie deels opmerkingen bij inleiding. Verder: uit de interviews over huidige werkwijze blijkt dat verschillende hulpverleners met vormen van gestructureerde behoeftetaxatie werken. In het onderzoek is slechts een vergelijking gemaakt tussen ongestructureerde (zonder tool) en gestructureerd (met tool) behoeftetaxatie. Het lijkt vrij logisch dat je met een gestructureerde taxatie meer risicofactoren vindt. Was het niet adequater/relevanter geweest om tevens de gestructureerde 'nieuwe' tool met reeds in de praktijk gebruikte gestructureerde tools te vergelijken. Want wat voegt jullie tool toe ten opzichte van bestaande tools – is het 'slechts' het zorgaanbod deel? Bovendien zou ook nog kritisch kunnen worden gekeken naar het zorgaanbod in relatie tot wetenschappelijk bewezen effectiviteit. Dan blijft een beperkt aantal interventies over. Ook dat mag kritisch benoemd worden.

Reactiebrief van auteur – 31-09-2023

Reviewer A:

Het onderzoek is inhoudelijk interessant en past qua onderwerp bij JiO. Het is prettig geschreven. Verder handelt het over een inhoudelijk relevant onderwerp, namelijk besluitvorming in de jeugdbescherming. Wel zijn er een aantal kritische kanttekeningen te plaatsen bij het artikel en dient er meer abstractieniveau aangebracht kunnen worden en consistentie in het artikel tussen doelstelling, theoretische inbedding en resultaten/discussie. Onderstaande opmerkingen kunnen helpen om het manuscript te verbeteren.

We bedanken de reviewer voor het erkennen van het belang van onze studie. We hebben elke opmerking zorgvuldig doorgenomen en suggesties verwerkt ter verbetering van ons manuscript.

1. In de inleiding mist theoretische inkadering van de factoren die besluitvormingsprocessen binnen jeugdbescherming beïnvloeden. Er zijn verschillende theoretische modellen die naast inhoudelijke factoren gerelateerd aan een casus ook andere factoren betrekken als eigenschappen van een professional of gedeelde besluitvorming. Waar focust dit onderzoek zich op als je inzoomt op factoren die van invloed zijn op besluitvorming?

Ons onderzoek zich richt op de evaluatie van een tool die professionals ondersteunt in diagnostiek van zorgbehoeften - het in kaart brengen van specifieke casusfactoren -, en niet op factoren van professionals, de organisatie, en/of andere (externe) factoren die besluitvorming van professionals beïnvloeden, zoals omschreven in het *decision-making ecology model* van Baumann et al. (2011). Om dit te verhelderen hebben we de volgende aanpassingen gedaan (p. 4/5):

... Het werken volgens het behoefteprincipe van het RNR-model sluit daarmee aan bij een holistische benadering binnen de jeugdbescherming door alle zorgbehoeften systematisch te inventariseren. Onder deze zorgbehoeften vallen specifieke casusfactoren, zoals de ontwikkelingsbehoeften van kinderen, de mogelijkheden van ouders en verzorgers om in deze ontwikkelingsbehoeften te voorzien, en familie- en omgevingsfactoren die mogelijk impact hebben op ouderlijke competenties en de ontwikkeling van kinderen (Spies et al., 2015; Wilson et al., 2008).'

... 'Deze theoretisch-empirische aanpak sluit aan bij het waarborgen van een gestructureerde analyse van beschikbare informatie in gezinssituaties ter bevordering van de kwaliteit van klinische besluitvorming door professionals in de jeugdbescherming (Helm & Roesch-Marsh, 2010; Lauritzen et al., 2018).'

Aanvullend bespreken we nu in de discussie kenmerken van professionals die mogelijk van invloed kunnen zijn op het besluitvormingsproces in de jeugdbescherming (zie punt 6).

2. Verder onduidelijk waarom een overzicht van Jeugdbescherming Regio Amsterdam met daarin de interventies en andere zorgvormen die worden toegepast bij cliënten is opgenomen. Waarom geen andere regio's? Waar draagt dit extra aan bij? Zegt dit iets over de interventies die in de database zijn opgenomen?

De tool was in eerste instantie ontwikkeld voor zorginstellingen in de regio Amsterdam. Dit hebben we nu specifiek vermeld (p. 6):

‘Belangrijk om te vermelden is dat (nog) niet alle Nederlandse interventies opgenomen zijn in [TOOL GEMASKEERD], omdat het instrument in eerste instantie is ontwikkeld voor zorginstellingen in de regio Amsterdam ([AUTEURS GEMASKEERD VOOR REVIEW]). Verreweg de meeste interventies worden ook in andere regio's in het land aangeboden, maar om ARIJ-Needs volledig bruikbaar te maken in heel Nederland kan de database van de tool worden uitgebreid met andere interventies die elders in het land worden aangeboden.’

3. Als ik het goed begrijp zou de analyse van verschillende databases met interventies geleid moeten hebben tot een overzicht van factoren/thema's waarop interventies aangrijpen. Is het zinvol om hier een overzicht van te geven? En in hoeverre sluit dit aan bij het overzicht wat eerder genoemd is (gezinsfactoren etc.)

Gezien de omvang van dat overzicht hebben wij dit niet in de huidige studie geïnccludeerd. Wij hebben wel verwezen naar een eerder rapport over de ontwikkeling van het instrument waarin een overzicht is opgenomen van alle interventies met interventiebeschrijvingen en contra-indicaties (p. 5):

‘(Zie [AUTEURS GEMASKEERD VOOR REVIEW] voor een overzicht van deze interventies met per interventie een beschrijving en contra-indicaties. Een volledig overzicht van de database is op te vragen bij de auteurs.)’

4. Verder zou de vraagstelling aangescherpt kunnen worden ten aanzien van de kennisvraag die erboven hangt. Aan welk kennistekort draagt dit onderzoek bij? Nu is de doelstelling gericht op: “Het doel van het huidige onderzoek was het verkrijgen van meer inzicht in (1) de wijze waarop het besluitvormingsproces over passende interventies plaatsvindt in de klinische jeugdbeschermingspraktijk, en (2) de praktische bruikbaarheid van [TOOL GEMASKEERD] ter ondersteuning van het besluitvormingsproces over passende interventies.” Met name de tweede doelstelling is het onduidelijk wat met praktische bruikbaarheid bedoeld wordt en welke theorie hieraan ten grondslag ligt. Hoe toets je of het praktisch bruikbaar is.

De aanleiding, het doel en de operationalisering van het onderzoek hebben we verduidelijkt in de eerste alinea's (p. 3/4):

‘Aangezien [TOOL GEMASKEERD] nog niet eerder is geïntroduceerd in de Nederlandse jeugdbeschermingspraktijk, was het doel van het huidige onderzoek om de praktische bruikbaarheid van de tool ter ondersteuning van besluitvorming over passende interventies voor gezinnen in de jeugdbescherming te onderzoeken. Daarvoor werd eerst inzicht verkregen in de huidige stand van zaken van het besluitvormingsproces over passende interventies in de klinische jeugdbeschermingspraktijk. Hiervoor werden semigestructureerde interviews bij professionals afgenomen. Daarnaast is in een vignet-studie onderzocht in welke mate besluitvormingsprocessen over passende interventies verschillen tussen geïnterviewde professionals die zonder en met [TOOL GEMASKEERD] te werk gaan. Aanvullend is een semigestructureerd interview afgenomen om gebruikerservaringen en relevante aanbevelingen voor doorontwikkeling van [TOOL GEMASKEERD] uit te vragen bij de professionals.’

5. Beschrijving van het design ontbreekt in de methode. Het betreft kwalitatief onderzoek en daarom is het goed om in het manuscript aan te geven hoe rekening is gehouden met de kwaliteitscriteria voor kwalitatief onderzoek (triangulatie, saturatie etc.)

We hebben de verantwoording voor het gebruikte design en kwaliteitscriteria toegelicht in de procedure (p. 7):

‘Dit onderzoek omvatte een mixed-methods-design waarbij interviews gecombineerd werden met kwantitatief onderzoek waarin een vergelijking werd gemaakt tussen het aantal getaxeerde behoeftefactoren zonder en met [TOOL GEMASKEERD].’

...

‘Aangezien alle professionals een zekere mate van expertise in de jeugdbescherming hadden, werd op basis van de consensustheorie – het principe dat experts op een specifiek gebied overeenkomstige bevindingen rapporteren – aangenomen dat een voldoende mate van saturatie bereikt kon worden met een relatief kleine sample (Guest et al., 2006). Na 15 interviews bleek dat de antwoorden van de professionals overeen kwamen en weinig nieuwe informatie naar voren kwam.’

6. Ik zou graag meer informatie willen over de samenstelling van de onderzoeksgroep. Wat zijn leeftijden, jaren werkervaring, scholingsachtergrond, ervaring met gebruik van de tool etc. Hier kan je ook een koppeling maken met factoren gelieerd aan professional die besluitvorming in jeugdbescherming beïnvloeden.

Wij hadden – ook met het oog op privacywetgeving (AVG) – enkel gegevens verzameld over het geslacht en functie binnen de jeugdbescherming van de geïnterviewde professionals. Wij hebben toegevoegd aan de tekst dat geen van de professionals bekend waren met de tool (participanten, p. 6). Wij zijn het eens met de reviewer dat kenmerken van professionals mogelijk van invloed zijn op het besluitvormingsproces in de jeugdbescherming. Daarom hebben we de volgende toelichting toegevoegd aan de discussie (p. 19):

‘Daarnaast waren geen kenmerken van de professionals gemeten die mogelijk van invloed zijn op het besluitvormingsproces in de jeugdbescherming. Zo zijn etnische achtergrond, aantal jaar werkervaring en diverse attitudes over overschrijdend opvoedingsgedrag mogelijk van invloed op de besluitvorming over passende ingrepen bij kindonveiligheid (Jent et al., 2011; Lauritzen et al., 2018; Lee et al., 2013). In vervolgonderzoek is het relevant om de eventuele invloed van dergelijke factoren op het gebruik van [TOOL GEMASKEERD] te toetsen.’

7. Het verdient de voorkeur om niet herleidbaar tot personen te gebruiken ipv geanonimiseerd.

We hebben dit aangepast (p. 7):

‘Er werd toestemming gevraagd voor het opnemen van het interview, met de expliciete vermelding dat de data vertrouwelijk en niet herleidbaar tot personen zouden worden verwerkt.’

8. De figuren zijn voor de reviewer niet zichtbaar waardoor het ingewikkeld is om inzicht te krijgen in het gebruik van de tool

We hebben een gemaskeerde schermafbeelding van de tool ingevoegd (p. 9).

9. Onderzoek lijkt te bestaan uit twee delen (ervaringen besluitvormingsprocedures en praktische bruikbaarheid) maar in de methode zie ik dat niet terug, evenals in de resultaten.

(Zie ons antwoord op het punt hieronder).

10. Onder procedure wordt het niet helder welke rol de casus speelt bij afname van het interview. Wordt er specifiek gekeken welke stappen zijn gezet? Of wordt na gebruik van de tool de ervaringen uitgevraagd?

Aansluitend op de opmerking van reviewer 2 hebben we de procedure verder toegelicht. Wij hebben de volgende alinea toegevoegd (p. 7):

'In het semigestructureerde interview werden allereerst vragen gesteld over het huidige besluitvormingsproces over passende behandeling in de jeugdbeschermingspraktijk. Vervolgens werd een fictieve casus voorgelegd aan de professional over de situatie van een multiprobleemgezin (Bijlage 1). Deze casus is in eerder onderzoek ontwikkeld en gebruikt ([AUTEURS GEMASKEERD VOOR REVIEW]). Uit onderzoek blijkt dat een zogenoemde 'vignetstudie' een geschikte methode is voor het vergelijken van besluitvormingsprocessen binnen diverse groepen sociaal maatschappelijk werkers (Lauritzen et al., 2018). Er werd aan de professionals gevraagd of zij alle behoeftefactoren van het gezin in de casus konden opschrijven, en vervolgens welke specifieke interventie(s) zij zouden aanbevelen die zij passend achtten bij de gesignaleerde behoeftefactoren. Vervolgens werd de [TOOL GEMASKEERD] geïntroduceerd waarbij het gebruik ervan werd toegelicht door de interviewer. Er werd aan de professionals nogmaals gevraagd om alle behoeftefactoren van het gezin in de casus te taxeren, maar dan nu met de tool. Ten slotte werden vragen gesteld over de gebruikservaringen met de tool en een evaluatie van de uitkomsten van de tool.'

11. Onduidelijk blijft hoe er precies gecodeerd is. Is er sprake geweest van deductief coderen? Kijkend naar de structuur van de resultaten lijkt de structuur van het interviewprotocol gehanteerd te zijn. Vergelijking tussen wel en niet gebruik van de tool lijkt niet kwalitatief bekeken te zijn maar kwantitatief. Onduidelijk is hoe deze data verzameld is. Ook vraag ik me af wat de waarde van deze uitkomsten is gezien de toch kleine onderzoeksgroep. Ook weten we weinig over de samenstelling van de onderzoeksgroep daarmee wordt de vraag naar representativiteit opgeroepen.

Zie ons antwoord op het punt hieronder.

12. In de data-analyse wordt geen aandacht besteed aan de betrouwbaarheid van de analyse. Kan daar informatie over toegevoegd worden?

Wij hebben het coderingsproces zoals beschreven bij de data-analyse (p. 9/10) verder toegelicht:

'In het codeerproces zijn zowel inductieve als deductieve methodes gebruikt om de data te coderen (Green & Thorogood 2018). Via deductieve analyse werden codes toegekend aan de hand van de interviewopzet. Per vragenlijstonderdeel werd een inductieve onderzoeksmethode toegepast, zodat eventuele assumpties of referentiekaders van de auteurs geen invloed hadden op het definitieve codeerschema (Green & Thorogood, 2018). De eerste auteur codeerde alle interviews waarbij onduidelijkheden werden gemarkeerd. Deze onduidelijkheden werden middels discussie met de coauteurs besproken tot volledige overeenstemming was bereikt over het codeerschema inclusief codegroepen en thema's. Aanvullend zijn t-toetsen voor onafhankelijke groepen uitgevoerd om verschillen in het aantal getaxeerde behoeftefactoren tussen de conditie zonder en met [TOOL GEMASKEERD] te toetsen. Deze toetsen waren uitgevoerd per factorcategorie.'

13. Het instrument wordt toegelicht er lijkt geen onderzoek te zijn gedaan naar de betrouwbaarheid en validiteit. Dit onderzoek lijkt een eerste onderzoek die hier een bijdrage aan moet leveren (?) en is iets terug te lezen is over inhoudsvaliditeit en ecologische validiteit?

Aangezien dit een eerste exploratief onderzoek was naar de bruikbaarheid van de tool, is gebruikgemaakt van een fictieve casus in een onderzoeksetting. In de discussie hebben wij het belang van het verder onderzoeken van de psychometrische kwaliteit van de tool (in combinatie met andere risicotaxatie-instrumenten) onderstreept (p. 19):

'Ten slotte is het belangrijk om naast gebruikerservaringen ook de psychometrische kwaliteit van [TOOL GEMASKEERD] te onderzoeken. Hoewel uit eerder onderzoek blijkt dat toepassing van instrumenten die gebaseerd zijn op de RNR-principes het besluitvormingsproces in de jeugdbescherming verbeteren ([AUTEURS GEMASKEERD VOOR REVIEW]), blijft een urgente vraag of implementatie van [TOOL GEMASKEERD] werkelijk leidt tot betere besluitvorming in de praktijk en dus tot minder kindermishandeling.'

14. In de resultaten komen veel verschillende termen/interventies voor die voor niet-ingewijden moeilijk te volgen zijn. Verder worden de resultaten er operationeel (ctrl-toets) en onduidelijk hoe dit bijdraagt aan het verkrijgen inzicht in...zie doelstelling.

Wij hebben de interventies toegelicht (p. 10):

'De participanten noemden verschillende instrumenten/methodieken in hun antwoord op de vraag hoe zij doorgaans een gestructureerde behoeftetaxatie uitvoeren, namelijk: (1) ARIJ-Risk, RiS 2.0 en SIRE-lijst (risicotaxatieinstrumenten), (2) de TOP-3 methodiek (redeneer- en beslismodel over onveiligheid in huiselijke kring), (3) het meten van de Big Five middels een persoonlijkheidsvragenlijst, (4) het Dialoog-model (een methode gericht op het bevorderen van afstemming, overeenstemming en samenwerking tussen hulpvragers en verleners), (5) Richtlijnen Jeugdhulp (globaal stappenplan), en (6) Ambulante Spoedhulp.'

Op grond van de resultaten zijn operationele aanbevelingen geformuleerd voor de doorontwikkeling van de tool en bevordering van het gebruiksgemak. Wij hebben het kopje aangepast om dit te verduidelijken (*Gebruikerservaringen en Suggesties voor Bevordering van Gebruiksgemak*, p. 15).

15. Er worden aantallen genoemd hoe vaak een professional iets weergeeft, bij kwalitatief onderzoek gaat het vooral om de ervaringen die naar voren gebracht worden en niet de aantallen (daarom hoeft de onderzoeksgroep niet representatief te zijn).

Wij hebben de aantallen steeds genoemd met het oog op het voorkomen van ambigue stellingen als *'de meeste' of 'enkele' professionals gaven aan dat ... etc.*, en ter aanduiding van een ordening in relevantie van stellingen. Mocht de editor ook van mening zijn dat dit overbodige informatie is, dan passen wij dit graag aan.

16. Verder dienen de auteurs goed te kijken in hoeverre de resultaten in lijn liggen met de onderzoeksvraag. De vergelijking tussen wel en niet gebruik van het instrument lijkt wat verder af te liggen van de onderzoeksvraag.

Door de aangescherpte vraagstelling (punt 4) en meer uitgebreide toelichting van de procedure (punt 10) en het design (punt 5) hebben we de verbinding tussen de resultaten en de onderzoeksvraag versterkt.

17. Wat betreft de discussie lijken deze niet gebaseerd te zijn op de resultaten. Een voorbeeld: "Daarnaast viel op dat professionals vooral gezins- en kindfactoren en relatief weinig ouderfactoren taxeerden in hun eigen klinisch oordeel van zorgbehoeften, terwijl zij significant meer gezins-, ouder-, en opvoedingsfactoren signaleerden wanneer gewerkt werd met..." Er is geen kwantitatief onderzoek gedaan.

De toelichting van het onderzoeksdesign (punt 5) en de data-analyse (punt 12) zijn aangescherpt waarmee de discussie in lijn is gebracht met de resultatensectie.

18. Een kritische reflectie op de kwaliteit van het onderzoek ontbreekt.

Naar aanleiding van de aanpassing bij punt 6 en de eerste suggestie van reviewer 2, hebben we een kritische reflectie op het onderzoek toegevoegd aan de discussie (p. 18/19):

'Een belangrijke kanttekening hierbij is dat [TOOL GEMASKEERD] interventiemogelijkheden biedt die meestal in geprotocolleerde vorm worden aangeboden. Hoewel deze interventies algemeen werkzame factoren kunnen bevatten die een interventie effectief maken, ongeacht het doel, de doelgroep en het soort interventies, kunnen deze interventies essentiële specifiek werkzame elementen missen die een interventie effectiever maken voor gezinnen waarin sprake is van kindonveiligheid (van Yperen et al., 2015). Hierbij is belangrijk te vermelden dat voor het terugdringen van kindermishandeling tot op heden nog geen specifiek werkzame elementen van interventies zijn gevonden die eventueel gekoppeld kunnen worden aan behoeftefactoren van de doelgroep ([AUTEURS GEMASKEERD VOOR REVIEW]). Daarnaast is ten tijde van dit schrijven slechts 1,82% van de interventies in de NJI-databank bewezen effectief volgens sterke aanwijzingen. Dit maakt dat de interventieaanbevelingen door [TOOL GEMASKEERD] kritisch geëvalueerd dienen te worden door de gebruikers, waarbij de geschiktheid van een interventie op papier afgewogen moet worden tegen de daadwerkelijke aangetoonde effectiviteit.'

Aanbeveling: Revisies benodigd

 Reviewer B:

Jeugdbescherming op maat: Een pilotonderzoek

Onderwerp

Het is een interessant en relevant onderwerp voor het tijdschrift en passend. Echter heb ik nog een aantal verzoeken tot aanpassing van het artikel/vragen aan de auteurs alvorens publicabel.

We bedanken de reviewer voor het erkennen van de relevantie van onze studie voor JiO. We hebben elke opmerking zorgvuldig doorgenomen en aanpassingssuggesties verwerkt ter verbetering van ons manuscript.

Inhoud: Het artikel bevat in principe de noodzakelijke theoretische en empirische onderbouwing. De auteurs hebben het terecht over de noodzaak van een gepersonaliseerde aanpak. Echter leidt de toepassing van het RNR-model niet per sé tot een gepersonaliseerde aanpak want inventarisatie van zorgbehoefte op basis van risicofactoren hoeft niet aan te sluiten bij wat gezinnen als behoeften zien. Verder wordt de zorgbehoefte gekoppeld aan bestaand aanbod. Dus dan is sprake van meer of minder geprotocolleerde interventies en is mij ook hierbij niet duidelijk in hoeverre recht wordt gedaan aan een gepersonaliseerde aanpak. Dat zou naar mijn idee nog explicieter benoemd mogen worden hier of in de discussie.

We zijn het met de reviewer eens dat een optimale vorm van gepersonaliseerde behandeling gebaseerd is op het laten aansluiten van interventies op (1) door een professional getaxeerde zorgbehoeften, en (2) factoren die door gezinnen zelf als zorgbehoeften worden beschouwd. Wij vinden ook dat sterk geprotocolleerde interventies een gepersonaliseerde aanpak kunnen bemoeilijken. Vanuit dit perspectief kan de door ons geëvalueerde tool niet volledig tegemoet komen aan personalisering van zorg, en dit bediscussiëren wij nu op pagina 18:

‘Een belangrijke kanttekening hierbij is dat [TOOL GEMASKEERD] interventiemogelijkheden biedt die meestal in geprotocolleerde vorm worden aangeboden. Hoewel deze interventies algemeen werkzame factoren kunnen bevatten die een interventie effectief maken, ongeacht het doel, de doelgroep en het soort interventies, missen deze interventies mogelijk essentiële specifiek werkzame elementen die aansluiten bij factoren die door gezinnen zelf als zorgbehoeften worden beschouwd (van Yperen et al., 2015). Hierbij is het wel belangrijk te vermelden dat voor het terugdringen van kindermishandeling er tot op heden nog geen specifiek werkzame elementen zijn gevonden binnen interventies die eventueel losstaand gekoppeld kunnen worden aan behoeftefactoren van de doelgroep ([AUTEURS GEMASKEERD VOOR REVIEW]).’

In de discussie wordt pas in de laatste zin besproken hoe belangrijk het is gezinnen te betrekken bij de besluitvorming. Als we het over evidence based practice hebben dan gaat het om een gezamenlijk besluitvormingsproces waarbij evidentie uit wetenschap, de klinisch expertise van de hulpverlener (en ook deze wordt in het artikel als weinig relevant weggezet) en de wensen, behoeften, ervaringen, normen en waarden van kind/ouder/gezin een evenwichtige inbreng hebben. Dus kortom: enigszins terughoudendheid over het gebruik van

gepersonaliseerd en een iets kritischer blik op de mogelijke opbrengsten van de tool in de discussie/conclusie lijken mij op z'n plaats.

Wij vinden ook dat terughoudendheid passend is en hebben onze tekst hierop aangepast. Zie ons antwoord hierboven.

Verder staan doelgroepfactoren in de inleiding onder dynamische factoren genoemd en dat is denk ik niet de bedoeling. Geef heir duidelijker aan dat deze factoren mee worden genomen maar niet dynamisch zijn.

Wij hebben dit inderdaad over het hoofd gezien. Dit is nu aangepast (p. 5):

'Daarnaast is er een overzicht van doelgroepfactoren (statische factoren, waaronder een cognitieve beperking), als mogelijke behoeftefactoren voor een passende interventie.'

Methode: De methode is naar mijn idee onvolledig. Bij de paragraaf procedure wordt alleen het interview maar niet het deel van het casus gedeelte met en zonder tool beschreven.

Wij zijn het met de reviewer eens dat de procedure verder toegelicht kon worden. Wij hebben daarom de volgende alinea toegevoegd (p. 7):

'In het semigestructureerde interview werden allereerst vragen gesteld over het huidige besluitvormingsproces over passende behandeling in de jeugdbeschermingspraktijk. Vervolgens werd een fictieve casus voorgelegd aan de professional over de situatie van een multiprobleemgezin (Bijlage 1). Deze casus is in eerder onderzoek ontwikkeld en gebruikt ([AUTEURS GEMASKEERD VOOR REVIEW]). Uit onderzoek blijkt dat een zogenoemde 'vignetstudie' een geschikte methode is voor het vergelijken van besluitvormingsprocessen binnen diverse groepen sociaal maatschappelijk werkers (Lauritzen et al., 2018). Er werd aan de professionals gevraagd of zij alle behoeftefactoren van het gezin in de casus konden opschrijven, en vervolgens welke specifieke interventie(s) zij zouden aanbevelen die zij passend achtten bij de gesignaleerde behoeftefactoren. Vervolgens werd de [TOOL GEMASKEERD] geïntroduceerd waarbij het gebruik ervan werd toegelicht door de interviewer. Er werd aan de professionals nogmaals gevraagd om alle behoeftefactoren van het gezin in de casus te taxeren, maar dan nu met de tool. Ten slotte werden vragen gesteld over de gebruikservaringen met de tool en een evaluatie van de uitkomsten van de tool.'

Bij paragraaf data-analyse wordt alleen op het kwalitatieve deel van de studie ingegaan maar niet op de analyses voor wel/niet gebruik van de tool.

Wij hebben dit nu aangevuld (p. 10):

'Aanvullend zijn t-toetsen voor onafhankelijke groepen uitgevoerd om verschillen in het aantal getaxeerde behoeftefactoren tussen de conditie zonder en met [TOOL GEMASKEERD] te toetsen. Deze toetsen waren uitgevoerd per factorcategorie.'

Resultaten: zijn correct weergegeven.

Discussie en conclusie: zie deels opmerkingen bij inleiding. Verder: uit de interviews over huidige werkwijze blijkt dat verschillende hulpverleners met vormen van gestructureerde

behoefte-taxatie werken. In het onderzoek is slechts een vergelijking gemaakt tussen ongestructureerde (zonder tool) en gestructureerd (met tool) behoefte-taxatie. Het lijkt vrij logisch dat je met een gestructureerde taxatie meer risicofactoren vindt. Was het niet adequater/relevanter geweest om tevens de gestructureerde ‘nieuwe’ tool met reeds in de praktijk gebruikte gestructureerde tools te vergelijken. Want wat voegt jullie tool toe ten opzichte van bestaande tools – is het ‘slechts’ het zorgaanbod deel?

Met een gestructureerde taxatie kan inderdaad verwacht worden dat meer casusfactoren gevonden worden, mits een professional het juiste instrument gebruikt en niet bijvoorbeeld een risicotaxatie-instrument voor behoefte-taxatie. Uit de interviews bleek dat professionals bekend waren met reeds bestaande risicotaxatie-instrumenten, maar nog niet met behoefte-taxatie-instrumenten. Wij hebben daarom dit punt concreter toegelicht in de discussie (p. 16):

‘Verder bleken de geïnterviewde professionals risicotaxatie-instrumenten (o.a., ARIJ-Risk, RiS 2.0 en SIRE-lijst) te gebruiken in het besluitvormingsproces over passende interventies in de jeugdbescherming, in plaats van behoefte-taxatie-instrumenten. Zoals beschreven in het visiedocument “Gefaseerde ketensamenwerking bij kindermishandeling, huiselijk geweld en seksueel misbruik” (Vogtländer & van Arum, 2016), is het belangrijk om onderscheid te maken tussen het werken aan kindveiligheid (1), risicogestuurde (2)- en herstelgerichte zorg (3). Zoals eerder genoemd maakt risicotaxatie inzichtelijk wie hulpverlening nodig heeft (gezinnen met een verhoogd risico) en met welke intensiteit deze hulpverlening aangeboden moet worden (risicogestuurde zorg), terwijl behoefte-taxatie zichtbaar maakt wat aangepakt moet worden met welke interventies (herstelgerichte zorg) ([AUTEURS GEMASKEERD VOOR REVIEW]). [TOOL GEMASKEERD] is daarmee een aanvulling op het reeds bestaande aanbod van instrumenten voor het opstellen van een systeemgericht hulpverleningsplan.’

Bovendien zou ook nog kritisch kunnen worden gekeken naar het zorgaanbod in relatie tot wetenschappelijk bewezen effectiviteit. Dan blijft een beperkt aantal interventies over. Ook dat mag kritisch benoemd worden.

Wij zijn het met de reviewer eens dat de huidige aanbod van effectieve interventies ontoereikend is. Dit hebben wij nu opgenomen in de discussie (p. 18/19):

‘Daarnaast is ten tijde van dit schrijven slechts 1,82% van de interventies in de NJI-databank bewezen effectief volgens sterke aanwijzingen. Dit maakt dat de interventieaanbevelingen door [TOOL GEMASKEERD] kritisch geëvalueerd dienen te worden door de gebruikers, waarbij de geschiktheid van een interventie op papier afgewogen moet worden tegen de daadwerkelijke aangetoonde effectiviteit.’

Aanbeveling: Opnieuw indienen voor review

Tweede versie van het manuscript – 31-09-2023

Jeugdbescherming op maat: Een pilotonderzoek naar [NAAM VAN TOOL GEMASKEERD VOOR REVIEW]

Samenvatting

Recentelijk is [NAAM VAN TOOL GEMASKEERD VOOR REVIEW, VERVANGEN DOOR ‘TOOL GEMASKEERD’ IN MANUSCRIPT] ontwikkeld om professionals te ondersteunen in zowel behoeftetaxatie als besluitvorming over passende zorg voor gezinnen met risico’s op (herhaling van) kindermishandeling. De huidige studie betrof een kwalitatief onderzoek naar de klinische waarde en bruikbaarheid van [TOOL GEMASKEERD]. Hiervoor werden professionals ($N = 15$) geïnterviewd waarbij een fictieve gezinscasus werd voorgelegd om verschillen te onderzoeken tussen ongestructureerde klinische behoeftetaxatie zonder gebruik van [TOOL GEMASKEERD] en gestructureerde behoeftetaxatie met gebruik van [TOOL GEMASKEERD]. Volgens de resultaten signaleren professionals significant meer zorgbehoeften als zij [TOOL GEMASKEERD] gebruiken dan wanneer zij zorgbehoeften taxeren volgens hun klinisch en ongestructureerde oordeel zonder gebruik van [TOOL GEMASKEERD]. Ook blijken professionals meer opvoedings- en ouderfactoren te signaleren bij gebruik van [TOOL GEMASKEERD], wat sterke voorspellers zijn voor kindermishandeling en daarmee belangrijke aanknopingspunten voor behandeling. De zorgkeuzemodule van [TOOL GEMASKEERD] werd beoordeeld als een ‘verbreding van de horizon’ in het voortdurend groeiende aanbod van zorg en interventies in de jeugdbeschermingspraktijk. In conclusie ondersteunen de resultaten in deze studie de klinische waarde en bruikbaarheid van [TOOL GEMASKEERD].

Abstract

The assessment instrument “[TOOL GEMASKEERD]” has recently been developed to support practitioners in assessing treatment needs and selecting appropriate interventions for families at risk for child maltreatment. This qualitative study examined the clinical value and usability of the [TOOL GEMASKEERD] assessment instrument. Practitioners ($N = 15$) were interviewed in which a case vignette was used to examine differences between unstructured clinical needs assessment without [TOOL GEMASKEERD] and structured needs assessment with [TOOL GEMASKEERD]. The results showed that practitioners identify significantly more treatment needs when they use [TOOL GEMASKEERD] than when they assess health

care needs according to their clinical and unstructured judgment. Professionals specifically identified more parent- and parenting factors when they use [TOOL GEMASKEERD], which are strong predictors of child abuse and therefore important treatment targets. The results also showed that [TOOL GEMASKEERD] supports practitioners in selecting interventions from a vastly growing range of child welfare interventions that directly target the assessed factors. In conclusion, this study supports the clinical value and usability of the [TOOL GEMASKEERD] assessment instrument.

Trefwoorden: *jeugdbescherming, [TOOL GEMASKEERD], zorgbehoefte, interventiekeuze*

Aantal woorden: 6302

1. Inleiding

1.1. Aanleiding van het Onderzoek

Kindermishandeling is een omvangrijk probleem met ernstige gevolgen voor zowel slachtoffers als de maatschappij (Stoltenborgh e.a., 2015). Een effectieve aanpak gericht op het voorkomen van mishandeling is dan ook essentieel. Hoewel de preventie van kindermishandeling steeds meer aandacht krijgt, laten diverse meta-analyses zien dat preventieve interventies ter voorkoming van kindermishandeling slechts in beperkte mate effectief zijn (e.g., Euser e.a., 2015; [AUTEURS GEMASKEERD VOOR REVIEW]). Een mogelijke oorzaak van deze matige effectiviteit is dat interventies niet of onvoldoende zijn toegespitst op zorgbehoeften en kenmerken die per risicogezin kunnen verschillen (Ng & Weisz, 2016; Weisz, 2014), terwijl het personaliseren van behandeling de effectiviteit van interventies juist kan vergroten, zoals onder meer het bewezen effectieve *Risk-Need-Responsivity*-model beschrijft (Andrews e.a., 1990; Andrews e.a., 2011; Bonta and Andrews, 2007; Bonta and Andrews, 2016). Dit model is ontwikkeld voor het effectief inrichten van strafrechtelijke zorg en schrijft in een aantal principes voor hoe interventies vormgegeven moeten worden om de effectiviteit van zorg te vergroten. In diverse studies is inmiddels betoogd dat dit model ook toepasbaar is in jeugdbescherming en dat het werken volgens de RNR-principes de effectiviteit van preventieve interventies voor kindermishandeling kan vergroten ([AUTEURS GEMASKEERD VOOR REVIEW]). Om de RNR-principes toepasbaar te maken in de context van jeugdbescherming zijn recentelijk de [TOOL GEMASKEERD] ([AUTEURS GEMASKEERD VOOR REVIEW]) en [TOOL GEMASKEERD] ([AUTEURS GEMASKEERD VOOR REVIEW]) ontwikkeld. [TOOL GEMASKEERD] heeft als doel professionals te ondersteunen in het taxeren van zorgbehoeften en het kiezen van passende zorg voor risicogezinnen in de jeugdbescherming ([AUTEURS GEMASKEERD VOOR REVIEW]).

Aangezien [TOOL GEMASKEERD] nog niet eerder is geïntroduceerd in de Nederlandse jeugdbeschermingspraktijk, was het doel van het huidige onderzoek om de praktische bruikbaarheid van de tool ter ondersteuning van besluitvorming over passende interventies voor gezinnen in de jeugdbescherming te onderzoeken. Daarvoor werd eerst inzicht verkregen in de huidige stand van zaken van het besluitvormingsproces over passende interventies in de klinische jeugdbeschermingspraktijk. Hiervoor werden semigestructureerde interviews bij professionals afgenomen. Daarnaast is in een vignet-studie onderzocht in welke mate besluitvormingsprocessen over passende interventies verschillen tussen geïnterviewde

professionals die zonder en met [TOOL GEMASKEERD] te werk gaan. Aanvullend is een semigestructureerd interview afgenomen om gebruikerservaringen en relevante aanbevelingen voor doorontwikkeling van [TOOL GEMASKEERD] uit te vragen bij de professionals.

1.2. RNR-Principes

In strafrechtelijke zorg worden de Risk-Need-Responsivity principes veelvuldig toegepast om recidive op effectieve wijze te voorkomen (Andrews e.a., 1990). Het RNR-model beschrijft op basis van drie principes hoe behandelingen gepersonaliseerd moeten worden voor individuele cliënten om de behandel-effectiviteit te vergroten. Zo stelt het risicoprincipe dat de intensiteit van een interventie moet aansluiten bij het recidiverisico en het behoefteprincipe dat een interventie gericht moet zijn op criminogene behoeften (i.e. de veranderbare risicofactoren die samenhangen met recidive). Tot slot schrijft het responsiviteitsprincipe voor dat interventies moeten aansluiten bij onder andere het cognitieve niveau, de motivatie en demografische eigenschappen (waaronder geslacht en etnische afkomst) van de cliënt. Interventies die zijn vormgegeven volgens deze principes blijken effectiever in het voorkomen van recidive dan interventies die niet voldoen aan deze principes (Andrews e.a., 1990; Dowden & Andrews, 1996; 2003; 2006; Hanson e.a., 2009).

Hoewel het RNR-model is ontwikkeld ter voorkoming van criminele recidive, is het model ook veelbelovend voor de preventie van (herhaling van) kindermishandeling ([AUTEURS GEMASKEERD VOOR REVIEW]). Zo kunnen zowel delinquent gedrag als kindermishandeling worden verklaard door de balans tussen risicofactoren (denk aan armoede, stress en psychopathologie) en beschermende factoren (zoals sociale steun) in verschillende sociale systemen rondom een cliënt of risicogezin, zoals de familie, een school, en de buurt waarin een cliënt of gezin woont (Belsky, 1993; [AUTEURS GEMASKEERD VOOR REVIEW]). Het werken volgens het behoefteprincipe van het RNR-model sluit daarmee aan bij een holistische benadering binnen de jeugdbescherming door alle zorgbehoeften systematisch te inventariseren. Onder deze zorgbehoeften vallen specifieke casusfactoren, zoals de ontwikkelingsbehoeften van kinderen, de mogelijkheden van ouders en verzorgers om in deze ontwikkelingsbehoeften te voorzien, en familie- en omgevingsfactoren die mogelijk impact hebben op ouderlijke competenties en de ontwikkeling van kinderen (Spies et al., 2015; Wilson et al., 2008).

Desondanks worden de RNR-principes tot op heden in beperkte mate toegepast in de jeugdbescherming. Om beter te kunnen werken volgens de RNR-principes is onlangs [TOOL GEMASKEERD] ontwikkeld ([AUTEURS GEMASKEERD VOOR REVIEW]), waarmee zorgbehoeften van risicogezinnen op gestructureerde wijze geïnventariseerd kunnen worden

(i.e., de veranderbare risicofactoren die samenhangen met kindermishandeling). Daarnaast wordt er met deze tool een koppeling gemaakt tussen getaxeerde zorgbehoeften en het beschikbare zorgaanbod, waarmee [TOOL GEMASKEERD] de toepassing van het behoefteprincipe van het RNR-model faciliteert.

1.3. [TOOL GEMASKEERD]

Professionals kunnen met [TOOL GEMASKEERD] inventariseren welke dynamische (veranderbare) risicofactoren voor kindermishandeling, oftewel zorgbehoeften, aanwezig zijn in een gezin en dus aanknopingspunten bieden voor behandeling. In dit onderdeel voor behoeftetaxatie gaat het om factoren die kindermishandeling veroorzaken en/of in stand houden ([AUTEURS GEMASKEERD VOOR REVIEW]). Voor de ontwikkeling van het behoeftetaxatie-onderdeel van [TOOL GEMASKEERD] werden zulke dynamische risicofactoren geselecteerd op basis van (overzicht)studies naar significante voorspellers van kindermishandeling (o.a., [AUTEURS GEMASKEERD VOOR REVIEW]; Cash, 2001; Hindley e.a., 2006; [AUTEURS GEMASKEERD VOOR REVIEW]; Stith e.a., 2009). Deze theoretisch-empirische aanpak sluit aan bij het waarborgen van een gestructureerde analyse van beschikbare informatie in gezinssituaties ter bevordering van de kwaliteit van klinische besluitvorming door professionals in de jeugdbescherming (Helm & Roesch-Marsh, 2010; Lauritzen et al., 2018). De geselecteerde dynamische risicofactoren werden vervolgens onderverdeeld in *opvoedingsfactoren* (o.a., problemen in de ouder-kind interactie, onvoldoende bescherming/toezicht/veiligheid), *gezinsfactoren* (o.a., weinig sociale steun, financiële problemen) en *ouderfactoren* (o.a., opvoeder ervaart stress, problemen met agressie-regulatie), *kindfactoren* (o.a., internaliserende problematiek, leerproblemen). Daarnaast is er een overzicht van doelgroepfactoren (statische factoren, waaronder een cognitieve beperking), als mogelijke behoeftefactoren voor een passende interventie.

Voor de ontwikkeling van de module om professionals te ondersteunen in het kiezen van passende zorg en/of interventies is een database gemaakt van in totaal 116 interventies voor (de preventie van) kindermishandeling die in Nederland beschikbaar zijn (Zie [AUTEURS GEMASKEERD VOOR REVIEW] voor een overzicht van deze interventies met per interventie een beschrijving en contra-indicaties. Een volledig overzicht van de database is op te vragen bij de auteurs). Elk van deze interventies grijpt tenminste aan op één risicofactor voor kindermishandeling. Om interventies te identificeren zijn meerdere bronnen geraadpleegd, namelijk de databank Effectieve Jeugdinterventies (Nederlands Jeugdinstituut), de databank Effectieve interventies huiselijk geweld en seksueel geweld (Ministerie van

Volksgezondheid, Welzijn en Sport & Movisie) en een overzicht van Jeugdbescherming Regio Amsterdam met daarin de interventies en andere zorgvormen die worden toegepast bij cliënten. **Belangrijk om te vermelden is dat (nog) niet alle Nederlandse interventies opgenomen zijn in [TOOL GEMASKEERD], omdat het instrument in eerste instantie is ontwikkeld voor zorginstellingen in de regio Amsterdam ([AUTEURS GEMASKEERD VOOR REVIEW]).** Verreweg de meeste interventies worden ook in andere regio's in het land aangeboden, maar om ARIJ-Needs volledig bruikbaar te maken in heel Nederland kan de database van de tool worden uitgebreid met andere interventies die elders in het land worden aangeboden. Van elke geïdentificeerde interventie werd de handleiding of het protocol gescreend op de dynamische factoren waarop de interventie aangrijpt. Zo was het mogelijk de interventies die werden opgenomen in de database van [TOOL GEMASKEERD] te koppelen aan de veranderbare risicofactoren (ofwel de mogelijke zorgbehoeften) die gemeten worden met het behoeftetaxatie-onderdeel van [TOOL GEMASKEERD] ([AUTEURS GEMASKEERD VOOR REVIEW]). De verwachting is dat [TOOL GEMASKEERD] de besluitvorming over het inzetten van passende hulpverlening in de jeugdbescherming effectiever, efficiënter en minder subjectief maakt, omdat deze tool het mogelijk maakt om het behoeftepincipe van het RNR model op systematische wijze in de klinische praktijk van jeugdbescherming toe te passen.

2. Methode

2.1. Participanten

In totaal werden bij 15 professionals uit de jeugdzorg, waarvan 12 vrouwen en 3 mannen, semigestructureerde interviews afgenomen. Deze professionals hadden de volgende functies: gezinscoach, medewerker Crisis Interventie Team, klinisch/GZ-(kinder- en jeugd)psycholoog, jeugdmaatschappelijk werker, gedragswetenschapper, schoolmaatschappelijk werker en jeugdbeschermer. Zij waren werkzaam bij GGD Hollands-Midden, Opvoedpoli, Levvel, Bureau Jeugdzorg (Limburg) en Bascule. **De professionals waren niet bekend met het gebruik van [TOOL GEMASKEERD].**

2.2. Procedure

De professionals werden geworven via de instellingen die betrokken waren bij het [PROJECT GEMASKEERD VOOR REVIEW] en via het netwerk van de onderzoekers in de periode september 2020 Tot en met januari 2021. De participanten ontvingen informatie over het doel

en de procedure van het onderzoek, waarna zij een informed-consent-formulier ondertekenden. De interviews duurden gemiddeld 37 minuten en werden op de werkplek van de participant of via een videoverbinding afgenomen. Er werd toestemming gevraagd voor het opnemen van het interview, met de expliciete vermelding dat de data vertrouwelijk en niet herleidbaar tot personen zouden worden verwerkt.

Dit onderzoek omvatte een mixed-methods-design waarbij interviews gecombineerd werden met kwantitatief onderzoek waarin een vergelijking werd gemaakt tussen het aantal getaxeerde behoeftefactoren zonder en met [TOOL GEMASKEERD]. In het semigestructureerde interview werden allereerst vragen gesteld over het huidige besluitvormingsproces over passende behandeling in de jeugdbeschermingspraktijk. Vervolgens werd een fictieve casus voorgelegd aan de professional over de situatie van een multiprobleemgezin (Bijlage 1). Deze casus is in eerder onderzoek ontwikkeld en gebruikt ([AUTEURS GEMASKEERD VOOR REVIEW]). Uit onderzoek blijkt dat een zogenoemde ‘vignetstudie’ een geschikte methode is voor het vergelijken van besluitvormingsprocessen binnen diverse groepen sociaal maatschappelijk werkers (Lauritzen et al., 2018). Er werd aan de professionals gevraagd of zij alle behoeftefactoren van het gezin in de casus konden opschrijven, en vervolgens welke specifieke interventie(s) zij zouden aanbevelen die zij passend achtten bij de gesignaleerde behoeftefactoren. Vervolgens werd de [TOOL GEMASKEERD] geïntroduceerd waarbij het gebruik ervan werd toegelicht door de interviewer. Er werd aan de professionals nogmaals gevraagd om alle behoeftefactoren van het gezin in de casus te taxeren, maar dan nu met de tool. Ten slotte werden vragen gesteld over de gebruikservaringen met de tool en een evaluatie van de uitkomsten van de tool.

Aangezien alle professionals een zekere mate van expertise in de jeugdbescherming hadden, werd op basis van de consensustheorie – het principe dat experts op een specifiek gebied overeenkomstige bevindingen rapporteren – aangenomen dat een voldoende mate van saturatie bereikt kon worden met een relatief kleine sample (Guest et al., 2006). Na 15 interviews bleek dat de antwoorden van de professionals overeen kwamen en weinig nieuwe informatie naar voren kwam.

2.3. Instrumenten

2.3.1 Interviews

In de semigestructureerde interviews werd een vooraf opgestelde vragenlijst gebruikt. De interviews bestonden uit de volgende onderdelen:

1. De Jeugdbeschermingspraktijk zonder [TOOL GEMASKEERD]

- 1.1. Vragen over de huidige behoeftetaxatie-procedure en het besluitvormingsproces over passende zorg in de jeugdbeschermingspraktijk;
- 1.2. Vragen over eventuele knelpunten in de toeleiding naar passende zorg;
- 1.3. Vragen over mogelijke oplossingen ter verbetering van de zorg;
- 1.4. Taxatie van zorgbehoeften bij een fictieve gezinscasus zonder [TOOL GEMASKEERD];
- 1.5. Vragen over welke zorg past bij de casus;

2. *Behoeftetaxatie en keuze passende zorg met [TOOL GEMASKEERD]*

- 2.1. Introductie van [TOOL GEMASKEERD]: taxatie van de zorgbehoeften in de casus
- 2.2. Evaluatie van de taxatieresultaten met [TOOL GEMASKEERD] en vragen over de gebruikerservaringen met [TOOL GEMASKEERD].

2.3.2 [TOOL GEMASKEERD]

[TOOL GEMASKEERD] is een computerapplicatie die is ontwikkeld om zorgprofessionals te ondersteunen in enerzijds het taxeren van zorgbehoeften en anderzijds het kiezen van zorg die aangrijpt op de getaxeerde zorgbehoeften ([AUTEURS GEMASKEERD VOOR REVIEW]). Deze zorg bestaat uit werkwijzen, behandelingen en interventies die volgens handleidingen en protocollen aangrijpen op tenminste één risicofactor voor kindermishandeling. In Figuur 1 is een schermafbeelding zichtbaar van de applicatie. Links in het scherm is het behoeftetaxatie-onderdeel zichtbaar waarin mogelijke zorgbehoeften in de vorm van dynamische risicofactoren voor kindermishandeling staan. Deze factoren zijn onderverdeeld in opvoedings-, gezins-, ouder-, en kindfactoren. In Figuur 1 worden de factoren getoond voor het tabblad opvoedingsfactoren. Rechts in het scherm worden de interventies getoond die aangrijpen op de geselecteerde factoren.

Hulpverleners kunnen meerdere factoren per tabblad selecteren. Met een aparte knop ('Toelichting risicofactoren') kan een toelichting voor elke factor in [TOOL GEMASKEERD] worden opgevraagd, zodat gebruikers de factoren zoveel mogelijk hetzelfde interpreteren. Nadat een gebruiker alle aanwezige zorgbehoeften voor een casus heeft geselecteerd, kunnen interventies worden gezocht die aangrijpen op deze getaxeerde zorgbehoeften (zie knop 'Zoek interventies/Profielen' in Figuur 1). Zo zijn in Figuur 1 drie opvoedingsfactoren geselecteerd, waarbij in de database van [TOOL GEMASKEERD] vier interventies zitten die elk aangrijpen op (tenminste) deze drie opvoedingsfactoren (zie de interventies rechts in het scherm; Figuur 1). Van elke getoonde interventie kan een toelichting worden opgevraagd (knop 'Toelichting interventie'), waarna informatie volgt over de doelstelling van een interventie, de doelgroep waarop een interventie zich richt, beschikbare

kennis over de effectiviteit van een interventie en waar in Nederland een interventie wordt aangeboden. Zoals Figuur 1 laat zien worden de interventies die aangrijpen op de selectie van zorgbehoeften getoond in afnemende volgorde van effectiviteitsbewijs waarbij de classificatie van Van Yperen, Veerman en Bijl (2017) is aangehouden. Na elke taxatie van zorgbehoeften en een zoekopdracht naar interventies die aangrijpen op de getaxeerde behoeften kunnen de resultaten van de behoeftetaxatie en de zoekopdracht worden opgeslagen.

Figuur 1

Schermafdruck van [TOOL GEMASKEERD] met daarin het Behoeftetaxatie-Onderdeel (Links) en Resultaten van de Zoekopdracht naar Interventies die Aangrijpen op de Getaxeerde Zorgbehoeften (Rechts)

The screenshot shows a software interface with a menu bar at the top (Bestand, Start, Maken, Externe gegevens, Hulpmiddelen voor databases, Help, Vertel wat u wilt doen) and a search bar. Below the menu bar, there are tabs for 'Opvoedingsfactoren', 'Gezinsfactoren', 'Ouderfactoren', 'Kindfactoren', and 'Doelgroepfactoren'. The 'Opvoedingsfactoren' tab is active, showing a list of risk factors. The 'Interventies' tab is also active, showing a table of interventions.

Behoeftetaxatie-Onderdeel (Links):

Selecteer per tabblad één of meer risicofactoren. Houd de CTRL-toets ingedrukt voor het selecteren van meerdere risicofactoren. Klik vervolgens op de knop ZOEK INTERVENTIES. Klik op TOELICHTING RISICOFACTOREN voor een toelichting van alle risicofactoren.

Opvoeder vertoont inadequaat opvoedgedrag (algemeen)

- Opvoeder heeft problemen met het stellen van regels en grenzen
- De opvoeding is onvoldoende gericht op bevordering van kindontwikkeling (kindstimulering)
- Er zijn problemen in de ouder-kind-interactie
- Opvoeder is onvoldoende sensitief
- Opvoeder is onvoldoende responsief
- Opvoeder heeft onrealistische verwachtingen van het kind
- Opvoeder heeft disfunctionele attitudes, waarden en normen
- Opvoeder ervaart het kind als een probleem
- Opvoeder hanteert een autoritaire opvoedingsstijl (algemeen)
- Opvoeder vertoont controlerend opvoedgedrag
- Opvoeder gebruikt lichamelijke straffen in de opvoeding
- Opvoeder gebruikt verbaal geweld in de opvoeding
- Opvoeder hanteert een nalatende opvoedingsstijl (algemeen)
- Opvoeder biedt onvoldoende bescherming, toezicht en veiligheid
- Opvoeder voorziet onvoldoende in de basale verzorging van het kind
- Er is onvoldoende stabiliteit in de opvoeding
- Opvoeder is onvoldoende emotioneel beschikbaar voor het kind

Resultaten van de Zoekopdracht naar Interventies die Aangrijpen op de Getaxeerde Zorgbehoeften (Rechts):

Onderstaande interventies grijpen aan op de geselecteerde risicofactoren. Selecteer een interventie, en klik vervolgens op TOELICHTING INTERVENTIE, om een toelichting van een interventie in te zien.

INTERVENTIE	EFFECTIVITEIT
Voorwaardelijke Interventies in Gezinnen (VIG)	2
Ambulante Spoedhulp	1
Multi Systeem Therapie - Child Abuse and Neglect (MST-CAN)	1
Veilig Sterk en Verder	1

Buttons at the bottom: TOELICHTING RISICOFACTOREN, ZOEK INTERVENTIES/PROFIELEN, TOELICHTING INTERVENTIE, OPSLAAN.

2.4. Data-Analyse

De audio-opnames van de interviews zijn getranscribeerd en vervolgens gecodeerd met ATLAS.ti (versie 8) volgens de codeerfases uit de richtlijnen van Boeije (2014). Eerst werden relevante fragmenten en quotes voorzien van een code (open coderen), waarna overlappende thema's werden verzameld in code-groepen (axiaal coderen). Ten slotte werden thema's vergeleken en aan elkaar gerelateerd in netwerken (selectief coderen). Het coderen van alle interviews resulteerde in een totaal van 159 codes, verdeeld over 16 codegroepen. In het codeerproces zijn zowel inductieve als deductieve methodes gebruikt om de data te coderen (Green & Thorogood 2018). Via deductieve analyse werden codes toegekend aan de hand van de interviewopzet. Per vragenlijstonderdeel werd een inductieve onderzoeksmethode

toegepast, zodat eventuele assumpties of referentiekaders van de auteurs geen invloed hadden op het definitieve codeerschema (Green & Thorogood, 2018). De eerste auteur codeerde alle interviews waarbij onduidelijkheden werden gemarkeerd. Deze onduidelijkheden werden middels discussie met de coauteurs besproken tot volledige overeenstemming was bereikt over het codeerschema inclusief codegroepen en thema's. Aanvullend zijn t-toetsen voor onafhankelijke groepen uitgevoerd om verschillen in het aantal getaxeerde behoeftefactoren tussen de conditie zonder en met [TOOL GEMASKEERD] te toetsen. Deze toetsen waren uitgevoerd per factorcategorie.

3. Resultaten

3.1. De Jeugdbeschermingspraktijk zonder [TOOL GEMASKEERD]

3.1.1. Huidige Werkwijze Behoeftetaxatie en Zorgkeuze

De meeste professionals ($n = 12$) taxeren zorgbehoeften tijdens een (intake)gesprek met de cliënt/het gezin (“*In eerste instantie gaan wij in gesprek met ouders en de kinderen om te bekijken wat de grootste zorg is in het gezin.*”). Zes professionals noemden daarbij dat overleg met collega's belangrijk is in het bepalen van de primaire zorgbehoeften en het opstellen van een behandelplan (“*Ik geef tijdens een vergadering aan waar ik naar op zoek ben en vervolgens geven collega's suggesties over welke zorg passend is.*”). Ook gaven sommige professionals ($n = 4$) aan dat ze over zorgbehoeften overleggen met andere betrokken instanties (waaronder zorgverwijzers en gemeenten) (“*Wij hebben een casuïstiekbespreking samen met ouders en het kind, ketenpartners en de gemeente.*”). Vijf professionals gaven aan meestal af te gaan op hun eigen ervaring in het taxeren van zorgbehoeften (“*Bij ons is er wel een gestructureerd protocol, maar in de praktijk doet iedere collega het toch op zijn eigen manier.*”). De participanten noemden verschillende instrumenten/methodieken in hun antwoord op de vraag hoe zij doorgaans een gestructureerde behoeftetaxatie uitvoeren, namelijk: (1) ARIJ-Risk, RiS 2.0 en SIRE-lijst (risicotaxatieinstrumenten), (2) de TOP-3 methodiek (redeneer- en beslismodel over onveiligheid in huiselijke kring), (3) het meten van de Big Five middels een persoonlijkheidsvragenlijst, (4) het Dialoog-model (een methode gericht op het bevorderen van afstemming, overeenstemming en samenwerking tussen hulpvragers en verleners), (5) Richtlijnen Jeugdhulp (globaal stappenplan), en (6) Ambulante Spoedhulp.

De meeste professionals ($n = 9$) gaven aan in overleg met collega's te bepalen welke zorg/interventies nodig zijn voor de cliënt/het gezin (“*Ik neem contact op met onze*

gedragwetenschapper en daarnaast werk ik veel samen met een zorgcoördinator. Soms raadpleeg ik een consulent van het gemeentelijke jeugdloket.”). Ook wordt vaak zorg ingezet op basis van eerdere ervaringen met cliënten (*“Ik kies voor zorg of een interventie waarmee ik bekend ben in mijn omgeving, of op grond van mijn eerdere ervaringen met bepaalde zorg of interventies”*). Drie professionals gebruikten de Zorgkeuze Midden-Holland (Gemeente Gouda, 2021), wat een website omvat met daarop een sociale kaart van zorgaanbieders in de regio. Daarnaast gaven drie professionals aan rekening te houden met het zorgverleden, het cognitief niveau of de identiteit (waaronder geloofsovertuiging) van cliënten bij het vinden van passende zorg. Ook noemde een professional vaak even te ‘Googelen’ naar welke zorg passend kan zijn.

3.1.2. Knelpunten in de Toeleiding naar Passende Zorg

Wachttijsten en Stroeve Samenwerking. De meeste professionals ($n = 12$) noemden de lange wachttijsten in de jeugdzorg als groot knelpunt in de toeleiding van cliënten naar passende zorg (*“De route die een cliënt moet afleggen om een doorverwijzing te krijgen van de gemeente verloopt echt moeizaam. Bij het gemeentelijke loket is de gemiddelde wachttijd zo’n zes weken. Vervolgens wordt daar de zorgbehoefte nogmaals in kaart gebracht en pas daarna wordt een aanmelding gedaan bij een zorginstantie, waarna eerst weer enkele maanden wachttijd verstrijkt.”*). Daardoor wordt soms ook een volgens de professionals minder passend alternatief aangeboden (*“Soms is een wachtlijst te lang, en dan ga ik op zoek naar een alternatief.”*, of *“Ik zit soms te stutten en steunen totdat ik de hulp kan inzetten die passend is.”*). Zes professionals gaven daarnaast aan dat een gebrek aan samenwerking of wisselingen tussen hulpverleners/instanties soms ook leidt tot vertraging in de zorg (*“Je moet net de goede persoon treffen die zegt: ik ga dat gelijk voor je regelen.”*).

Beperkte Inkoop, Ingewikkeld Financieringsbeleid en Lastige Verwijscriteria.

Zeven professionals noemden een beperkte inkoop van zorg door gemeenten als knelpunt bij het inzetten van passende zorg (*“Het inkoopbeleid van de gemeente is zó leidend dat je niet altijd de best passende zorg kan inzetten.”*). Twee professionals gaven daarnaast aan dat het verkrijgen van de benodigde financiering van zorg ook vaak leidt tot vertraging (*“Ik wacht vaak op administratieve rompslomp, terwijl het juist zo belangrijk is om aan een gezin te laten zien dat sprake is van een goede samenwerking met andere partijen.”*). Drie professionals noemden ook dat ingewikkelde verwijscriteria een barrière vormen voor aanmelding bij zorginstanties (*“Soms valt een gezin net niet onder de ingewikkelde verwijscriteria van een instantie, en dat is niet altijd even makkelijk.”*).

Onvoldoende zicht op Interventieaanbod en Zorgbehoeften. Drie professionals vertelden niet altijd bekend te zijn met het hulpaanbod (“*Ik vraag me wel eens af: ‘Is er niet een betere oplossing?’ Ik kies toch snel voor een interventie waarmee ik bekend ben, terwijl er misschien een beter alternatief is.*”). Ook gaven twee professionals aan dat ze soms te weinig informatie over problematiek en zorgbehoeften hebben om een inschatting te kunnen maken van welke zorg of hulp nodig is (“*Het is lastig wanneer je nog weinig informatie hebt en snel beslissingen moet nemen.*”).

Complexe Problematiek en Gebrekkige Motivatie. Vijf professionals gaven aan dat er niet altijd passende hulp beschikbaar is, omdat de problematiek te ingewikkeld is voor het zorgaanbod waaruit gekozen kan worden (“*Bij complexe echtscheidingen vind ik het heel lastig om passende hulp in te zetten, omdat ouders tegelijkertijd met een rechtszaak bezig zijn.*”). Vier professionals benoemden het gebrek aan motivatie bij cliënten als knelpunt (“*Wat doe je als mensen gewoon geen hulp willen?*”).

3.1.3. Suggesties ter Verbetering van de Jeugdbeschermingspraktijk

Professionals gaven meerdere suggesties op de vraag wat er nodig zou zijn om de knelpunten in het vinden van passende zorg in de jeugdbescherming te verhelpen. Zo is er behoefte aan meer kennis over (1) het hulpaanbod (“*Het zou helpend zijn om meer kennis te hebben over het aanbod in de regio.*”), (2) een betere samenwerking tussen betrokken instanties, (3) een specifiek zoekstelsel met mogelijke interventies (“*Thema’s zijn nu zo algemeen, zoals ‘autisme’ of ‘psychiatrie’, ik wil specifiekere kunnen zoeken.*”), (4) het inzetten van wachtlijstcoaches of een casusregisseur, en (5) een aangepast financieringssysteem (“*Het zou mooi zijn als ieder kind een persoonlijk budget krijgt, want dan maakt het niet uit met welke instanties er wel of geen contracten zijn.*”).

3.2. Behoeftetaxatie Met- en Zonder [TOOL GEMASKEERD]

Na het lezen van de fictieve casus bleek dat professionals significant meer zorgbehoeften taxeerden met gebruik van [TOOL GEMASKEERD] ($M = 21.67$, $SD = 5.98$), dan tijdens de eerste taxatieronde die gebaseerd was op alleen de eigen klinische inschatting ($M = 9.07$, $SD = 3.88$) ($t(28) = -6.84$, $p < .001$) (zie Tabel 2). Professionals taxeerden significant meer *familiefactoren* (o.a. huiselijk geweld, een problematische relatie tussen opvoeders en financiële problemen), *opvoedingsfactoren* (o.a. inconsistente opvoeding, verstoorde ouder-kindrelatie en problemen met het stellen van grenzen) en *ouderfactoren* (o.a. ouderlijke stress en agressieregulatie-problematiek).

Tabel 2

Aantal Gesignaleerde Zorgbehoeften (zonder en met [TOOL GEMASKEERD])

Categorie	Zonder [TOOL GEMASKEERD]		Met [TOOL GEMASKEERD]		<i>t</i>
	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	
Familie	3.53	1.06	5.87	0.83	-6.70***
Kind	2.53	2.48	3.60	1.50	-1.43
Opvoeding	1.73	0.88	6.40	3.38	-5.18***
Ouder	1.27	1.28	5.80	2.11	-7.11***
Totaal	9.07	3.88	21.67	5.98	-6.84***

Noot. Voor elke categorie is een independent samples *t*-test uitgevoerd om het verschil in het gemiddeld aantal getaxeerde zorgbehoeften te toetsen tussen de behoefte-taxaties met en zonder [TOOL GEMASKEERD]. Het gebruik van de tool was een dichotome variabele (0 = zonder [TOOL GEMASKEERD], 1 = met [TOOL GEMASKEERD]), wat betekent dat een negatieve *t*-waarde een hoger gemiddeld aantal getaxeerde zorgbehoeften impliceert met [TOOL GEMASKEERD].

*** $p < .001$.

3.3. Zorgkeuze Met- en Zonder [TOOL GEMASKEERD]

Na de behoefte-taxatie zonder [TOOL GEMASKEERD] gaven professionals verschillende suggesties voor zorg die zij zouden inzetten bij de fictieve gezinscasus, waaronder: sociaal maatschappelijke ondersteuning, systeemgerichte zorg, opvoedondersteuning, ambulante spoedhulp, mediation, stresshantering, veiligheidsafspraken, logopedie voor het kind en een kindercoach. Ook werden verschillende interventies genoemd, waaronder: ‘Kinderen uit de Knel’, ‘En nu ik...!’, ‘Families First’, ‘KIES’, ‘Ouderschap Blijft’, ‘Piep zei de muis’, ‘SCHIP-aanpak’, ‘Signs of Safety’ en ‘Words and Pictures’.

Bij de eerste taxatie met [TOOL GEMASKEERD], waarbij de respondenten alle aanwezige zorgbehoeften bij de casus taxeerden, bleek dat maar twee interventies in [TOOL GEMASKEERD]-database zaten die op *alle* aanwezige zorgbehoeften in de casus aansloten. Multisysteemtherapie – Child Abuse and Neglect (MST-CAN) kwam bijna altijd naar voren (93.3%), en Intensieve Ambulante Gezinsbegeleiding kwam één keer naar voren (6.7%). Bij de tweede taxatie met [TOOL GEMASKEERD], waarbij respondenten de zorgbehoeften met de meeste prioriteit taxeerden, bleken meer interventies in [TOOL GEMASKEERD]-database passend te zijn bij de casus. MST-CAN bleek ook hierbij het vaakst passend (93.3%), gevolgd door Ambulante Spoedhulp (66.7%), Voorwaardelijke Interventies in Gezinnen (60%),

Families First en Veilig Sterk en Verder (46.7%) en Parent Child Interaction Therapy en Intensieve Ambulante Gezinsbegeleiding (40%).

3.4 Evaluatie van [TOOL GEMASKEERD]

Interventieaanbevelingen. De meeste professionals (10) konden zich vinden in de interventieaanbevelingen die volgden uit de behoefte-taxatie met [TOOL GEMASKEERD]. Acht professionals gaven aan niet bekend te zijn met één of meerdere interventies. Twee professionals vonden de interventies wel passend, maar pas in een later behandelstadium (*“Ouders moeten eerst een stabiele woonomgeving krijgen.”*).

Module voor Behoeftetaxatie. Bij de beoordeling van de bruikbaarheid van het behoefte-taxatie-onderdeel was het merendeel van de feedback positief. Professionals benoemden onder andere dat het overzicht van de verschillende domeinen structuur biedt en handig is tijdens het schrijven van rapportages. Ook gaven ze aan dat de tool makkelijk in teamverband gebruikt kan worden en goed inzetbaar is bij multiprobleemgezinnen. Daarnaast gaven professionals aan dat er overlap zit in sommige factoren (*“Vooral bij opvoedingsfactoren, waarbij veel andere factoren lijken te vallen onder de algemene factor ‘er is onvoldoende stabiliteit in de opvoeding’.*”). Ook benoemden twee professionals het belang van het kunnen aangeven van de ernst van de problematiek, oftewel dat de huidige dichotome antwoordoptie (wel/niet aanwezig van een factor) uitgebreid zou moeten worden met een ernstgradatie.

Module voor Zorgkeuze. Vijf professionals noemden het interventieoverzicht in [TOOL GEMASKEERD] een verbreding van de horizon (*“Het geeft een mooi beeld van de mogelijkheden, ook van interventies waaraan je in eerste instantie niet zou denken. Normaal gesproken kies je toch voor een interventie waar je bekend mee bent, of waar je goede ervaringen mee hebt.”*). Zes professionals vonden het interventieoverzicht in het algemeen nuttig (*“Het is best concreet en ik denk dat het heel helpend kan zijn.”*). Drie professionals benoemden dat het interventieoverzicht moet aansluiten bij het beschikbare aanbod in de regio waarin gewerkt zou worden met [TOOL GEMASKEERD]. Verder gaven professionals aan dat het zinvol zou zijn om het interventieaanbod uit te breiden met diverse methodieken, waaronder Signs of Safety en de MPG-Aanpak (ketenaanpak voor zorg aan multiprobleemgezinnen, regio Amsterdam).

Ontbrekende Informatie. Alle professionals gaven aan dat er geen overbodige informatie in de tool staat. Wel zeiden drie professionals de ruimte voor positieve/beschermende factoren (zoals sociale steun) te missen in de tool. Andere missende

factoren waren volgens de professionals: hulpverleningsgeschiedenis, (complexe) scheiding van ouders, specifieke kindfactoren (i.p.v. algemeen externaliserende- en internaliserende problematiek), trauma (kind), leeftijdsindicaties van interventies en het mentaliserend vermogen van ouders. Daarnaast misten enkele participanten de optie voor prioritering van factoren (ernstgradatie) en de verwerking van veiligheidsafspraken als eventuele bijlage.

Gebruikerservaringen en Suggesties voor Bevordering van Gebruiksgemak. Twaalf professionals vonden [TOOL GEMASKEERD] over het algemeen een gebruiksvriendelijke en overzichtelijke tool (“*Ik vind hem heel overzichtelijk, vooral de tabbladen/categorieën zijn erg helpend.*”). Drie professionals suggereerden dat het makkelijker zou zijn als de Ctrl-toets niet ingedrukt hoeft te worden bij de selectie van behoeftefactoren. Andere suggesties waren: een minder ‘basic’ ontwerp, online beschikbaarheid van de tool, minder gebruik van jargon, ‘onbekend’ toevoegen als antwoordoptie en continue zichtbaarheid van de toelichtingen bij de factoren. Eén professional benoemde dat het ook nuttig zou zijn als ouders de tool zelf kunnen gebruiken om te onderzoeken welke hulp zij nodig zouden hebben.

Gebruik in de Praktijk. Alle professionals gaven aan de tool in de praktijk te willen gebruiken (“*Je kiest toch snel voor een interventie die je vaker hebt ingezet en waar je positieve ervaringen mee hebt in plaats van een minder bekende, maar eventueel beter passende interventie. Ik denk dat je hiermee een heel goed beeld krijgt van de mogelijkheden.*”). Ook zouden de professionals de tool bij collega’s aanraden (“*Zeker voor wijkteams of spoedhulp.*”). Ten slotte benoemden meerdere professionals dat het nuttig zou zijn om [TOOL GEMASKEERD] te koppelen aan de [TOOL GEMASKEERD].

4. Discussie

[TOOL GEMASKEERD] is, voor zover bekend, het eerste instrument dat specifiek ontwikkeld is om professionals in jeugdbescherming te ondersteunen in behoeftetaxatie én besluitvorming over passende zorg. Het doel van deze studie was om te onderzoeken of [TOOL GEMASKEERD] daadwerkelijk ondersteunend is in behoeftetaxatie en het besluitvormingsproces. Om hier uitspraken over te kunnen doen werd in deze studie ook nagegaan wat de huidige werkwijze is van professionals voor het taxeren van zorgbehoeften en het kiezen van passende zorg.

4.1. [TOOL GEMASKEERD] in de Jeugdbeschermingspraktijk

Uit dit onderzoek blijkt dat ongestructureerde besluitvorming over passende zorg nog altijd gebruikelijk is in de jeugdbeschermingspraktijk, ondanks de voordelen van een gestructureerd

klinisch besluitvormingsproces (Douglas et al., 2002; [TOOL GEMASKEERD]). De jeugdzorgprofessionals die in dit onderzoek zijn geïnterviewd baseren hun keuze over passende zorg doorgaans op eigen ervaringen en adviezen van collega's. Valkuilen van dergelijke intuïtieve besluitvorming zijn het optreden van vooroordelen (bias) in het beoordelen van een casus en het diagnostisch proces (Helm, 2011; Saltiel, 2016). Zo kan een professional door bias in de oordeelsvorming vooral aandacht hebben voor bepaalde herkenbare problematiek door eerdere ervaringen met dezelfde problemen en minder aandacht hebben voor andere (complexe) informatie die niet direct samenhangt met de gesignaleerde problematiek (Saltiel, 2016).

Verder bleken de geïnterviewde professionals risicotaxatie-instrumenten (o.a., ARIJ-Risk, RiS 2.0 en SIRE-lijst) te gebruiken in het besluitvormingsproces over passende interventies in de jeugdbescherming, in plaats van behoeftetaxatie-instrumenten. Zoals beschreven in het visiedocument “Gefaseerde ketensamenwerking bij kindermishandeling, huiselijk geweld en seksueel misbruik” (Vogtländer & van Arum, 2016), is het belangrijk om onderscheid te maken tussen het werken aan kindveiligheid (1), risicogestuurde (2)- en herstelgerichte zorg (3). Zoals eerder genoemd maakt risicotaxatie inzichtelijk wie hulpverlening nodig heeft (gezinnen met een verhoogd risico) en met welke intensiteit deze hulpverlening aangeboden moet worden (risicogestuurde zorg), terwijl behoeftetaxatie zichtbaar maakt wat aangepakt moet worden met welke interventies (herstelgerichte zorg) ([AUTEURS GEMASKEERD VOOR REVIEW]). [TOOL GEMASKEERD] is daarmee een aanvulling op het reeds bestaande aanbod van instrumenten voor het opstellen van een systeemgericht hulpverleningsplan.

De resultaten van dit onderzoek sluiten aan op eerder onderzoek waaruit blijkt dat gestructureerde besluitvorming (zoals met [TOOL GEMASKEERD]) tunnelvisie kan voorkomen en een holistische benadering van probleemgezinnen in de jeugdbescherming bevordert (Bartelink et al., 2015). Bij een holistisch besluitvormingsproces wordt naast zorgbehoeften van kinderen rekening gehouden met zorgbehoeften van de opvoeders en/of de bredere sociale omgeving. Zo bleek dat professionals tijdens de gestructureerde behoeftetaxatie met [TOOL GEMASKEERD] significant meer zorgbehoeften signaleerden dan tijdens de klinische taxatie zonder [TOOL GEMASKEERD]. Daarnaast viel op dat professionals vooral gezins- en kindfactoren en relatief weinig ouderfactoren taxeerden in hun eigen klinisch oordeel van zorgbehoeften, terwijl zij significant meer gezins-, ouder-, en opvoedingsfactoren signaleerden wanneer gewerkt werd met [TOOL GEMASKEERD]. Hiermee lijkt [TOOL GEMASKEERD] een belangrijke meerwaarde te hebben ten opzichte

van ongestructureerde behoefte-taxatie, omdat opvoedingsfactoren (o.a. problematische ouder-kind relatie) en ouderfactoren (o.a. crimineel gedrag en psychopathologie) aanzienlijk sterkere voorspellers zijn voor kindermishandeling dan kindfactoren ([AUTEURS GEMASKEERD VOOR REVIEW]; Stith et al., 2009). Ook uit een calamiteitenonderzoek (Toezicht Sociaal Domein, 2015) bleek dat professionals onvoldoende aandacht hadden voor risicofactoren van ouders in kwetsbare gezinnen waarin de veiligheid van kinderen in gevaar was. [TOOL GEMASKEERD] kan professionals dus ondersteunen in het signaleren van de meest relevante risicofactoren die fungeren als aanknopingspunten voor behandeling.

Over de gebruiksvriendelijkheid van [TOOL GEMASKEERD] zeiden professionals dat het overzicht van risicofactoren in verschillende domeinen structuur biedt bij het in kaart brengen van hulpbehoeften. Daarbij werd benoemd dat zo'n overzicht belangrijke en inzichtelijke aanknopingspunten biedt voor het maken van doeltreffende rapportages of gezinsplannen, wat gezinswerkers vooral beschouwen als een administratieve last (Sekreve et al., 2020). De zorgkeuzemodule van [TOOL GEMASKEERD] werd gezien als een 'verbreding van de horizon' in het voortdurend groeiende aanbod van zorg en interventies, die al dan niet effectief zijn bevonden. Uit recent onderzoek blijkt dat veel gezinswerkers onbekend zijn met effectieve behandelvormen die bijvoorbeeld zijn opgenomen in de NJI Databank Effectieve Jeugdinterventies (DEJ) (Sekreve et al., 2020). [TOOL GEMASKEERD] vergroot niet alleen de kennis van professionals over het beschikbare aanbod van behandelingen en interventies, maar maakt ook inzichtelijk welke interventies aangrijpen op de risicofactoren die professionals selecteren in een behoefte-taxatie. Bovendien maakt [TOOL GEMASKEERD] inzichtelijk in welke mate de werkzaamheid van interventies wetenschappelijk is onderbouwd ([TOOL GEMASKEERD]), wat een belangrijk aspect is om mee te wegen in besluitvorming over passende zorg.

4.2. Aanbevelingen voor de Doorontwikkeling van [TOOL GEMASKEERD]

De geïnterviewde professionals noemden ook suggesties om de tool te verbeteren en de inzetbaarheid ervan in de jeugdbeschermingspraktijk te vergroten. Ten eerste noemden zij dat niet alle interventies in de database van [TOOL GEMASKEERD] beschikbaar zijn in elke regio in Nederland vanwege het decentrale, gemeentelijk inkoopbeleid van jeugdzorg. Enerzijds belemmert dat de algemene bruikbaarheid van de tool, maar anderzijds biedt de tool op deze manier wel inzicht in lacunes in het regionale hulpaanbod. Zo kunnen jeugdzorginstanties in overleg met gemeenten aangeven welke nog niet beschikbare interventies geschikt zouden zijn voor hun cliënten.

Ten tweede merkten professionals op dat de huidige module voor behoeftetaxatie alleen (veranderbare) risicofactoren meet en geen beschermende factoren. In de doorontwikkeling van [TOOL GEMASKEERD] kan worden overwogen om beschermende factoren op te nemen in de behoeftetaxatie om het draagvlak voor deze tool in de jeugdbescherming te vergroten (Wilkins, 2015). Het blijft daarbij wel belangrijk dat dynamische risicofactoren centraal staan in zowel behoeftetaxatie als de behandeling van gezinnen, zodat de klinische praktijk aansluit bij het bewezen effectieve Needs-principe van het RNR-model (REF). Bovendien blijkt uit onderzoek dat kinderen uit hoog-risico-gezinnen minder profiteren van de eventuele aanwezigheid van beschermende factoren dan laag-risico-gezinnen (Luthar & Goldstein, 2004).

Ten derde bleek dat er slechts een beperkt aanbod is van interventies die aansluiten op multidimensionale gezinsproblematiek waarbij veel en diverse behoeftefactoren aanwezig zijn, zoals het geval was in de casus die in dit onderzoek werd voorgelegd aan de respondenten. Naarmate minder risicofactoren worden geselecteerd in [TOOL GEMASKEERD] neemt het aantal interventies dat aangrijpt op alle geselecteerde risicofactoren toe en daardoor toont [TOOL GEMASKEERD] ook meer passende interventies in de zorgkeuze-module. Professionals gaven de suggestie om een ernstgradatie toe te voegen aan de dichotome antwoordcategorieën (wel/niet aanwezig) bij de behoeftetaxatie. Met deze mogelijkheid kunnen professionals zicht krijgen op behoeftefactoren met de meeste prioriteit en interventiemogelijkheden die aansluiten op deze factoren. Een belangrijke kanttekening hierbij is dat [TOOL GEMASKEERD] interventiemogelijkheden biedt die meestal in geprotocolleerde vorm worden aangeboden. Hoewel deze interventies algemeen werkzame factoren kunnen bevatten die een interventie effectief maken, ongeacht het doel, de doelgroep en het soort interventies, kunnen deze interventies essentiële specifiek werkzame elementen missen die een interventie effectiever maken voor gezinnen waarin sprake is van kindonveiligheid (van Yperen et al., 2015). Hierbij is belangrijk te vermelden dat voor het terugdringen van kindermishandeling tot op heden nog geen specifiek werkzame elementen van interventies zijn gevonden die eventueel gekoppeld kunnen worden aan behoeftefactoren van de doelgroep ([AUTEURS GEMASKEERD VOOR REVIEW]). Daarnaast is ten tijde van dit schrijven slechts 1,82% van de interventies in de NJI-databank bewezen effectief volgens sterke aanwijzingen. Dit maakt dat de interventieaanbevelingen door [TOOL GEMASKEERD] kritisch geëvalueerd dienen te worden door de gebruikers, waarbij de geschiktheid van een interventie op papier afgewogen moet worden tegen de daadwerkelijke aangetoonde effectiviteit.

Ten slotte is het belangrijk om naast gebruikerservaringen ook de psychometrische kwaliteit van [TOOL GEMASKEERD] te onderzoeken. Hoewel uit eerder onderzoek blijkt dat toepassing van instrumenten die gebaseerd zijn op de RNR-principes het besluitvormingsproces in de jeugdbescherming verbeteren ([AUTEURS GEMASKEERD VOOR REVIEW]), blijft een urgente vraag of implementatie van [TOOL GEMASKEERD] werkelijk leidt tot betere besluitvorming in de praktijk en dus tot minder kindermishandeling. Daarnaast waren geen kenmerken van de professionals gemeten die mogelijk van invloed zijn op het besluitvormingsproces in de jeugdbescherming. Zo zijn etnische achtergrond, aantal jaar werkervaring en diverse attitudes over overschrijdend opvoedingsgedrag mogelijk van invloed op de besluitvorming over passende ingrepen bij kindonveiligheid (Jent et al., 2011; Lauritzen et al., 2018; Lee et al., 2013). In vervolgonderzoek is het relevant om de eventuele invloed van dergelijke factoren op het gebruik van [TOOL GEMASKEERD] te toetsen.

4.3. Klinische Implicaties

In dit onderzoek benoemden professionals knelpunten in de huidige jeugdbeschermingspraktijk die ook in eerdere onderzoeken naar voren zijn gekomen, zoals lange wachtlijsten, beperkingen in het hulpaanbod door het inkoopbeleid van gemeenten en een gebrekkige samenwerking tussen instanties ([AUTEURS GEMASKEERD VOOR REVIEW]; Inspectie Gezondheidszorg en Jeugd, 2020). Een gestructureerd, consistent besluitvormingsproces over passende interventies kan de samenwerking tussen partijen in de jeugdbeschermingspraktijk verbeteren (De Bortoli et al., 2016). Het is belangrijk om te benadrukken dat [TOOL GEMASKEERD] geen vervanging is van het oordeel van professionals, omdat casus-specifieke en tijdgebonden factoren, zoals ernst en spoedeisendheid, meewegen in de keuze voor de best passende interventie ([AUTEURS GEMASKEERD VOOR REVIEW]). Daarnaast is een inschatting van responsiviteitsfactoren van cliënten in de jeugdbescherming (waaronder niveau van motivatie, het cognitief functioneren van opvoeder(s) en de culturele achtergrond van een gezin) belangrijk om zorg op maat te kunnen bieden wat de effectiviteit van de zorg ten goede komt ([AUTEURS GEMASKEERD VOOR REVIEW]). Ook worden ouders en kinderen bij voorkeur betrokken bij het besluitvormingsproces om het vertrouwen en de kans op een positieve samenwerking te vergroten (van Bijleveld et al., 2019; Helm, 2011; [AUTEURS GEMASKEERD VOOR REVIEW]).

4.4. Conclusie

Uit deze pilotstudie blijkt dat [TOOL GEMASKEERD] professionals ondersteunt in behoeftetaxatie en besluitvorming over passende zorg voor het verlagen van het risico op (de herhaling van) kindermishandeling. Professionals signaleerden significant meer zorgbehoeften bij een multiprobleemgezin wanneer zij zorgbehoeften taxeerden met [TOOL GEMASKEERD] dan wanneer zij taxeerden volgens hun eigen klinische ongestructureerde oordeel. Ook signaleerden professionals meer opvoedings- en ouderfactoren met [TOOL GEMASKEERD], wat de sterkste voorspellers zijn voor kindermishandeling en daarmee belangrijke aanknopingspunten voor behandeling. De zorgkeuzemodule van [TOOL GEMASKEERD] werd beoordeeld als een ‘verbreding van de horizon’ in het voortdurend groeiende aanbod van zorg en interventies in de jeugdbeschermingspraktijk.

Korte Verklaring over de Data en Analyses

Analyse codes en output: De analyse codes en output zijn niet vrij toegankelijk, maar wel opvraagbaar bij de auteur(s).

Data: Wegens gevoeligheid van de (interview)data zijn deze niet vrij toegankelijk.

Informatie/vragen over de data is op te vragen bij/kunnen gesteld worden aan
[GEMASKEERD VOOR REVIEW].

Financiering

Deze studie is gesubsidieerd door ZonMw (projectnummer: [GEMASKEERD VOOR REVIEW]).

Belangen

[GEMASKEERD VOOR REVIEW]

Literatuur

[REFERENTIE GEMASKEERD VOOR REVIEW]

Andrews, D. A., Bonta, J., & Hoge, R. D. (1990). Classification for effective rehabilitation: Rediscovering psychology. *Criminal Justice and Behavior*, 17, 19-52.

<https://doi.org/10.1177/0093854890017001004>

Andrews, D. A., Bonta, J., & Wormith, J. S. (2011). The Risk-Need-Responsivity (RNR) Model: Does adding the Good Lives Model contribute to effective crime prevention? *Criminal Justice and Behavior*, 38(7), 735-755.

<https://doi.org/10.1177/0093854811406356>

Bartelink, C., van Yperen, T. A., & ten Berge, I. J. (2015). Deciding on child maltreatment: A literature review on methods that improve decision-making. *Child Abuse & Neglect*, 49, 142-153. <https://doi.org/10.1016/j.chiabu.2015.07.002>

Belsky, J. (1993). Etiology of child maltreatment: A developmental ecological analysis. *Psychological Bulletin*, 114(3), 413-434.

<http://dx.doi.org/10.1037/0033-2909.114.3.413>

Bijleveld, van, G. G., Bunders-Aelen, J. F. G., Dedding, C. W. M. (2019). Exploring the essence of enabling child participation within child protection services. *Child & Family Social Work*, 25(2), 286-293. <https://doi.org/10.1111/cfs.12684>

[REFERENTIE GEMASKEERD VOOR REVIEW]

Boeije, H. (2014). *Analyseren in kwalitatief onderzoek*. Amsterdam: Boom

Bonta, J., & Andrews, D. A. (2016). *The psychology of criminal conduct (sixth edition)*. New York: Routledge

De Bortoli, L., Ogloff, J., Coles, J., & Dolan, M. (2017). Towards best practice: Combining evidence-based research, structured assessment and professional judgement. *Child & Family Social Work*, 22(2), 660-669. <https://doi.org/10.1111/cfs.12280>

Douglas, K. S., & Kropp, P. R. (2002). A prevention-based paradigm for violence risk assessment: Clinical and research applications. *Criminal Justice and Behavior*, 29(5), 617-658. <https://doi.org/10.1177/009385402236735>

Dowden, C., & Andrews, D. A. (1996). What works for female offenders: A meta-analytic review. *Crime and Delinquency*, 45(4), 438-452.

<https://doi.org/10.1177/0011128799045004002>

Dowden, C., & Andrews, D. A. (2000). Effective correctional treatment and violent reoffending: A meta-analysis. *Canadian Journal of Criminology*, 42, 449-467.

<https://doi.org/10.3138/cjcrim.42.4.449>

- Dowden, C., & Andrews, D. A. (2003). Does family intervention work for delinquents? Results of a meta-analysis. *Canadian Journal of Criminology and Criminal Justice*, 45(3), 327-342. <https://doi.org/10.3138/cjccj.45.3.327>
- Euser, S., Alink, L. R. A., Stoltenborgh, M., & Bakermans-Kranenburg, M. J. (2015). A gloomy picture: A meta-analysis of randomized controlled trials reveals disappointing effectiveness of programs aiming at preventing child maltreatment. *BMC Public Health*, 15, 1068. <https://doi.org/10.1186/s12889-015-2387-9>
- Gemeente Gouda. (2021). *Zorgkeuze Midden Holland*. Verkregen via: <https://www.zorgkeuzemiddenholland.nl>
- Green, J. & Thorogood, N. (2018). *Qualitative Methods for Health Research. 4th edition*. London: SAGE.
- Guest, G., Bunce, A., & Johnson, L. (2006). How many interviews are enough? An experiment with data saturation and variability. *Field methods*, 18(1), 59-82. <https://doi.org/10.1177/1525822X05279903>
- [REFERENTIE GEMASKEERD VOOR REVIEW]
- Hanson, R. K., Bourgon, G., Helmus, L., & Hodgson, S. (2009). The principles of effective correctional treatment also apply to sexual offenders: A meta-analysis. *Criminal Justice and Behavior*, 36(9), 865-891. <https://doi.org/10.1177/0093854809338545>
- Helm, D. (2011). Judgments or assumptions? The role of analysis in assessing children and young people's needs. *British Journal of Social Work*, 41, 894-911. <https://doi.org/10.1093/bjsw/bcr096>
- Helm, D., & Roesch-Marsh, A. (19 oktober, 2010). *Ecology of judgment in child welfare and protection*. University of Stirling. <https://www.iriss.org.uk/sites/default/files/2016-07/iriss-ecology-of-judgement-report-2011-07-13.pdf>
- Hindley, N., Ramchandani, P. G., & Jones, D. P. (2006). Risk factors for recurrence of maltreatment: A systematic review. *Archives of Disease in Childhood*, 91(9), 744-752. <http://dx.doi.org/10.1136/adc.2005.085639>
- Inspectie Gezondheidszorg en Jeugd, Ministerie van Volksgezondheid, Welzijn en Sport. (2020). *Voortgangsrapportage kwetsbare kinderen onvoldoende beschermd*. Verkregen via: <https://www.rijksoverheid.nl/documenten/rapporten/2020/10/15/tk-bijlage-voortgangsrapportage-igj-en-ijenv-kwetsbare-kinderen-onvoldoende-beschermd>
- Jent, J. F., Eaton, C. K., Knickerbocker, L., Lambert, W. F., Merrick, M. T., &

- Dandes, S. K. (2011). Multidisciplinary child protection decision making about physical abuse: Determining substantiation thresholds and biases. *Children and Youth Services Review*, 33(9), 1673–1682. <https://doi.org/10.1016/j.childyouth.2011.04.029>
- Lauritzen, C., Vis, S. A., & Fossum, S. (2018). Factors that determine decision making in child protection investigations: A review of the literature. *Child & Family Social Work*, 23(4), 743-756. <https://doi.org/10.1111/cfs.12446>
- Lee, S. J., Sobock, J. L., Djelaj, V., & Agius, E. (2013). When practice and policy collide: Child welfare workers' perceptions of investigation processes. *Children and Youth Services Review*, 35(4), 634–641. <https://doi.org/10.1016/j.childyouth.2013.01.004>
- Luthar, S. S., & Goldstein, A. (2004). Children's exposure to community violence: Implications for understanding risk and resilience. *Journal of Clinical Child and Adolescent Psychology*, 33(3), 499-505. https://doi.org/10.1207/s15374424jccp3303_7
- [REFERENTIE GEMASKEERD VOOR REVIEW]
- Ng, M. Y., & Weisz, J. R. (2016). Annual research review: Building a science of personalized intervention for youth mental health. *Journal of Child Psychology and Psychiatry*, 57(3), 216-236. <https://doi.org/10.1111/jcpp.12470>
- [REFERENTIE GEMASKEERD VOOR REVIEW]
- [REFERENTIE GEMASKEERD VOOR REVIEW]
- [REFERENTIE GEMASKEERD VOOR REVIEW]
- [REFERENTIE GEMASKEERD VOOR REVIEW]
- [REFERENTIE GEMASKEERD VOOR REVIEW]
- Saltiel, D. (2016). Observing front line decision making in child protection. *British Journal of Social Work*, 46, 2104-2119. <https://doi.org/10.1093/bjsw/bcv112>
- Sekreve, A., Campmans, M., van Bon-Martens, M., & Orobio de Castro, B. (2020). Doelgericht werken in wijkteams laat nog te wensen over. *Kind & Adolescent Praktijk*, 19, 6-12. <https://doi.org/10.1007/s12454-020-0627-3>
- Spies, G. M., Delport, C. S. L., & le Roux, M. P. (2015). Developing safety and risk assessment tools and training materials: A researcher-practice dialogue. *Research on Social Work Practice*, 25(6), 670-680. <https://doi.org/10.1177/1049731514565393>
- Stoltenborgh, M., Bakermans-Kranenburg, M.J., Alink, L. R. A., & Van IJzendoorn, M.H. (2015). The Prevalence of child maltreatment across the globe: Review of a series of meta-analyses. *Child Abuse Review*, 24(1), 37-50. <https://doi.org/10.1002/car.2353>

Stith, S. M., Ting Liu, L., Davies, C., Boykin, E. L., Alder, M. C., Harris, J. M., Som, A., McPherson, M., & Dees, J. E. M. E. G. (2009). Risk factors in child maltreatment: A meta analytic review of the literature. *Agression and Violent Behavior, 14*(1), 13-29. <https://doi.org/10.1016/j.avb.2006.03.006>

Toezicht Jeugd/Sociaal Domein. (2015). *Calamiteitenonderzoek Amsterdam: Borgen van veiligheid in kwetsbare gezinnen*. Verkregen via: <https://www.toezichtsociaaldomein.nl/documenten/rapporten/2015/09/30/calamiteit-amsterdam-borgen-van-veiligheid-in-kwetsbare-gezinnen>

[REFERENTIE GEMASKEERD VOOR REVIEW]

[REFERENTIE GEMASKEERD VOOR REVIEW]

Vogtländer, L., & Arum, van, S. (mei 2016). *Eerst samenwerken voor veiligheid, dan samenwerken voor risicogestuurde zorg*. VNG & GGD GHOR Nederland.

https://vng.nl/files/vng/201605_visiedocument_gefaseerde_ketensamenwerkingvogtlander_van_arum_0.pdf

Weisz, J. R. (2014). Building robust psychotherapies for children and adolescents.

Perspectives on Psychological Science, 9(1), 81-84.

<https://doi.org/10.1177/1745691613512658>

Wilkins, D. (2015). Balancing risk and protective factors: How do social workers and social work managers analyse referrals that may indicate children are at risk of significant harm. *British Journal of Social Work, 45*, 395-411.

<https://doi.org/10.1093/bjsw/bct114>

Wilson K., Ruch G., Lymberly M., Cooper A. (2008). *Social work. An introduction to contemporary practice*. London, England: Pearson Longman.

Yperen, van, T., Veerman, J. W., & Bijl, B. (2017). *Zicht op effectiviteit. Handboek voor resultaatgerichte ontwikkeling van interventies in de jeugdsector*. Rotterdam:

Lemniscaat

Yperen, van, T., Veerman, J.W., & Berg, van den, G. (juni 2015). *Elementen die er toe doen: Overzicht van begrippen over werkzame elementen en een voorstel voor een*

indeling. ZonMw. https://www.zonmw.nl/sites/zonmw/files/typo3-migrated-files/Begrippenkader_Werkz_Elem_versie_juni_2015.pdf

Bijlage 1 – Casus

Het gezin bestaat uit vader, moeder en hun drie zoons K. (9), I. (7) en U. (3). Ouders zijn na 12 jaar huwelijk uit elkaar gegaan. Over hun relatie is bekend dat er veel ruzie was. Volgens vader is moeder regelmatig vreemdgegaan, al geeft moeder aan dat op deze momenten de relatie tussen vader en moeder over was. Na het verbreken van de relatie heeft vader nog lang hoop gehouden dat het weer goed zou komen. Dit heeft voor veel onduidelijkheid gezorgd over de relatie, ook bij de kinderen.

Ouders strijden nog altijd met elkaar, waarbij ook af en toe sprake is van fysiek geweld. Zo heeft vader recent moeder in het gezicht geslagen en heeft moeder een week later tijdens een ruzie een brandende kaars naar vader gegooid. Vader geeft aan dat moeder hem meerdere malen fysiek heeft aangevallen. Tijdens een van deze ruzies is de politie betrokken geweest.

Moeder heeft geen eigen woning. Ze woont momenteel in een huis dat haar ouders voor haar en de kinderen hebben gehuurd. Het jongste kind woont bij haar, de middelste af en toe. Vader heeft ook geen eigen woning. Hij woont bij zijn ouders. Het oudste kind woont bij vader en de middelste is ook meestal daar.

Moeder heeft momenteel geen werk, ze heeft schulden en staat onder bewind voering. Vader heeft een fulltimebaan bij een bakkerij. Ook vader heeft schulden en is bezig met een plan voor schuldhulpverlening. Moeder voelt zich nog steeds erg afhankelijk van vader en vader probeert hier zoveel mogelijk aan tegemoet te komen. Zo vraagt moeder veel hulp in de opvoeding en vraagt zij regelmatig om geld bij vader.

De kinderen hebben veel moeite met de scheiding van ouders. De oudste zoon K. geeft moeder de schuld van de scheiding. Hij wil dan ook niet graag naar zijn moeder toe. Als hij gaat is hij tegendraads en luistert hij slecht. Moeder belt dan naar vader om hem weer op te halen. Verder heeft K. veel woede-uitbarstingen (schopt en slaat moeder). De woede-uitbarstingen komen volgens vader niet voor als de oudste bij hem is.

De middelste zoon I. neemt het zijn vader kwalijk dat hij moeder heeft geslagen en vindt het lastig om te kiezen waar hij wil wonen. Ouders maken hier geen duidelijke keuze in voor hem. Hij is getuige geweest van het feit dat vader zijn moeder heeft geslagen en dat de politie hieraan te pas moest komen. Dit heeft hij uitgebreid op school in de kring verteld.

De jongste zoon U. moet veel huilen en klampt zich veel aan moeder op het moment dat zij weggaat. Daarnaast slaapt hij veel bij moeder in bed, omdat hij bang is. Opvallend aan de jongste is, is dat hij niet goed verstaanbaar praat.

In de eerste paar gesprekken komt naar voren dat de situatie in het gezin niet verbetert. Vader en moeder willen verbeteren, maar ze blijven veel ruzie maken. Er zijn sterke twijfels of de kinderen wel thuis kunnen blijven wonen. Moeder is bereid met de hulpverlening mee te werken. Ze geeft aan stress te ervaren en moeite te hebben met de opvoeding. Vader wil echter niet meewerken. Hij schaamt zich voor waar hij woont, zijn ouders roken en hij is bang dat de kinderen uit huis worden geplaatst.

Tweede reactie reviewer 1 – 04-10-2023

Het manuscript is verbeterd en tegemoet gekomen aan de opmerkingen van reviewers.

Tweede reactie reviewer 2 - 04-10-2023

Review 2: Jeugdbescherming op maat: Een pilotonderzoek

Ik bedank de auteurs voor het zorgvuldig omgaan met de feedback op het artikel en de gemaakte aanpassingen.

Het artikel is nu publicabel.

Wel geef ik de auteurs nog twee suggesties mee:

1)

In reactie op mijn feedback hebben auteurs de volgende zinnen toegevoegd:
'In het semigestructureerde interview werden allereerst vragen gesteld over het huidige besluitvormingsproces over passende behandeling in de jeugdbeschermingspraktijk. Vervolgens werd een fictieve casus voorgelegd aan de professional over de situatie van een multiprobleemgezin (Bijlage 1).'

Ik adviseer de auteurs om i.p.v. 'multiprobleemgezinnen' gebruik te maken van 'gezinnen met meervoudige en complexe problemen'. Dit wordt door gezinnen minder stigmatiserend ervaren en is in lijn met de Richtlijn Gezinnen met meervoudige en complexe problemen uit 2020.

2)

De auteurs hebben in de discussie een alinea toegevoegd:

'Daarnaast is ten tijde van dit schrijven slechts 1,82% van de interventies in de NJI-databank bewezen effectief volgens sterke aanwijzingen. Dit maakt dat de interventieaanbevelingen door [TOOL GEMASKEERD] kritisch geëvalueerd dienen te worden door de gebruikers, waarbij de geschiktheid van een interventie op papier afgewogen moet worden tegen de daadwerkelijke aangetoonde effectiviteit.'

Ik vermoed dat de 1,82% betrekking heeft op alle interventies in de databank? Of gaat het hier om interventies gericht op de doelgroep van het artikel?. Dus graag nog specificeren en zo mogelijk een referentie toevoegen.

Tweede reactie reviewer 3 - 14-3-2023

Er is een mooi artikel en goed tegemoet gekomen aan de aandachtspunten van de reviewers.

Tweede reactiebrief van auteur - 15-5-2023

Utrecht, 15 mei 2023

Onderwerp: Tweede revisie manuscript

Geachte Prof. dr. Boendermaker en meneer Veenstra,

Bijgevoegd vindt u de tweede revisie van ons manuscript: “Stevig als een rots, flexibel als water? Effectiviteit van een schoolinterventie gericht op het intrapersonlijk en interpersoonlijk domein”. We bedanken de editor en de reviewers voor de laatste punten van feedback en de complimenten.

We hebben de laatste punten van Reviewer A verwerkt in ons manuscript en de gevraagde informatie toegevoegd. Ter vergemakkelijking van het reviewproces hebben wij onze reacties op de reviewers hieronder als appendix toegevoegd en de aanpassingen in het manuscript geel gemarkeerd.

We zien uw reactie op onze revisie met belangstelling tegemoet.

Met vriendelijke groet,

Namens mijn co-auteurs Deković, Van Londen, en Reitz,

Esther Mertens

e.c.a.mertens@fsw.leidenuniv.nl

Reviewer A:

Review Sterk als rots, flexibel als water – R1

Zoals ik in mijn eerste review schreef is dit een waardevol onderzoek dat goed zou passen in JiO. Het gereviseerde artikel is daarenboven een duidelijke verbetering. De auteurs hebben de feedback zorgvuldig verwerkt.

Wij danken de Reviewer voor dit compliment.

Enkele zaken blijven voor mij echter nog onduidelijk en/of verdienen naar mijn mening meer aandacht.

- 1) Ik denk dat de auteurs R&W nog sterker in de verf kunnen zetten. Het is mij nog altijd niet duidelijk welke interventieprogramma's met vergelijkbare doelen er zijn en hoe R&W zich van deze interventieprogramma's onderscheidt. Een extra paragraaf in de inleiding over bestaande interventieprogramma's is nodig om de meerwaarde van R&W helder te krijgen. Er wordt ook verwezen naar "de theorie van R&W", maar de theorie wordt niet benoemd noch uitgewerkt. Het is belangrijk dat R&W niet alleen evidence-based is, maar ook theory-based.

Eén van de belangrijkste redenen om R&W te onderzoeken was niet zozeer de mogelijke superioriteit van R&W in vergelijking met andere interventieprogramma's, maar vooral het gegeven dat R&W in Nederland vaak geïmplementeerd wordt zonder dat er voldoende informatie over de effectiviteit van dit programma is; ons onderzoek is dus vooral een toets van de effectiviteit van het huidige beleid. Zoals wij in de inleiding aangeven wordt R&W bijvoorbeeld ook ingezet "binnen de kaders van de Wet Veiligheid op school" (p. 6). Een onderzoek naar de effectiviteit van dit programma in de Nederlandse context is dus belangrijk. In de inleiding benadrukken wij dit nu nog meer:

"Hoewel R&W in veel landen geïmplementeerd wordt (bijv. Frankrijk, Australië, China), en met name in Nederland vaak toegepast wordt, is er weinig bekend over de effectiviteit." (p. 6)

In de inleiding benoemen wij dat R&W gebaseerd is op de theorie van het 'R&W huis' (Ykema, 2002, 2018). Vervolgens leggen wij uit wat deze theorie inhoudt: "Dit huis is gebouwd op vier pilaren: zelfcontrole, zelfreflectie, zelfvertrouwen, en emotionele zelfregulatie. Deze vier competenties worden beschouwd als specifieke uitkomsten van de interventie en vormen de basisvaardigheden voor de ontwikkeling van de bredere, overkoepelende, competenties." (p. 6). Vervolgens koppelen wij deze theorie ook aan onze uitkomstmaten: "Volgens de theorie van R&W, zou het verbeteren van deze vier competenties samengaan met verbeteringen in psychosociaal welzijn, veerkracht, en seksuele autonomie en het verminderen van internaliserend gedrag (het intrapersoonlijk domein). Door het creëren van een veilige en respectvolle omgeving en door leerlingen alternatieve gedragingen aan te bieden zou volgens de theorie van R&W externaliserend gedrag, agressie en pesten verminderen en positieve relaties tussen leerlingen toenemen (het interpersoonlijk domein)." (p. 6).

- 2) De auteurs lijken zich sterk te maken dat R&W zich onderscheidt door goed aan te sluiten bij de leerstijl en interesse van jongeren, maar er wordt niet uitgelegd hoe R&W dit doet (zie eerste zin van p.6). Dit blijft toch een beetje een "black box" voor de lezer die R&W niet kent.

We hebben nu expliciet beschreven op welke manier R&W aansluit bij de leerstijl van jongeren.

“Tijdens de R&W lessen worden klassikale moment van informatie, instructie of reflectie afgewisseld met actieve oefeningen waarin nieuwe vaardigheden geoefend kunnen worden. De R&W trainer (de leerkracht) ondersteunt leerlingen op een positieve manier en geeft, als rolmodel, tijdens de lessen het goede voorbeeld.” (p. 6)

- 3) De validiteit van de instrumenten/vragenlijsten is een cruciaal onderdeel van de kwaliteit van de methodologie en dient dan ook besproken te worden in de methode. Het is gebruikelijk in onderzoeksartikelen dat auteurs validiteit in 1 à 2 zinnen bespreken per instrument in de methode sectie.

Wij zijn het eens met de Reviewer dat het belangrijk is om valide instrumenten te gebruiken voor onderzoek. Binnen ons onderzoek hebben wij dan ook standaard vragenlijsten gebruikt die veelvuldig toegepast worden binnen onderzoek en waarbij psychometrisch onderzoeken uitgevoerd zijn die de validiteit uitgebreid onderzocht hebben. Voor lezers die geïnteresseerd zijn in de validiteit van een bepaald instrument, achten wij het daarom zinvoller om de psychometrische artikelen hierover te lezen dan dat wij de validiteit oppervlakkig benoemen. Bovendien hebben wij 13 instrumenten gebruikt in ons onderzoek waardoor het bespreken van de validiteit per instrument ongeveer een extra pagina aan het artikel zou toevoegen en ongeveer 13 aanvullende referenties. Binnen de ruimte die wij hebben voor het rapporteren van ons onderzoek, hebben wij ervoor gekozen om details te rapporteren die uniek zijn voor ons onderzoek en nodig zijn voor replicatie in plaats van een (nog) uitgebreidere beschrijving van onze (standaard) instrumenten.

- 4) Mijn vraag hoeveel clusters er nodig zijn om kleine effecten te detecteren wordt niet beantwoord. Ik blijf mij afvragen of het aantal clusters per conditie niet te weinig is (13 scholen verdeeld over 4 condities). Mocht dat het geval zijn, dan moet dat ook benoemt worden in de discussie als een limitatie.

De analyses zijn uitgevoerd op het niveau (level) van de individuele leerlingen, zoals ook aangegeven onder punt 9. De power voor de analyses wordt daarom ook bepaald op dit niveau, niet op het schoolniveau. Met 1299 leerlingen hebben wij voldoende power om kleine effecten te detecteren.

- 5) De informatie over de leerkrachten moet gegeven worden in het onderdeel “Participanten”. Ook leerkrachten zijn immers deelnemers aan het onderzoek. Is er niet meer informatie beschikbaar dan enkel leeftijd en geslacht? Juist omdat de interventie getoetst werd in “real world” omstandigheden en de implementatie dus grotendeels afhankelijk is van de leraren(teams) is het belangrijk dat leraar kenmerken goed beschreven worden. Het viel mij ditmaal ook op dat de steekproef van scholen niet beschreven staat in “Participanten”. Schoolkenmerken en schoolpopulatiekenmerken dienen nog toegevoegd te worden aan de steekproefbeschrijving.

We hebben de beschrijving van de leerkrachten die Rots en Water hebben gegeven uitgebreid en verplaatst naar het kopje ‘participanten’. Over de scholen hebben wij geen gegevens, aangezien wij binnen het onderzoek geen reden hadden deze gegevens te verzamelen en het de kans op het identificeren van de participerende scholen zou vergroten. Wij kunnen dus geen verdere beschrijving geven van de scholen.

*“De R&W lessen werden in totaal verzorgd door 49 leerkrachten (trainers) tussen de 23 en 64 jaar oud ($M = 36,77$, $SD = 10,02$) met gemiddeld 12,98 jaar aan leerkrach-
ervaring. De meeste trainers waren gymdocent ($n = 30$; 71%). Iets meer dan de helft
van de trainers was man ($n = 27$; 56%) en bijna alle trainers hadden een Westerse
achtergrond ($n = 46$; 96%).”* (p. 9)

- 6) Daarnaast blijft het mij onduidelijk wat de “kernteams” zijn. Hoe zijn deze samengesteld? Mijns inziens is dat een heel belangrijke vraag juist omdat alleen het kernteam effectief bleek om succes te boeken met de interventie. Als ik als schooldirecteur zou overwegen om R&W te implementeren in mijn school dan zou ik graag willen weten hoe ik “mijn kernteam” moet samenstellen.

Wij danken de Reviewer voor het aanhalen van dit punt. We beschrijven de leerkrachten van het kernteam (degene die de Rots en Water lessen gaven) nu uitgebreider wat meer inzicht geeft in de samenstelling van de kernteams. Deze beschrijving kan scholen verder helpen bij het bepalen van de kernteams.

- 7) Er wordt niet gerapporteerd of de condities verschilden op de uitkomstmaten op T1. Het is wel gebruikelijk dat dit getoetst wordt in interventiestudies (zelfs al wordt hiervoor gecontroleerd door het intercept op te nemen in het model), dus ik zou dat graag terugzien in het manuscript.

We hebben dit toegevoegd in de methode sectie bij het rapporteren van de verschillen tussen condities.

“Een MANOVA liet zien dat er bij de start van de studie kleine, verwaarloosbare, verschillen tussen de condities waren op de uitkomstmaten ($F(39, 3591) = 2.12$, $p < .001$, $\eta^2_{\text{partial}} = .023$). De Light conditie scoorde enigszins slechter dan de Controle conditie op zelfcontrole, veerkracht, seksuele autonomie, internaliserende problemen, en externaliserend gedrag. Verder scoorde de Plus conditie iets beter op interpersoonlijke relaties in de klas dan de Light en Controle condities. In alle analyses is gecontroleerd voor de voormeting van de desbetreffende uitkomst, ongeacht of er bij aanvang significante verschillen waren tussen de condities.” (p. 9)

- 8) Op mijn vraag of subpopulaties anders reageerden op de interventie wordt wel een antwoord gegeven in de rebuttal, maar dit wordt niet verwerkt in het manuscript (bv. in de discussie als suggestie voor verder onderzoek, zodat lezers deze kennis ook hebben).

We hebben dit toegevoegd in de discussie sectie.

“Daarnaast zou het interessant zijn voor vervolgonderzoek om meer inzicht te krijgen in verschillen in effectiviteit voor bepaalde subgroepen (bijv. persoonlijkheid: Mertens et al., 2022).” (p. 23)

- 9) Als ik het goed begrijp is er een 1-level model geanalyseerd met correctie voor clustering. Dit moet expliciet vermeld worden in het analyse plan.

We hebben de beschrijving van onze modellen aangepast in lijn met de suggestie van de Reviewer.

“We hebben 1-level modellen geanalyseerd met een correctie voor participanten geclusterd in scholen middels de complex sample cluster in Mplus (versie 8.2; Muthén & Muthén, 1998-2010).” (p. 13)

- 10) De condities waren niet geheel gelijk bij aanvang (qua leeftijd en migratie achtergrond). Zelfs al wordt hiervoor gecontroleerd in de analyses, dan nog kan dit effect hebben gehad op het succes van de interventie. Het kan ook een indicatie zijn dat de randomisatie niet helemaal gelukt is (wat niet verrassend is als het aantal deelnemende scholen klein is en de scholen door heel NL verspreid lagen). Daarom is het mijns inziens belangrijk om dit kort te noemen bij de limitaties van de studie. Daarnaast is niet bekeken of SES ook verschilde tussen condities. Dat is ook een beperking die in de discussie kort genoemd kan worden.
We hebben de verschillen bij aanvang aangevuld als limitatie.

“Daarnaast waren er aan het begin van de studie kleine verschillen tussen de condities op uitkomstmaten wat effect gehad zou kunnen hebben op de resultaten, ondanks dat we hier in de analyses voor gecontroleerd hebben, en hebben we SES niet onderzocht.” (p. 23)

- 11) De auteurs merken terecht op dat de resultaten niet generaliseerbaar zijn naar niet-vmbo scholen. Daarnaast is het belangrijk om ook te bespreken in hoeverre deze steekproef van 13 scholen representatief is voor de vmbo scholen in NL.
We hebben de suggestie van de Reviewer verwerkt in de sectie over de limitaties van het onderzoek.

“Tot slot, in deze studie is R&W onderzocht onder een specifieke doelgroep, namelijk vmbo-leerlingen, en bij een beperkt aantal scholen verspreid over Nederland. Hoewel het belangrijk is om inzicht te krijgen in welke interventies effectief zijn voor deze doelgroep, kunnen de bevindingen niet zonder verder onderzoek gegeneraliseerd worden naar andere groepen (leerlingen) en alle vmbo-leerlingen van Nederland.” (p. 23)

Minor feedback:

- 12) Ik denk niet dat het klopt dat metacognitieve vaardigheden “nog in de eerste fase van ontwikkeling zijn”; ik zou dat iets genuanceerder formuleren.
We hebben deze zin genuanceerd.

“ (...), nog volop in ontwikkeling zijn.” (p. 21)

Reviewer B:

Dank aan de auteurs voor het zorgzaam overwegen van suggesties van de reviewers, en voor de gevraagde aanpassingen. Wat mij betreft is dit manuscript aanvaardbaar voor publicatie in JiO.

Wij danken de Reviewer voor dit compliment.

Reviewer C:

Er is een mooi artikel en goed tegemoet gekomen aan de aandachtspunten van de reviewers.

Wij danken de Reviewer voor dit compliment.

Derde versie van het manuscript - 15-5-2023

Sterk als een rots, flexibel als water? Effectiviteit van een schoolinterventie gericht op het intrapersonlijk en interpersoonlijk domein

Samenvatting

De huidige studie had als doel om te onderzoeken of 1) de schoolinterventie Rots en Water (R&W) effectief was in het verbeteren van competenties en voorkomen van problemen in het intra- en interpersoonlijk domein van leerlingen op het voorbereidend middelbaar beroepsonderwijs (vmbo), en of 2) interventie-effecten veranderden naarmate er meer mensen betrokken waren bij de interventie (kernteam van leerkrachten, het hele schoolteam, het hele schoolteam en ouders). We hebben een gerandomiseerde gecontroleerde studie uitgevoerd onder brugklassers ($N = 1299$; $M_{leeftijd} = 12,38$; 54% jongens) die vier keer digitale vragenlijsten hebben ingevuld. De resultaten lieten zien dat R&W een aantal interventie-specifieke uitkomsten (zelfcontrole en emotionele zelfregulatie), uitkomsten in het intrapersoonlijk domein (psychologisch welzijn, seksuele autonomie, en internaliserend gedrag) en in het interpersoonlijk domein (agressie) verbeterde, maar uitsluitend wanneer een kernteam van leerkrachten betrokken was bij de implementatie van de interventie. De (gematigde) interventie-effecten waren het grootste in het eerste jaar van de tweejarige interventie. Deze resultaten laten zien dat R&W positieve effecten kan bewerkstelligen bij vmbo-leerlingen wanneer deze geïmplementeerd wordt door een kernteam. Het lijkt dus waardevol voor scholen om te investeren in interventies met een leerstijl die aansluit bij die van de doelgroep, met een hoge kwaliteit van implementatie en een beperkt aantal betrokken personen.

Abstract

The aim of the present study was to examine 1) the effectiveness of the school-based intervention Rock and Water (R&W) in stimulating competencies and preventing problems in the intra- and interpersonal domains of prevocational students, and 2) whether intervention effects changed as more people were involved in the intervention (core team of teachers, whole school, and whole school and parents). We have conducted a Randomized Controlled Trial among 7th Grade students ($N = 1299$, $M_{age} = 12.38$, 54% boys) who completed four digital questionnaires. The results showed that R&W improved some intervention specific outcomes (self-control and emotional self-regulation), outcomes in the intrapersonal domain (psychological wellbeing, sexual autonomy, and internalizing behavior) and in the interpersonal domain (aggression), but only when solely a core team of teachers was involved in the implementation of the intervention. The (moderate) intervention effects were strongest in the first year of the two-year intervention. These results show that R&W can establish positive effects among prevocational students when implemented by a core team. Thus, it appears worthwhile for schools to invest in interventions that fit with the learning style of the

target group, with a high level of implementation quality and involving a limited number of people.

Trefwoorden: Schoolinterventie; Rots en Water; Intrapersoonlijk domein; Interpersoonlijk domein; Vmbo-leerlingen

Bijsluiter voor de praktijk

1. Vmbo-leerlingen rapporteren meer problemen in het intra- en interpersoonlijk domein dan havo- en vwo-leerlingen.
2. Wij onderzochten: 1) de effectiviteit van de schoolinterventie Rots en Water voor het stimuleren van het intra- en interpersoonlijk domein bij vmbo-leerlingen en 2) of de effectiviteit veranderde naarmate verschillende mensen betrokken waren bij de interventie (kernteam leerkrachten, hele schoolteam, hele schoolteam en ouders).
3. De interventie liet gematigde positieve effecten zien, met name op uitkomsten in het intrapersoonlijk domein, wanneer alleen een kernteam van leerkrachten betrokken was bij de uitvoering van de interventie.
4. Interventie-effecten waren het grootste in het eerste jaar van de interventie.
5. Het lijkt waardevol om te investeren in goede training voor en kwalitatief hoogstaande implementatie door een kernteam.

Aantal woorden:

5818 (hoofdtekst)

Introductie

Scholen spelen, naast andere factoren zoals gezin, leeftijdsgenoten en maatschappelijke omstandigheden, een belangrijke rol in het bevorderen van de ontwikkeling van leerlingen in het intrapersoonlijk en interpersoonlijk domein (Pellegrino & Hilton, 2012). Het *intrapersoonlijk domein* refereert naar subjectief persoonlijk functioneren (bijv. psychologisch welzijn) en is gerelateerd aan schoolresultaten en zelfregulatie. Het *interpersoonlijk domein*, daarentegen, refereert naar sociaal functioneren (bijv. agressie) en is onder andere gerelateerd aan relaties met leeftijdsgenoten en conflict oplossend vermogen (Park et al., 2017). Leerlingen kunnen in deze domeinen competenties ontwikkelen door relevante cognitieve, affectieve en sociale vaardigheden te leren (bijv. identificeren van emoties en perspectief wisselen). Wanneer het verwerven van (sommige) vaardigheden achterwege blijft, kunnen leerlingen problemen ontwikkelen (Modecki et al., 2017). Hoewel de twee domeinen elkaar beïnvloeden, worden ze als apart beschouwd. Scholen zouden bewust kunnen proberen de competenties van leerlingen op beide domeinen te stimuleren om mogelijke problemen te voorkomen (Pellegrino & Hilton, 2012).

Met name leerlingen op het voorbereidend middelbaar beroepsonderwijs (vmbo) zouden baat kunnen hebben bij stimulering van competenties en het voorkomen van problemen. Zij rapporteren, vergeleken met leerlingen op het hoger algemeen voortgezet onderwijs (havo) en voorbereidend wetenschappelijk onderwijs (vwo), over het algemeen een lager niveau van welzijn, meer gedragsproblemen, en meer problemen met leeftijdsgenoten (Stevens & De Looze, 2018). Tegelijkertijd kunnen vmbo-leerlingen een uitdagende groep zijn om te betrekken bij interventielessen op school. Over het algemeen laten deze leerlingen minder intrinsieke motivatie voor school en (verbaal) leren zien dan havo- en vwo-leerlingen (Timmermans et al., 2017). Hoewel schoolinterventies positieve effecten laten zien (e.g., Mertens et al., 2020), is het onduidelijk in hoeverre deze ook effectief zijn voor specifiek vmbo-leerlingen; vooral voor het betrekken van deze groep leerlingen bij een schoolinterventie lijkt het namelijk belangrijk een interventie te implementeren die goed aansluit bij hun interesses en leerstijl. Vmbo-leerlingen worden het beste bereikt en blijven het meest gemotiveerd en betrokken met een praktijkgerichte leer methode. Deze leer methode houdt in dat er een leeromgeving gecreëerd wordt met een praktijkgerichte insteek met activerende, stapsgewijze, gestructureerde taken; leerlingen kunnen hun nieuw geleerde vaardigheden toepassen en eigen doelen stellen. Tijdens de les worden korte momenten van instructie of reflectie afgewisseld met het toepassen van nieuwe vaardigheden en leerkrachten functioneren als mentor en rolmodel (De Bruijn et al., 2005; Smit et al., 2014).

Een universele schoolinterventie die aansluit bij deze actieve leerstijl van vmbo-leerlingen is Rots en Water (R&W; Ykema, 2002, 2018; <https://www.rotsenwater.nl>). Tijdens de R&W lessen worden klassikale moment van informatie, instructie of reflectie afgewisseld met actieve oefeningen waarin nieuwe vaardigheden geoefend kunnen worden. De R&W trainer (de leerkracht) ondersteunt leerlingen op een positieve manier en geeft, als rolmodel, tijdens de lessen het goede voorbeeld. R&W is gebaseerd op de theorie van het 'R&W huis' (Ykema, 2002, 2018). Dit huis is gebouwd op vier pilaren: zelfcontrole, zelfreflectie, zelfvertrouwen, en emotionele zelfregulatie. Deze vier competenties worden beschouwd als specifieke uitkomsten van de interventie en vormen de basisvaardigheden voor de ontwikkeling van de bredere, overkoepelende, competenties. Volgens de theorie van R&W, zou het verbeteren van deze vier competenties samengaan met verbeteringen in psychosociaal welzijn, veerkracht, en seksuele autonomie en het verminderen van internaliserend gedrag (het intrapersoonlijk domein). Door het creëren van een veilige en respectvolle omgeving en door leerlingen alternatieve gedragingen aan te bieden zou volgens de theorie van R&W externaliserend gedrag, agressie en pesten verminderen en positieve relaties tussen leerlingen toenemen (het interpersoonlijk domein).

Hoewel R&W in veel landen geïmplementeerd wordt (bijv. Frankrijk, Australië, China), en met name in Nederland vaak toegepast wordt, is er weinig bekend over de effectiviteit. Met name in Nederland is het belangrijk meer inzicht te krijgen in de effectiviteit van R&W, niet alleen omdat deze interventie vaak wordt geïmplementeerd in Nederland, maar ook omdat het ingezet wordt binnen de kaders van de Wet Veiligheid op school – een wet die scholen verplicht zich actief in te spannen om pesten tegen te gaan en een veilig sociaal schoolklimaat te creëren. In Australië hebben professionals in het werkveld verschillende kleinschalige studies uitgevoerd, vaak op basis- en middelbare scholen, die hebben aangetoond dat deelnemers zich na het afronden van R&W veerkrachtiger voelden, een meer positieve identiteit hadden, en meer actieve copingstrategieën toepasten (Ykema et al., 2006). Een onderzoek in Nederland onder (mannelijke) vmbo-leerlingen van De Graaf en collega's (2016) heeft de effectiviteit van R&W op seksuele agressie onderzocht. Na het afronden van R&W verminderden zelfgerapporteerde dwingende strategieën en verbale manipulatie, terwijl zelfregulatie toenam. Ondanks deze veelbelovende resultaten is een grootschalige studie nodig die zich richt op bredere uitkomstmaten, zowel jongens als meisjes includeert, en in het bijzonder de effectiviteit onderzoekt voor vmbo-leerlingen. Daarom was het eerste doel van de huidige studie om te onderzoeken of R&W effectief is voor vmbo-leerlingen in het verbeteren van competenties en het voorkomen van problemen in het

intrapersoonlijk en het interpersoonlijk domein. Verwacht werd dat leerlingen in de R&W conditie een significante verbetering zouden laten zien op alle gemeten interventie-specifieke en overkoepelende (intra- en interpersoonlijke) uitkomsten ten opzichte van leerlingen in de controle conditie.

Het tweede doel was om te onderzoeken of interventie-effecten afhankelijk zijn van de mate waarin verschillende systemen betrokken zijn bij de interventie. Volgens het ecologisch model van Bronfenbrenner (1979), wordt gedrag bepaald door interacties tussen verschillende systemen zoals individu-, familie-, klas-, en schoolsystemen. Het actief betrekken van gerelateerde systemen bij een interventie, zoals de school- en familiesystemen, zou mogelijk zowel generalisatie als behoud van nieuwe vaardigheden kunnen bevorderen (Mahoney et al., 2021); de inhoud van de interventie wordt dan namelijk ingebed in een groot deel van de sociale omgeving van het individu wat de mogelijke invloed ervan op het individu vergroot. Zo concludeerden Mertens en collega's (2020) in hun meta-analyse naar componenten van universele schoolinterventies dat betrokkenheid van de volledige school en/of ouders gerelateerd was aan sterkere interventie-effecten. Echter, Durlak en collega's (2011) vonden in hun meta-analyse van *social emotional learning* (SEL) interventies dat interventies waarbij de hele school en/of ouders betrokken waren niet effectiever waren dan interventies die alleen in de klas gegeven werden. Deze inconsistente bevindingen geven aan dat er meer inzicht nodig is in hoe betrokkenheid van meerdere systemen in een interventie de effectiviteit kan beïnvloeden (Kuosmanen et al., 2019).

De huidige studie onderzocht daarom of er verschil is in de effectiviteit van drie verschillende implementatieversies van R&W: betrokkenheid van 1) een kernteam van leerkrachten, 2) een kernteam en het volledige schoolteam, en 3) een kernteam, het volledige schoolteam en ouders. We hadden geen specifieke hypothese aangezien onderzoek naar de betrokkenheid van meerdere ecologische systemen in interventies inconsistente resultaten laat zien.

Methode

Verschillende versies van R&W

We hebben drie versies van R&W onderzocht die verschilden in het aantal systemen dat betrokken was bij de interventie: 1) *Light*: alleen een kernteam van leerkrachten (die de interventielessen gaven), 2) *Standaard*: een kernteam van leerkrachten en de rest van het schoolteam, en 3) *Plus*: een kernteam, het schoolteam, en ouders. In alle drie de condities werden de interventielessen gegeven door een kernteam van leerkrachten (R&W trainers) die de 3-daagse training succesvol hadden afgerond. Dit kernteam ontving supervisie van een

persoonlijke R&W coach. In de Light conditie was alleen een kernteam getraind en op de hoogte van R&W. In de Standaard en de Plus condities had, naast het kernteam van R&W trainers, de rest van het schoolteam een 3-daagse introductie cursus gevolgd over de basisprincipes van R&W, ondersteuning van R&W trainers, en toepassing van de interventietechnieken in reguliere lessen. In de Plus conditie kregen daarnaast ouders de mogelijkheid een documentaire over R&W te bekijken, werden ze uitgenodigd voor een R&W les op school, en ontvingen ze wekelijks e-mails met daarin informatie over de R&W les van die week en werden ze aangemoedigd om met deze informatie thuis aan de slag te gaan.

Design en procedure

De effectiviteit van deze drie verschillende versies van R&W (Light, Standaard en Plus) is onderzocht in scholen met een vmbo-traject in een gerandomiseerde gecontroleerde studie met een actieve controle conditie. In de Controle conditie voerden scholen hun gebruikelijk beleid voor het verbeteren van competenties en het voorkomen van problemen bij leerlingen. Scholen werden geëxcludeerd als ze in de afgelopen twee jaar R&W hadden geïmplementeerd. Dertien scholen verspreid door Nederland werden willekeurig toegewezen aan de condities door gestratificeerde blok randomisatie met blokken van vier. Scholen werden gestratificeerd op schoolgrootte (< 100 leerlingen in de eerste klas, > 100 leerlingen in de eerste klas) om een gelijke verdeling van leerlingen over de condities te bevorderen. Eén school uit de Controle conditie viel uit na randomisatie maar voor het begin van de dataverzameling vanwege een verandering in schoolmanagement. Deze school werd vervangen door een andere geïnteresseerde school (zie Figuur 1 voor de flowchart).

[Figuur 1]

Leerlingen vulden digitale vragenlijsten in, afgenomen door getrainde onderzoeksassistenten, voor de start van de interventie (T1 = oktober/november 2017), na het afronden van het eerste jaar van de R&W lessen (T2 = maart/april 2018), voor de start van de R&W lessen in het tweede jaar van het R&W programma (T3 = oktober 2018), en na de interventie (T4 = januari 2019). Leerlingen gaven actieve, geïnformeerde toestemming voor het invullen van de vragenlijsten. Ouders gaven passieve, geïnformeerde toestemming voor deelname van hun kind. Deze studie is goedgekeurd door de Ethische Commissie [GEMASKEERD VOOR REVIEW] en geregistreerd [GEMASKEERD VOOR REVIEW] (voor protocol zie [GEMASKEERD VOOR REVIEW]).

Tabel 1

Demografische gegevens van leerlingen op de voormeting per conditie

	Light	Standaard	Plus	Controle	Verschillen op T1		
					F/χ^2	p	$\eta^2_{\text{partial}}/\phi$
<i>N</i>	373	303	249	374			
Leeftijd, <i>M</i> (<i>SD</i>)	12,33 (0,57)	12,38 (0,66)	12,34 (0,60)	12,47 (0,64)	3,89	,009	0,009
Jongens, <i>n</i> (%)	170 (48%)	161 (56%)	131 (55%)	199 (57%)	7,38	,061	0,077
Westerse achtergrond, <i>n</i> (%)	291 (82%)	115 (43%)	211 (91%)	198 (59%)	182,01	< ,001	0,392

Participanten

Op de voormeting bestond de steekproef uit 1299 eerstejaars vmbo-leerlingen ($M_{\text{leeftijd}} = 12,38$ jaar; $SD = 0,62$; 54% jongens; 69% met een Westerse achtergrond; zie Tabel 1).

Tussen de condities waren geen verschillen in sekse, maar wel in leeftijd en migratieachtergrond. Leerlingen in de Controle conditie waren ouder dan leerlingen in de Light conditie. De Controle en Standaard condities hadden ongeveer een gelijke verdeling tussen leerlingen met een Westerse en een niet-Westerse achtergrond, terwijl in de Light en Plus condities leerlingen vooral een Westerse achtergrond hadden. Daarom werd in alle analyses gecontroleerd voor leeftijd en migratieachtergrond. Een MANOVA liet zien dat er bij de start van de studie kleine, verwaarloosbare, verschillen tussen de condities waren op de uitkomstmaten ($F(39, 3591) = 2.12, p < .001, \eta^2_{\text{partial}} = .023$). De Light conditie scoorde enigszins slechter dan de Controle conditie op zelfcontrole, veerkracht, seksuele autonomie, internaliserende problemen, en externaliserend gedrag. Verder scoorde de Plus conditie iets beter op interpersoonlijke relaties in de klas dan de Light en Controle condities. In alle analyses is gecontroleerd voor de voormeting van de desbetreffende uitkomst, ongeacht of er bij aanvang significante verschillen waren tussen de condities.

Over alle meetmomenten ontbrak 15% van de data. Vergelijkingen tussen leerlingen die uitgevallen zijn ($n_{T2} = 68, n_{T3} = 83, n_{T4} = 60$) en leerlingen die de studie volledig hebben afgerond, lieten geen verschillen zien op demografische kenmerken (leeftijd, sekse verdeling en migratieachtergrond) noch op de uitkomstvariabelen.

De R&W lessen werden in totaal verzorgd door 49 leerkrachten (trainers) tussen de 23 en 64 jaar oud ($M = 36,77, SD = 10,02$) met gemiddeld 12,98 jaar aan leerkracht-ervaring. De meeste trainers waren gymdocent ($n = 30; 71\%$). Iets meer dan de helft van de trainers was man ($n = 27; 56\%$) en bijna alle trainers hadden een Westerse achtergrond ($n = 46; 96\%$).

Conditie en Uitvoering

R&W

R&W (Ykema, 2002, 2018) integreert spel en opdrachten om leerlingen te begeleiden in het maken van (fysiek) contact met anderen en het ontdekken, respecteren en stellen van zowel eigen grenzen als grenzen van anderen. Middels oefeningen worden lichaamsbewustzijn (bijv. spanning in de spieren, ademhaling), non-verbale communicatie, en copingstrategieën van leerlingen versterkt. In een oefening in tweetallen moeten leerlingen bijvoorbeeld proberen de andere leerling uit balans te tikken met als doel leerlingen te laten ervaren dat het soms beter is om rigide te zijn en druk van anderen te weerstaan (als een rots) en dat het soms beter is om flexibel te zijn en open te staan voor de mening, gedachten, en gevoelens van anderen (als water).

Naast deze spellen hanteert R&W ook een gedragsmatige aanpak, zoals het gebruik van rolspellen om te oefenen met bijvoorbeeld het stellen en aangeven van grenzen, kalm reageren op provocaties, en het helpen van leerlingen die gepest worden. Tijdens en na de oefeningen geven leerlingen feedback en reflecteren ze. Trainers tonen en bekrachtigen gewenst gedrag, en creëren een veilige omgeving waarin leerlingen fouten mogen maken.

De interventie bestaat uit 22 lessen verspreid over twee schooljaren: 14 lessen in het eerste en 8 lessen in het tweede leerjaar. De lessen werden wekelijks gegeven tijdens gymlessen van 90 minuten.

Controle conditie

In de Controle conditie voerden scholen hun gebruikelijk beleid. Leerlingen hadden bijvoorbeeld een mentor of persoonlijke coach met wie ze regelmatig afspraken hadden, hun welzijn bespraken en naar wie ze toe konden voor advies. Ook was er bijvoorbeeld een projectweek over ‘anders zijn’, tekenden leerlingen een anti-pestcontract, werd pesten in de klas besproken en was er een ‘anti-pestcoördinator’.

Uitvoering interventie

De uitvoering van R&W is gemeten op twee manieren. Ten eerste, R&W trainers vulden een vragenlijst in na elke derde les in het eerste jaar en elke tweede les in het tweede jaar van de interventie. Deze vragenlijst bestond uit vragen over het verloop van de R&W lessen, zoals “Heeft u in de afgelopen drie/twee lessen van R&W alles kunnen behandelen?” (1 = *alle lessen volledig kunnen geven* tot 4 = *geen les volledig kunnen geven*) en “In hoeverre bent u afgeweken van het R&W handboek?” (1 = *bijna niet* tot 6 = *heel veel*). Ten tweede, drie R&W experts, die alle drie de R&W training voor gevorderden hebben afgerond en

meerdere jaren ervaring hebben met het implementeren van R&W, hebben 67 R&W lessen (Light = 28 lessen; Standaard = 17 lessen, Plus = 22 lessen) geobserveerd. Tijdens deze lesobservatie hebben zij een schema ingevuld, gebaseerd op Bishop en collega's (2014), waarin ze onder andere hun subjectieve beoordeling gaven over de mate waarin de handleiding werd gevolgd, aanpassingen die werden gedaan en de kwaliteit van de les.

Volgens de zelfrapportages hebben R&W trainers over het algemeen de lessen volledig kunnen geven (65%) en zijn ze niet of nauwelijks afgeweken van de handleiding (72%). De R&W experts observeerden dat de meeste lessen volledig of bijna volledig uitgevoerd waren (86%). Trainers weken niet veel af van de handleiding (91%). Eventuele aanpassingen aan de uitvoering werden over het algemeen gezien als verbeteringen (62%). Over het algemeen beoordeelden de R&W experts de lessen van goede (38%) tot zeer goede (54%) kwaliteit. De uitvoering van de R&W lessen verschilde niet tussen condities noch tussen de twee jaar van implementatie, zowel volgens de zelfrapportages als volgens de observaties. Concluderend, het grootste deel van de R&W lessen was daadwerkelijk gegeven en de uitvoering lijkt goed tot zeer goed te zijn geweest.

Om te bepalen of de Standaard en Plus versies geïmplementeerd waren zoals bedoeld, hebben R&W trainers en ouders vragenlijsten ingevuld na het eerste jaar en na het tweede jaar van de interventie. Zoals gepland waren er in de Standaard en Plus condities meer leerkrachten betrokken bij de interventie dan in de Light conditie ($F(2,24) = 5,73; p = ,009$). Ouders in de Plus conditie die de vragenlijst hebben ingevuld ($n = 47$) rapporteerden dat zij de wekelijkse informatie soms (57%) tot vaak (39%) lazen, en sommige ouders gaven aan de documentaire over R&W gekeken te hebben (15%) en deelgenomen te hebben aan een R&W les op school (23%). Verder rapporteerden ouders in de Plus conditie, vergeleken met ouders in de Standaard conditie, thuis vaker over R&W gesproken te hebben ($F(2,75) = 4,37; p = ,016$) en het vaker bij hun eigen kind toegepast te hebben ($F(2,75) = 3,44; p = ,037$). Kortom, ouders in de Plus conditie waren zoals gepland meer betrokken bij de interventie dan ouders in de Light en Standaard condities. Dus, zowel de leerkracht- als de oudercomponent blijkt geïmplementeerd te zijn.

Interventie-Specifieke Uitkomsten

Zelfcontrole

De vaardigheid van leerlingen om hun impulsen te controleren en ongepast gedrag te onderbreken werd gemeten met de korte versie van de *Self-Control Scale* (Finkenauer et al., 2005) die bestaat uit 11 items (bijv. "Ik kan goed weerstand bieden tegen verleidingen", $\alpha = 0,62-0,72$) beantwoord op een 5-puntsschaal ($1 = \text{helemaal niet}$ tot $5 = \text{helemaal wel}$).

Zelfreflectie

Leerlingen vulden de *Engage* subschaal in van de *Self-Reflection and Insight Scale* (Sauter et al., 2010) om de mate van reflectie op persoonlijke gedachten, gevoelens en gedrag te meten. De subschaal bestaat uit 6 items (bijv. “Ik denk vaak na over hoe ik me voel over iets”) beantwoord op een 6-puntsschaal (1 = *helemaal mee oneens* tot 6 = *helemaal mee eens*). De betrouwbaarheid was laag ($\alpha = 0,53-0,60$). Verwijdering van 3 negatief geformuleerde items (“Ik ben bijna nooit bezig met ‘zelfreflectie’”, “Ik denk niet vaak na over mijn gedachten”, en “Ik denk niet echt na over de reden waarom ik me gedraag zoals ik dat doe”) verbeterde de betrouwbaarheid ($\alpha = 0,74-0,89$).

Zelfvertrouwen

Globaal zelfvertrouwen van leerlingen is gemeten met de subschaal *Global self-perception* van de *Self-Perception Profile* (Harter, 1988). Deze subschaal bestaat uit 5 items (bijv. “Ik ben best tevreden met mezelf”) beantwoord op een 4-puntsschaal (1 = *helemaal niet waar* tot 4 = *helemaal waar*; $\alpha = 0,73-0,75$).

Emotionele zelfregulatie

Vaardigheden in het controleren van emoties en het gebruik van emotieregulatie strategieën werd gemeten met de *Difficulties in Emotion Regulation Scale* (Anderson et al., 2016). De vragenlijst bestaat uit 14 items (bijv. “Als ik van streek ben, kan ik iets bedenken om ervoor te zorgen dat ik me weer beter voel”) beantwoord op een 5-puntsschaal (1 = *bijna nooit* tot 5 = *bijna altijd*; $\alpha = 0,88-0,91$).

Overkoepelende Uitkomsten

Intrapersoonlijk domein

Psychologisch welzijn. Om de aanwezigheid van positieve emoties te meten, hebben leerlingen de subschaal *Psychological Wellbeing* ingevuld van de KIDSCREEN-27 (Ravens-Sieberer & The European KIDSCREEN Group, 2006). De subschaal bestaat uit 7 items (bijv. “In de afgelopen week, heb je lol gehad?”) beantwoord op een 5-puntsschaal (1 = *nooit* tot 5 = *altijd*; $\alpha = 0,76-0,83$).

Veerkracht. De vaardigheid om terug te veren na uitdagingen werd gemeten middels de *Connor-Davidson Resilience Scale-Short Version* (Davidson & Connor, 2017) bestaande uit 10 items (bijv. “Aanpassen aan verandering”) beantwoord op een 5-puntsschaal (0 = *helemaal niet waar* tot 4 = *bijna altijd waar*; $\alpha = 0,79-0,92$).

Seksuele autonomie. Coping vaardigheden in seksuele situaties werd gemeten met 5 items van de *Seks onder je 25^{ste}* vragenlijst (bijv. “Wanneer ik ben met iemand die ik heel erg

leuk vind, voel ik me helemaal op mijn gemak”; De Graaf et al., 2005) beantwoord op een 4-puntsschaal ($1 = \text{nooit}$ tot $4 = \text{altijd}$). Betrouwbaarheid was laag op T1 ($\alpha = 0,53$) en redelijk op T2, T3, en T4 ($\alpha = 0,62-0,65$).

Internaliserend gedrag. Leerlingen vulden de internaliserende subschaal in van de korte versie van de *Youth Self Report* (Chorpita et al., 2010). De subschaal bestaat uit 6 items (bijv. “Ik maak me vaak zorgen”) beantwoord op een 3-puntsschaal ($0 = \text{nooit}$ tot $2 = \text{vaak}$; $\alpha = 0,79-0,86$).

Interpersoonlijk domein

Interpersoonlijke relaties in de klas. De manier waarop leerlingen relaties in de klas ervaren, werd gemeten met de subschalen Conflict, Cohesie en Comfort van de *Classroom Peer Context Questionnaire* (Boor-Klip et al., 2016). Deze 12 items (bijv. “In deze klas vinden kinderen elkaar aardig”) werden beantwoord op een 5-puntsschaal ($1 = \text{helemaal niet waar}$ tot $5 = \text{helemaal waar}$; $\alpha = 0,80-0,85$).

Externaliserend gedrag. Leerlingen hebben de externaliserende subschaal ingevuld van de korte versie van de *Youth Self Report* (Chorpita et al., 2010). De subschaal bestaat uit 6 items (bijv. “Ik maak veel ruzie”) beantwoord op een 3-puntsschaal ($0 = \text{nooit}$ tot $2 = \text{vaak}$; $\alpha = 0,65-0,79$).

Agressie. Agressie is gemeten met de *Reactive and Proactive Aggression Questionnaire* (Dodge & Coie, 1987). De vragenlijst bestaat uit 6 items (bijv. “Als ze mij plagen of tegen mij liegen, reageer ik kwaad”) beantwoord op een 5-puntsschaal ($1 = \text{nooit}$ tot $5 = \text{bijna altijd}$; $\alpha = 0,65-0,83$).

Pesten en gepest worden. De frequenties van pesten en ervaringen van gepest worden zijn gemeten met 2 items van de *Olweus Bully/Victim Questionnaire* (Solberg & Olweus, 2003): “Hoe vaak heb je de afgelopen twee maanden meegedaan aan het pesten van andere leerlingen op school?” en “Hoe vaak ben je de afgelopen twee maanden op school gepest?”. Leerlingen beantwoordden deze vragen op een 5-puntsschaal ($1 = \text{nooit}$ tot $5 = \text{meerdere keren per week}$).

Statistische Analyses¹

Data zijn geanalyseerd op basis van een *intention-to-treat* benadering waarbij leerlingen toegewezen aan de interventie geïnccludeerd zijn in de analyse ongeacht of ze daadwerkelijk hebben deelgenomen aan de interventie. We hebben 1-level-modellen geanalyseerd met een correctie voor participanten geclusterd in scholen middels de *complex*

¹ Uit de power analyse in het protocol paper (zie GEMASKEERD VOOR REVIEW) blijkt een minimale steekproefgrootte van 720 participanten nodig voor deze analyses.

sample cluster in Mplus (versie 8.2; Muthén & Muthén, 1998-2010). Om alle participanten in de modellen te kunnen includeren, hebben we *Full Information Maximum Likelihood* (FIML) procedures gebruikt. Parameters werden verkregen met *Robust Maximum Likelihood* schattingen (MLR; Muthén & Muthén, 1998-2010).

Om de effecten van de interventie te analyseren, hebben we een serie latente groeicurve (LGC) modellen getest. LGC modellen schatten voor elke participant een individueel groeitraject gebaseerd op hun basisniveau (= intercept) en verandering over tijd (= slope). De individuele groeicurves zijn de indicatoren van latente variabelen die de groeitrajecten van de groep beschrijven (Muthén & Muthén, 1998-2010). De slope is het belangrijkste; wanneer de interventie vergeleken met de Controle conditie effectief is, verandert de interventie de slope in de gewenste richting.

Voor het bepalen van de interventie-effecten, zijn drie dummy variabelen gemaakt (Light, Standaard, en Plus conditie) met de Controle conditie als referentiegroep en toegevoegd als voorspellers van de intercept en slope. Leeftijd en migratieachtergrond van de leerlingen zijn toegevoegd als covariaten. Wanneer twee of meer interventiecondities vergeleken met de Controle conditie effectief bleken, zijn de condities onderling vergeleken met een multi-groep model door de slopes aan elkaar gelijk te stellen en vrij te schatten. De model fits van deze twee geneste modellen zijn vergeleken met de *Satorra-Bentler Scale Chi-Square test* (Satorra & Bentler, 2010). Wanneer deze significant is, past het vrij geschatte model beter bij de data en is dus de ene interventieconditie effectiever dan de andere.

We hebben effectgroottes berekend door de mate van verandering te vermenigvuldigen met de tijdsspanne van de interventie en dat vervolgens gedeeld door de standaarddeviatie van de desbetreffende uitkomst ($d = (\text{slope} * \text{tijdsspanne}) / SD$; Feingold, 2013). We hebben effectgroottes berekend tussen de verschillende meetmomenten (verandering van T1 naar T2, van T2 naar T3, en van T3 naar T4) en de algemene verandering (verandering van T1 naar T4).

Resultaten

Effecten van R&W

In Tabel 2 staan de gemiddelden en standaarddeviaties van de uitkomsten per conditie op elk meetmoment. De LGC modellen hadden in het algemeen een acceptabele fit (zie voor details [GEMASKEERD VOOR REVIEW]).

De gestandaardiseerde regressiecoëfficiënten van de interventie condities op de slope zijn gerapporteerd in Tabel 3. De effectgroottes tussen de verschillende meetmomenten en de

algemene effectgroottes van de interventie condities (in vergelijking met de Controle conditie) zijn gerapporteerd per uitkomst in Tabel 4.

Tabel 2

Gemiddelden en standaarddeviaties van uitkomsten per conditie per meetmoment

	Light				Standaard				Plus				Controle			
	T1	T2	T3	T4	T1	T2	T3	T4	T1	T2	T3	T4	T1	T2	T3	T4
	<i>M</i> (<i>SD</i>)	<i>M</i> (<i>SD</i>)	<i>M</i> (<i>SD</i>)	<i>M</i> (<i>SD</i>)	<i>M</i> (<i>SD</i>)	<i>M</i> (<i>SD</i>)	<i>M</i> (<i>SD</i>)	<i>M</i> (<i>SD</i>)	<i>M</i> (<i>SD</i>)	<i>M</i> (<i>SD</i>)	<i>M</i> (<i>SD</i>)	<i>M</i> (<i>SD</i>)	<i>M</i> (<i>SD</i>)	<i>M</i> (<i>SD</i>)	<i>M</i> (<i>SD</i>)	<i>M</i> (<i>SD</i>)
<i>Interventie-specifieke uitkomsten</i>																
Zelfcontrole	3,33 (0,58)	3,53 (0,60)	3,52 (0,61)	3,54 (0,60)	3,49 (0,57)	3,66 (0,66)	3,56 (0,65)	3,61 (0,69)	3,41 (0,60)	3,53 (0,59)	3,43 (0,61)	3,53 (0,65)	3,49 (0,56)	3,59 (0,62)	3,57 (0,65)	3,58 (0,64)
Zelfreflectie	2,92 (1,20)	2,83 (1,32)	2,80 (1,27)	2,76 (1,36)	3,02 (1,31)	2,83 (1,31)	2,76 (1,32)	2,90 (1,42)	2,90 (1,22)	2,96 (1,33)	2,85 (1,36)	2,81 (1,52)	3,09 (1,21)	2,95 (1,44)	2,98 (1,35)	2,86 (1,43)
Zelfvertrouwen	3,23 (0,62)	3,28 (0,66)	3,24 (0,65)	3,28 (0,62)	3,28 (0,57)	3,33 (0,60)	3,29 (0,66)	3,29 (0,62)	3,25 (0,59)	3,33 (0,57)	3,18 (0,68)	3,12 (0,70)	3,31 (0,59)	3,29 (0,64)	3,22 (0,65)	3,22 (0,65)
Emotionele zelfregulatie	3,83 (0,77)	3,94 (0,85)	3,98 (0,76)	3,97 (0,75)	3,85 (0,69)	3,99 (0,80)	3,91 (0,82)	3,91 (0,84)	3,87 (0,71)	3,88 (0,74)	3,81 (0,80)	3,83 (0,85)	3,93 (0,71)	3,94 (0,75)	3,90 (0,76)	3,92 (0,76)
<i>Intrapersoonlijk domein</i>																
Psychologisch welzijn	3,94 (0,62)	3,98 (0,69)	3,91 (0,71)	3,87 (0,78)	4,02 (0,61)	4,10 (0,71)	4,02 (0,73)	3,93 (0,77)	3,97 (0,62)	4,06 (0,63)	3,87 (0,72)	3,85 (0,80)	4,06 (0,59)	4,02 (0,68)	3,90 (0,75)	3,92 (0,78)
Veerkracht	2,36 (0,65)	2,49 (0,75)	2,15 (0,87)	2,47 (0,82)	2,43 (0,67)	2,70 (0,71)	2,56 (0,70)	2,53 (0,90)	2,43 (0,64)	2,57 (0,68)	2,47 (0,71)	2,51 (0,84)	2,51 (0,65)	2,68 (0,69)	2,57 (0,74)	2,61 (0,85)
Seksuele autonomie	3,23 (0,44)	3,39 (0,48)	3,40 (0,44)	3,40 (0,46)	3,30 (0,44)	3,44 (0,52)	3,38 (0,54)	3,39 (0,51)	3,29 (0,40)	3,38 (0,46)	3,34 (0,45)	3,35 (0,57)	3,32 (0,44)	3,36 (0,49)	3,32 (0,50)	3,36 (0,52)
Internaliserend gedrag	0,53 (0,42)	0,41 (0,44)	0,44 (0,44)	0,42 (0,46)	0,44 (0,42)	0,33 (0,39)	0,40 (0,44)	0,37 (0,43)	0,49 (0,40)	0,45 (0,42)	0,51 (0,44)	0,48 (0,51)	0,43 (0,39)	0,39 (0,41)	0,45 (0,46)	0,42 (0,44)
<i>Interpersoonlijk domein</i>																

Interpersoonlijke relaties in de klas	3,90 (0,64)	3,81 (0,68)	3,83 (0,67)	3,73 (0,69)	3,98 (0,63)	3,87 (0,72)	3,81 (0,73)	3,80 (0,75)	4,07 (0,62)	4,00 (0,63)	3,92 (0,69)	3,84 (0,81)	3,91 (0,59)	3,78 (0,65)	3,77 (0,67)	3,75 (0,72)
Externaliserend gedrag	0,59 (0,38)	0,47 (0,39)	0,48 (0,39)	0,45 (0,41)	0,54 (0,39)	0,43 (0,38)	0,50 (0,44)	0,42 (0,42)	0,55 (0,35)	0,46 (0,36)	0,55 (0,38)	0,49 (0,45)	0,49 (0,34)	0,42 (0,37)	0,45 (0,39)	0,43 (0,39)
Agressie	1,79 (0,56)	1,71 (0,59)	1,69 (0,59)	1,66 (0,66)	1,85 (0,61)	1,84 (0,73)	1,94 (0,73)	1,83 (0,81)	1,73 (0,46)	1,72 (0,60)	1,80 (0,58)	1,74 (0,81)	1,78 (0,56)	1,76 (0,65)	1,81 (0,67)	1,85 (0,78)
Pesten	1,13 (0,48)	1,18 (0,61)	1,15 (0,55)	1,21 (0,64)	1,13 (0,41)	1,18 (0,60)	1,15 (0,58)	1,19 (0,64)	1,10 (0,41)	1,15 (0,47)	1,20 (0,68)	1,20 (0,64)	1,08 (0,36)	1,18 (0,59)	1,22 (0,71)	1,22 (0,71)
Gepest worden	1,32 (0,77)	1,43 (0,96)	1,30 (0,77)	1,31 (0,86)	1,28 (0,73)	1,35 (0,87)	1,22 (0,77)	1,22 (0,70)	1,36 (0,88)	1,37 (0,85)	1,40 (0,96)	1,40 (0,93)	1,37 (0,94)	1,56 (1,16)	1,41 (1,04)	1,34 (0,95)

Tabel 3

Interventie-effecten over tijd van de R&W condities in vergelijking met de controle conditie

	Light			Standaard			Plus		
	β	SE	<i>p</i>	β	SE	<i>p</i>	β	SE	<i>p</i>
<i>Interventie-specifieke uitkomsten</i>									
Zelfcontrole ¹	0,12	0,03	< ,001	0,00	0,04	,951	0,02	0,04	,532
Zelfreflectie	0,01	0,04	,803	-0,02	0,06	,761	0,04	0,03	,238
Zelfvertrouwen	0,18	0,10	,060	0,09	0,11	,394	-0,05	0,06	,327
Emotionele zelfregulatie	0,17	0,05	< ,001	0,07	0,05	,163	-0,01	0,02	,676
<i>Intrapersoonlijk domein</i>									
Psychologische welzijn	0,15	0,06	,009	0,08	0,08	,287	0,04	0,04	,228
Veerkracht	0,09	0,73	,902	0,18	0,27	,507	-0,10	0,31	,736
Seksuele autonomie ¹	0,18	0,04	< ,001	0,08	0,06	,144	0,06	0,03	,074
Internaliserend gedrag ¹	-0,13	0,04	< ,001	-0,07	0,04	,048	0,01	0,03	,812
<i>Interpersoonlijk domein</i>									
Interpersoonlijke relaties in de klas	0,05	0,03	,056	-0,05	0,07	,499	-0,04	0,04	,278
Externaliserend gedrag	-0,20	0,12	,107	-0,13	0,11	,223	-0,02	0,07	,756
Agressie	-0,19	0,09	,030	0,01	0,10	,933	-0,01	0,05	,822
Pesten	-0,12	0,06	,056	-0,12	0,09	,149	-0,05	0,09	,624
Gepest worden	-0,01	0,04	,817	-0,04	0,06	,527	0,03	0,04	,530

Noot. ¹ Variantie van de voormeting van de desbetreffende uitkomst variabele was vastgezet op nul vanwege negatieve residuele variantie van de geobserveerde variabele op T1.

Tabel 4

Effectgroottes van verandering over tijd in R&W condities in vergelijking met de controle conditie

	Light				Standaard				Plus			
	T1-T2	T2-T3	T3-T4	T1-T4	T1-T2	T2-T3	T3-T4	T1-T4	T1-T2	T2-T3	T3-T4	T1-T4
<i>Interventie-specifieke uitkomsten</i>												
Zelfcontrole ¹	0,19	0,03	0,03	0,25	0,00	0,00	0,00	0,01	0,04	0,01	0,01	0,06
Zelfreflectie	0,01	0,00	0,00	0,02	-0,02	0,00	0,00	-0,03	0,06	0,01	0,00	0,07
Zelfvertrouwen	0,04	0,15	0,04	0,22	0,02	0,08	0,02	0,12	-0,01	-0,05	-0,01	-0,08
Emotionele zelfregulatie	0,17	0,09	0,04	0,29	0,08	0,04	0,02	0,13	-0,01	0,00	0,00	-0,01
<i>Intrapersoonlijk domein</i>												
Psychologisch welzijn	0,07	0,18	-0,01	0,24	0,04	0,10	-0,01	0,14	0,02	0,06	0,00	0,08
Veerkracht	0,04	-0,03	0,02	0,03	0,09	-0,07	0,04	0,06	-0,05	0,04	-0,02	-0,04
Seksuele autonomie ¹	0,32	0,06	0,01	0,38	0,15	0,03	0,00	0,18	0,12	0,02	0,00	0,15
Internaliserend gedrag ¹	0,19	0,04	0,03	0,26	0,11	0,03	0,02	0,15	-0,01	0,00	0,00	-0,01
<i>Interpersoonlijk domein</i>												
Interpersoonlijke relaties in de klas	0,05	0,03	0,02	0,10	-0,05	-0,03	-0,02	-0,10	-0,04	-0,02	-0,02	-0,08
Externaliserend gedrag	0,13	-0,05	0,03	0,11	0,09	-0,03	0,02	0,08	0,02	-0,01	0,00	0,02
Agressie	0,18	0,10	0,03	0,31	-0,01	-0,01	0,00	-0,02	0,01	0,01	0,00	0,02
Pesten	0,10	0,08	0,06	0,24	0,11	0,09	0,07	0,28	0,05	0,04	0,03	0,11
Gepest worden	0,01	0,01	0,00	0,01	0,04	0,02	0,00	0,06	-0,03	-0,02	0,00	-0,04

Noot. ¹ Variantie van de voormeting van de desbetreffende uitkomst variabele was vastgezet op nul vanwege negatieve residuele variantie van de geobserveerde variabele op T1; Effectgroottes (*d*) zijn dusdanig berekend dat positieve effectgroottes een verandering in de gewenste richting voor R&W aangeven.

Interventie-Specifieke Uitkomsten

Leerlingen in de Light conditie toonden, vergeleken met de Controle conditie, gunstigere veranderingstrajecten voor zelfcontrole ($d = 0,25$) en emotionele zelfregulatie ($d = 0,29$; zie Figuur 2). Zij hadden een sterkere stijging in zelfcontrole (slope = 1,75) dan leerlingen in de Controle conditie (slope = 1,44). Met betrekking tot emotionele zelfregulatie lieten leerlingen in de Light conditie een lichte verbetering over tijd zien (slope = 0,24), terwijl leerlingen in de Controle conditie een lichte afname lieten zien (slope = -0,10). Op beide uitkomsten verbeterden leerlingen met name in het eerste jaar van de interventie. Verder laten de resultaten een relevant ($d = 0,22$), maar niet significant ($p = ,060$), verschil zien tussen de veranderingstrajecten van zelfvertrouwen in de Light conditie en Controle conditie. Leerlingen in de Light conditie bleven relatief stabiel (slope = -1,68), terwijl leerlingen in de Controle conditie een kleine afname in zelfvertrouwen rapporteerden (slope = -2,58). Dit interventie-effect was het sterkst tussen T2 en T3. We hebben geen interventie-effecten gevonden voor zelfreflectie noch voor leerlingen in de Standaard en Plus condities. Over het algemeen leek de Light conditie de meest gunstige veranderingstrajecten te bewerkstelligen op de interventie specifieke uitkomsten, met name in het begin van de interventie.

[Figuur 2]

Overkoepelende Uitkomsten

Intrapersoonlijk domein. Ook voor psychologisch welzijn, seksuele autonomie, en internaliserend gedrag hadden leerlingen in de Light conditie positievere veranderingstrajecten dan leerlingen in de Controle conditie (range $d = 0,26-0,38$; zie Figuur 3). Zij bleven stabiel in hun psychologisch welzijn (slope = 0,08), terwijl leerlingen in de Controle conditie verslechterden (slope = -0,25). Het interventie-effect was het sterkste tussen T2 en T3. Op seksuele autonomie verbeterden leerlingen in de Light conditie sterker (slope = 0,73) dan leerlingen in de Controle conditie (slope = 0,32), vooral in het eerste jaar van de interventie. Wat betreft internaliserend gedrag rapporteerden leerlingen in de Light en Standaard condities een sterkere afname (slope_{Light} = -0,60; slope_{Standaard} = -0,53) dan leerlingen in de Controle conditie (slope = -0,41), met name in het eerste jaar. Er waren geen interventie-effecten op veerkracht noch voor leerlingen in de Plus conditie.

[Figuur 3]

Wanneer we de leerlingen in de Light conditie vergeleken met de leerlingen in de Standaard conditie op verandering in internaliserend gedrag, paste het vrij geschatte model significant beter op de data, $\Delta\chi^2_{SB}(1) = 6,60$; $p = ,010$, dan het model met gelijkgestelde slopes. Internaliserend gedrag nam dus sterker af bij leerlingen in de Light conditie (slope = -

0,60) dan bij leerlingen in de Standaard conditie (slope = -0,53). In het intrapersonlijk domein toonden leerlingen in de Light conditie over het algemeen dus de positiefste veranderingen over tijd, waarbij de interventie de grootste impact had aan het begin van het programma.

Interpersoonlijk domein. Leerlingen in de Light conditie hadden tevens een voordeliger veranderingstraject voor agressie vergeleken met de Controle conditie ($d = 0,31$; zie Figuur 3), vooral in het eerste jaar. Zij bleven over tijd relatief stabiel op agressie (slope = 0,09), terwijl leerlingen in de Controle conditie een toename lieten zien (slope = 0,56). Verder lieten de resultaten een relevant ($d = 0,24$), hoewel niet significant ($p = ,056$), verschil zien tussen veranderingstrajecten in pesten in de Light conditie ten opzichte van de Controle conditie. Leerlingen in de Light conditie vertoonden een minder sterke toename in pesten (slope = 2,34) dan leerlingen in de Controle conditie (slope = 2,80; zie Figuur 3). Het interventie-effect was het sterkste in het eerste jaar. Er zijn geen interventie-effecten gevonden op de andere uitkomsten noch voor leerlingen in de Standaard en Plus condities. Ook in het interpersoonlijk domein hadden leerlingen in de Light conditie over het algemeen betere veranderingstrajecten met, wederom, de sterkste interventie-effecten in het eerste jaar van de interventie.

Discussie

Het doel van de huidige studie was om 1) te onderzoeken of R&W een positief effect heeft op vmbo-leerlingen en 2) om te bepalen of verschillende versies van deze interventie verschillende effecten zouden bewerkstelligen. De interventie bleek gematigd effectief in het stimuleren van interventie-specifieke uitkomsten (zelfcontrole, emotionele zelfregulatie, en een trend voor zelfvertrouwen) en in het bevorderen van sommige aspecten van het intra- en interpersoonlijk domein, maar uitsluitend wanneer een kernteam betrokken was bij de implementatie van de interventie en vooral in het eerste jaar. R&W Light was met name effectief in het stimuleren van competenties en het voorkomen van problemen in het intrapersonlijk domein (psychologisch welzijn, seksuele autonomie, en internaliserend gedrag). In het interpersoonlijk domein leek R&W Light een buffer te zijn tegen toenames in agressie en pesten. Interventie-effecten waren over het algemeen gematigd (effecten van 0,24-0,38) en vergelijkbaar met effecten van andere universele schoolinterventies met soortgelijke doelen. Zo rapporteerden Durlak en collega's (2011) in hun meta-analyse effectgroottes van 0,23 voor positieve houdingen naar jezelf en anderen, 0,24 voor positief sociaal gedrag, 0,22 voor gedragsproblemen, en 0,24 voor emotionele problemen. Mertens en collega's (2020)

vonden in hun meta-analyse effectgroottes van 0,19 en 0,15 op respectievelijk het intrapersoonlijk en interpersoonlijk domein.

Met betrekking tot de interventie-specifieke uitkomsten vonden we op drie van de vier basiscompetenties, die van belang zijn volgens de R&W theorie, positieve interventie-effecten (zelfcontrole, emotionele zelfregulatie, en een trend voor zelfvertrouwen) wat de potentie van de interventie aangeeft. Echter, er is geen effect gevonden op de mate van zelfreflectie van leerlingen. Ook eerdere interventies zijn ineffectief gebleken in het verbeteren van zelfreflectie bij vmbo-leerlingen. Ter Vrugte en collega's (2015), bijvoorbeeld, onderzochten de effectiviteit van een spel om de rekenvaardigheid van vmbo-leerlingen te verbeteren. Dit spel bevatte een reflectiestimulerend component, maar ze vonden hier geen effect op. Mogelijk is de vroege adolescentie niet de beste ontwikkelingsfase om zelfreflectie te meten omdat de metacognitieve vaardigheden, die adolescenten in staat stellen om betekenisvol te reflecteren op gedachten, gevoelens en gedrag (Sauter et al., 2010), nog **volop** in ontwikkeling zijn. Desalniettemin, aangezien zelfreflectie een potentieel effectief component in schoolinterventies op het voortgezet onderwijs is (Mertens et al., 2020), zou toekomstig onderzoek moeten nagaan wanneer en hoe zelfreflectie het beste gestimuleerd kan worden bij vmbo-leerlingen.

Wat betreft de overkoepelende uitkomsten vonden we de meeste effecten in het intrapersoonlijk domein. Slechts één interventie-effect (agressie en een trend voor pesten) werd gevonden in het interpersoonlijk domein. Mogelijk komt dit doordat in het intrapersoonlijk domein een meer praktische aanpak gebruikt wordt, terwijl een meer verbale aanpak gebruikt wordt bij het interpersoonlijk domein. In de handleiding van de interventie (Ykema, 2002, 2018) wordt de nadruk in de oefeningen en spellen gelegd op gevoelens, emoties, en attitudes van leerlingen (dus het intrapersoonlijk domein; "Was je in balans tijdens het spel?"). De attitudes en gedragingen van leerlingen in relatie met anderen (dus het interpersoonlijk domein) worden voornamelijk aangekaart tijdens (verbale) rolspellen en discussies. Zo bespreken leerlingen bijvoorbeeld wat pesten is en wat leerlingen er tegen kunnen doen (bijv. "Hoe voelde de pester zich toen zij samen 'stop' zeiden?"). Deze meer verbale aanpak sluit minder goed aan bij een praktische leer methode waardoor vmbo-leerlingen mogelijk niet optimaal profiteren van de interventie op het interpersoonlijk domein.

De interventie-effecten waren het sterkste tijdens het eerste gedeelte van de interventie en namen af tot onbeduidende effecten in het tweede jaar; mogelijk is het dus voldoende om alleen het eerste jaar van de interventie te implementeren. Een afname in effecten in relatief lange interventies is eerder gevonden in meta-analyses (Bakermans-Kranenburg et al., 2003;

Cuijpers, 2002; De Mooij et al., 2020). Deze meta-analyses suggereren dat korte-termijn interventies met een beperkte hoeveelheid sessies de voorkeur hebben. Onderzoek heeft ook aangetoond dat participanten die profiteren van een interventie vaak vroeg in de interventie verbetering laten zien, ongeacht de tijdsspanne van de interventie (Lutz et al., 2014; Tadić et al., 2010). Deze afname in interventie-effecten zou mogelijk gerelateerd kunnen zijn aan een afname in de motivatie van leerlingen (Philips & Wennberg, 2014). Toekomstig onderzoek zou kunnen onderzoeken wat de effectiviteit van R&W is indien het tweede jaar van de interventie niet gegeven wordt, of wanneer alleen enkele boostersessies gegeven worden in dat jaar. Alleen het eerste jaar van de interventie implementeren zou wellicht de belasting voor leerlingen kunnen verlichten zonder de effectiviteit van de interventie in gevaar te brengen.

Verder lieten de resultaten zien dat de Light conditie (kernteam) het meest effectief was, wat suggereert dat voor sommige interventies geldt “*less is more*”. Dit is in lijn met de resultaten van de meta-analyse van Durlak en collega’s (2011) die aantoonde dat het positieve effect van het betrekken van meerdere systemen in interventies niet altijd gevonden wordt. In een interventie met een smalle ecologische focus (alleen kernteam) zijn slechts een paar personen betrokken; zij zijn dus als enige verantwoordelijk voor het correct en effectief toepassen van de interventie, wat de implementatie zou kunnen versterken (Durlak, 2016). In een interventie met een brede ecologische focus (bijv. toevoegen van een leerkracht- en/of oudercomponent) zijn veel mensen betrokken die de verantwoordelijkheid voor implementatie kunnen verdelen. Deze gedeelde verantwoordelijkheid zou mogelijk een “bijstandereffect” kunnen uitlokken, verminderde verantwoordelijkheidsgevoelens van individuen wanneer meerdere mensen aanwezig zijn (Fischer et al., 2011). Toekomstig onderzoek zou zich moeten richten op de mogelijk invloed van bijstandereffecten wanneer de verantwoordelijkheid voor implementatie verdeeld is onder een relatief grote groep mensen.

De resultaten van de huidige studie moeten geïnterpreteerd worden in het licht van zijn sterke en zwakke punten. Sterke punten van de studie zijn de meerdere meetmomenten, de drie condities die verschilden in de ecologische focus, en de grote steekproef. Een zwak punt is het gebrek aan informatie over de inhoud van de betrokkenheid van het schoolteam en de oudercomponent. Hoewel we weten dat deze schoolteam- en oudercomponenten geïmplementeerd zijn, weten we relatief weinig over *hoe* interventietechnieken toegepast zijn door schoolteam en ouders. Ten tweede, de R&W trainers gaven de interventie voor de eerste keer, na drie trainingdagen. Meer ervaren trainers zouden mogelijk meer verandering kunnen bewerkstelligen. Ten derde, de uitkomsten zijn enkel gemeten door middel van zelfrapportage

van de leerlingen. Daarnaast waren er aan het begin van de studie kleine verschillen tussen de condities op uitkomstmaten wat effect gehad zou kunnen hebben op de resultaten, ondanks dat we hier in de analyses voor gecontroleerd hebben, en hebben we SES niet onderzocht. Ten vierde, we hebben effecten meteen na de interventie onderzocht. Toekomstig onderzoek zou ook lange-termijn effecten van de interventie moeten onderzoeken. Tot slot, in deze studie is R&W onderzocht onder een specifieke doelgroep, namelijk vmbo-leerlingen, en bij een beperkt aantal scholen verspreid over Nederland. Hoewel het belangrijk is om inzicht te krijgen in welke interventies effectief zijn voor deze doelgroep, kunnen de bevindingen niet zonder verder onderzoek gegeneraliseerd worden naar andere groepen (leerlingen) en alle vmbo-leerlingen van Nederland. Daarnaast zou het interessant zijn voor vervolgonderzoek om meer inzicht te krijgen in verschillen in effectiviteit voor bepaalde subgroepen (bijv. persoonlijkheid: Mertens et al., 2022).

Conclusie

Met name vmbo-leerlingen lijken behoefte te hebben aan een effectieve interventie die hun competenties kan stimuleren en de ontwikkeling van problemen kan voorkomen in zowel het intra- als het interpersoonlijk domein (Stevens & De Looze, 2018). R&W kan een positieve invloed hebben op vmbo-leerlingen, met name in het verbeteren van het intrapersoonlijk domein en met name in het eerste jaar van de interventie. In het interpersoonlijk domein fungeert de interventie mogelijk als een buffer tegen toenames in agressie en pesten. Verder is gebleken dat effecten niet noodzakelijkerwijs toenemen bij een langere interventie of door meer mensen erbij te betrekken (dus door een bredere ecologische implementatiestrategie).

Onze bevindingen hebben belangrijke implicaties voor onderzoek en de praktijk. Zo laat onze studie zien dat verschillende implementatiestrategieën van dezelfde interventie effect kan hebben op de effectiviteit ervan. Dit onderstreept het belang om bij de evaluatie van een interventie niet alleen het programma zelf maar ook de implementatiestrategie te onderzoeken. Voor de praktijk laat onze studie zien dat interventies met een actieve en praktijkgerichte leerstijl positieve effecten kunnen hebben op vmbo-leerlingen, een kortere interventie periode mogelijk voldoende is om verandering te bewerkstelligen en het betrekken van meer leerkrachten en ouders niet altijd een positief effect heeft. Voor scholen zou het daarom waardevol kunnen zijn om te investeren in interventies met een leerstijl die aansluit bij die van de doelgroep, een goede kwaliteit training voor een kernteam van het schoolteam en een hoge kwaliteit van implementatie met een smalle ecologische focus.

Trial Registratie

Nederlands Trial Register nummer NL6371 (oud nummer: NTR 6554), geregistreerd op 3 juli 2017.

Verklaring van Concurrerende Belangen

Geen.

Dankbetuigingen

Deze studie is gefinancierd door een beurs van ZonMw, nummer 531001106

Auteurschap (CRediT)

EM: Conceptualisatie, methodologie, formele analyses, dataverzameling, schrijven – conceptversie, schrijven – revisies, projectmanagement; **MD:** Conceptualisatie, methodologie, schrijven – revisies, supervisie; **MvL:** Conceptualisatie, methodologie, schrijven – revisies, supervisie; **ER:** Conceptualisatie, methodologie, schrijven – revisies, supervisie, subsidie verworven.

Korte Verklaring over de Data en Analyses

De analyse codes en output zijn vrij toegankelijk via Open Science Framework ([GEMASKEERD VOOR REVIEW]). De data zijn wegens gevoeligheid niet vrij toegankelijk, maar op aanvraag beschikbaar bij de eerste auteur.

Literatuur

- Anderson, L. M., Reilly, E. E., Gorrell, S., Schaumberg, K., & Anderson, D. A. (2016). Gender-based differential item function for the difficulties in emotion regulation scale. *Personality and Individual Differences, 92*, 87–91.
- Bakermans-Kranenburg, M. J., IJzendoorn, M. H. van, & Juffer, F. (2003). Less is more: Meta-analyses of sensitivity and attachment interventions in early childhood. *Psychological Bulletin, 129*, 195–215.
- Bishop, D. C., Pankratz, M. M., Hansen, W. B., Albritton, J., Albritton, L., & Strack, J. (2014). Measuring fidelity and adaptation: Reliability of an instrument for school-based prevention programs. *Issues and Solutions in Risky Behavior Prevention Research, 37*, 231-257.
- Boor-Klip, H. J., Segers, E., Hendrickx, M. M. H. G., & Cillessen, A. H. N. (2016). Development and psychometric properties of the classroom peer context questionnaire. *Social Development, 25*, 370–389.
- Bronfenbrenner, U. (1979). *The ecology of human development*. Harvard University Press.
- Bruijn, D. de, Overmaat, M., Glaudé, M., Heemskerk, I., Leeman, Y., Roeleveld, J., & Venne, L. van de (2005). Krachtige leeromgevingen in het middelbaar beroepsonderwijs: Vormgeving en effecten. *Pedagogische Studiën, 82*, 77–95.
- Chorpita, B. F., Reise, S., Weisz, J. R., Grubbs, K., Becker, K. D., Krull, J.L., & The Research Network on Youth Mental Health (2010). Evaluation of the brief problem checklist: Child and caregiver interviews to measure clinical progress. *Journal of Consulting and Clinical Psychology, 78*, 526–536.
- Cuijpers, P. (2002). Effective ingredients of school-based drug prevention programs. A systematic review. *Addictive Behaviors, 27*, 1009–1023.
- Davidson, J. R. T., & Connor, K. M. (2017). *Connor-Davidson Resilience Scale (CD-RISC) manual*. Unpublished and partly accessible at www.cd-risc.com.
- Dodge, K. A., & Coie, J. D. (1987). Social-information-processing factors in reactive and proactive aggression in children's peer groups. *Journal of Personality and Social Psychology, 53*, 1146–1158.
- Durlak, J. A. (2016). Programme implementation in social and emotional learning: Basic issues and research findings. *Cambridge Journal of Education, 4*(3), 333–345.
<https://doi.org/10.1080/030574X.2016.1142504>

- Durlak, J. A., Weissberg, R. P., Dymnicki, A. B., Taylor, R. D., & Schellinger, K. B. (2011). The impact of enhancing students' social and emotional learning: A meta-analysis of school-based universal interventions. *Child Development, 82*, 405–432.
- Feingold, A. (2013). A regression framework for effect size assessments in longitudinal modeling of group differences. *Review of General Psychology, 17*, 111–121.
- Finkenauer, C., Engels, R. C. M. E., Baumeister, R. F. (2005). Parenting behaviour and adolescent behavioural and emotional problems: The role of self-control. *International Journal of Behavioral Development, 29*, 58–69.
- Fischer, P., Krueger, J. I., Greitemeyer, T., Vogrincic, C., Kastenmüller, A., Frey, D., Heene, M., Wicher, M., & Kainbacher, M. (2011). The bystander-effect: A meta-analytic review on bystander intervention in dangerous and non-dangerous emergencies. *Psychological Bulletin, 137*, 517–537.
- Graaf, H. de, Meijer, S., Poelman, J., & Vanwesenbeeck, I. (2005). *Seks onder je 25e*. Uitgeverij Eburon.
- Graaf, I. de, Haas, S. de, Zaagsma, M., & Wijssen, C. (2016). Effects of rock and water: An intervention to prevent sexual aggression. *Journal of Sexual Aggression, 22*, 4–19.
- Harter, S. (1988). *Manual for the Self-Perception Profile for Adolescents*. University of Denver.
- Kuosmanen, T., Clarke, A., & Barry, M. M. (2019). Promoting adolescents' mental health and wellbeing: Evidence synthesis. *Journal of Public Mental Health, 18*(1), 73–83. <https://doi.org/10.1108/JPMH-07-2018-0036>
- Lutz, W., Hofmann, S. G., Rubel, J., Boswell, J. F., Shear, M. K., Gorman, J. M., Woods, S. W., & Barlow, D. H. (2014). Patterns of early change and their relationship to outcome and early treatment termination in patients with panic disorder. *Journal of Consulting and Clinical Psychology, 82*, 287–297.
- Mahoney, J. L., Weissberg, R. P., Greenberg, M. T., Dusenburg, L., Jagers, R. J., Niemi, K., Schlinger, M., Schlund, J., Shriver, T. P., VanAusdal, K., & Yoder, N. (2021). Systemic social and emotional learning: Promoting educational success for all preschool to high school students. *American Psychologist, 76*(7), 1128–1142. <https://doi.org/10.1037/amp0000701>
- Mertens, E. C. A., Deković, M., Leijten, P., Londen, M. van, & Reitz, E. (2020). Components of school-based interventions stimulating students' intrapersonal and interpersonal domains: A meta-analysis. *Clinical Child and Family Psychology Review, 23*, 605–631. <https://doi.org/10.1007/s10567-020-00328-y>

Mertens, E. C. A., Deković, M., Londen, M. van, & Reitz, E. (2022). Personality as a moderator of intervention effects: Examining differential susceptibility. *Personality and Individual Differences, 186 Part A*. Article number 111323.

<https://doi.org/10.1016.j.paid.2021.111323>

- Modecki, K. L., Zimmer-Gembeck, M. J., & Guerrà, N. (2017). Emotion regulation, coping, and decision making: Three linked skills for preventing externalizing problems in adolescence. *Child Development, 88*, 417–426.
- Mooij, B. de, Fekkes, M., Scholte, R. H. J., & Overbeek, G. (2020). Effective components of social skills training programs for children and adolescents in nonclinical samples: A multilevel meta-analysis. *Clinical Child and Family Psychology Review, 23*, 250–264.
- Muthén, L. K., & Muthén, B. O. (1998–2010). *Mplus user's guide* (6th ed.). Muthén & Muthén.
- Park, D., Tsukayama, E., Goodwin, G., Patrick, S., & Duckworth, A. L. (2017). A tripartite taxonomy of character: Evidence for intrapersonal, interpersonal, and intellectual competencies in children. *Contemporary Educational Psychology, 48*, 16–27.
- Pellegrino, J. W., & Hilton, M. L. (2012). *Education for life and work: Developing transferable knowledge and skills in the 21st century*. National Academy of Sciences.
- Philips, B., & Wennberg, P. (2014). The importance of therapy motivation for patients with substance use disorders. *Psychotherapy, 51*, 555–562.
- Ravens-Sieberer, U., & The European KIDSCREEN Group (2006). *The KIDSCREEN questionnaires. Quality of life questionnaires for children and adolescents – Handbook*. Pabst Science Publisher.
- Satorra, A., & Bentler, P. M. (2010). Ensuring positiveness of the scaled difference chi-square test statistic. *Psychometrika, 75*, 243–248.
- Sauter, F. M., Heyne, D., Blöte, A. W., Widenfelt, B. M. van, & Westenberg, P. M. (2010). Assessing therapy-relevant cognitive capacities in young people: Development and psychometric evaluation of the self-reflection and insight scale for youth. *Behavioral and Cognitive Psychotherapy, 38*, 303–317.
- Smit, K., Brabander, C. J. de, & Martens, R. L. (2014). Student-centred and teacher-centred learning environment in pre-vocational secondary education: Psychological needs, and motivation. *Scandinavian Journal of Educational Research, 58*(6), 695–712.
- <https://doi.org/10.1080/00313831.2013.821090>
- Solberg, M. E., & Olweus, D. (2003). Prevalence estimation of school bullying with the Olweus bully/victim questionnaire. *Aggressive Behavior, 29*, 239–268.

- Stevens, G., & Looze, M. de (2018). Welbevinden en psychische problemen. In M. Boer & G. Stevens (Eds.), *HBSC 2017: Gezondheid, welzijn en opvoeding van jongeren in Nederland* (pp. 63–75). Ridderprint BV.
- Tadić, A., Helmreich, I., Mergl, R., Hautzinger, M., Kohnen, R., Henkel, V., & Hegerl, U. (2010). Early improvement is a predictor of treatment outcome in patients with mild major, minor or subsyndromal depression. *Journal of Affective Disorders*, *120*, 86–93.
- Timmermans, A., Naaijer, H., Keuning, J., & Zijssling, D. (2017). *Cohortonderzoek COOL5-18 basisrapport meting VO-3 in 2014*. GION (onderwijs/onderzoek).
- Vrugte, J. ter, Jong, T. de, Wouters, P., Vandercruysse, S., Elen, J., & Oostendorp, H. van (2015). When a game supports prevocational math education but integrated reflection does not. *Journal of Computer Assisted Learning*, *31*, 462–480.
- Ykema, F. (2002). *Het Rots en Water perspectief. Een psychofysieke training voor jongens. Basisboek*. Uitgeverij SWP.
- Ykema, F. (2018). *Rots en Water praktijkboek. Psychofysieke training voor jongens én meisjes*. Uitgeverij SWP.
- Ykema, F., Hartman, D., & Imms, W. (2006). *Bringing it together. Includes 22 case studies of Rock & Water in practice in various settings*. Callaghan: University of Newcastle, Family action Centre and Gadaku institute.

Derde reactie reviewer 1 – 14-6-2023

Zoals aangegeven in mijn eerdere review kent dit onderzoek veel sterke kanten en zou het goed passen in JiO. Los van mijn waardering voor de sterke punten, zie ik ook nog punten voor verbetering.

1. A) Dat R&W veel geïmplementeerd wordt op scholen zonder dat er veel bekend is over de effectiviteit ervan, is een belangrijk argument om wetenschappelijk onderzoek hiernaar te doen. Daarin volg ik de auteurs volledig. Dat neemt niet weg dat het ook belangrijk is om te weten wat de meerwaarde is van R&W ten opzichte van bestaande interventies. Dat heeft naar mijn mening niets te maken met "superioriteit". In wetenschappelijk onderzoek is het heel gebruikelijk dat men de stand van zaken beschrijft en men vervolgens beargumenteert hoe de eigen benadering/methode/interventie daarop aansluit, zodat de lezer zicht krijgt op de meerwaarde van de nieuwe benadering. B) Het kan zijn dat ik niet duidelijk was in mijn vorige review met wat ik bedoelde met een onderbouwde theorie. De auteurs benoemen inderdaad, zij het slechts beknopt, de theorie van R&W. Dit is een specifieke interventietheorie, inderdaad belangrijk om te beschrijven. Maar deze theorie is, blijkend uit de referenties, niet peer-reviewed en niet gepubliceerd in de academische (Engelstalige) literatuur. Een interventie kan pas als beloftevol beschouwd worden wanneer de specifieke interventietheorie verantwoord kan worden vanuit - ik citeer - "algemeen aanvaarde en met onderzoek ondersteunde theorieën" (zie effectladder, <https://www.nji.nl/effectieve-jeugdhulp/werken-aan-effectiviteit>). Met "algemeen aanvaard" wordt bedoeld theorieën die in de internationale onderzoeksgemeenschap aanvaard zijn en ondersteund worden door (voldoende) empirisch onderzoek. C) Wat mij (nu pas) opvalt is dat de interventie-specifieke theorie van Ykema zich vooral concentreert op jongens in de adolescentie met als doelstelling het bevorderen van mannelijkheid (<https://files.eric.ed.gov/fulltext/ED451886.pdf>). Kun jullie uitleggen hoe de theorie van Ykema zich ontwikkeld heeft tot een theorie die voor zowel meisjes als jongens geldt? Wellicht kan in de discussie hier ook nog kort aandacht aan besteed worden (bv. vervolgonderzoek naar verschillen in effectiviteit voor jongens en meisjes).
1. Bedankt voor het beschrijven van hoe R&W aansluit bij de leerstijl van jongeren. Dat is nu duidelijker.
1. Ik heb begrip voor het argument dat de woordlimiet dit eigenlijk niet toelaat zonder andere relevante informatie te schrappen.

1. De interventie is uitgevoerd op het niveau van de scholen. Het zou mij sterk lijken dat het aantal clusters in verhouding tot de condities niet zou uitmaken voor de poweranalyse. Hoe hebben jullie het aantal clusters en condities meegenomen in de berekening van de power? Ik ben geen statisticus, maar ik meen dat het benodigde aantal clusters samenhangt met de mate van multilevel variantie (intraclasscorrelatie, ICC) en het aantal personen per cluster? Kunnen jullie de lezer informatie geven over de power analyse? b) Een klein aantal clusters per conditie maakt ook de kans groot dat de verdeling van clustereigenschappen ongelijk is tussen condities, ondanks de randomisatie. Veel onderzoekers worstelen hiermee (vaak vanwege beperkte middelen; interventieonderzoek is duur). Dit leidt tot methodologische beperkingen die lang niet altijd te voorkomen zijn, maar ze moeten wel duidelijk benoemd worden in de limitatie sectie om de lezer hier bewust van te maken (meer dan enkel een bijzin).
1. Bedankt voor het toevoegen van kenmerken van de steekproef van leerkrachten. Ik vind het opvallend dat er geen informatie over schoolkenmerken gegeven kan worden. Schoolkenmerken als locatie (stedelijk, rand-stedelijk, landelijk), schoolgrootte en bij voorkeur ook leerlingenpopulatie kunnen in de meeste gevallen vermeld worden zonder dat de anonimiteit in gevaar komt.
1. Waar kan ik de informatie over de selectie en samenstelling van het kernteam terugvinden in het artikel?
1. Bedankt voor het toevoegen van de analyses naar conditiever verschillen op uitkomstmaten op T1. Ik twijfel of de participanten sectie de meeste logische plaats is voor de rapportage van deze analyses. Mijns inziens is het wat logischer om dit te vermelden aan het begin van de resultaten, maar anderzijds vind ik het geen probleem als het bij participanten zou blijven staan als de auteurs daar de voorkeur aan geven.
1. Bedankt voor het toevoegen van deze suggestie voor verder onderzoek.
1. Bedankt voor deze toevoeging.
1. Bedankt om hier aandacht aan te besteden in de discussie.
1. Bedankt voor deze toevoeging.

1. Bedankt voor deze herformulering. Zo klinkt het veel beter.

Derde reactiebrief auteur – 23-6-2023

Utrecht, 23 juni 2023

Onderwerp: Derde revisie manuscript

Geachte Prof. dr. Boendermaker en meneer Veenstra,

Bijgevoegd vindt u de derde revisie van ons manuscript: “Stevig als een rots, flexibel als water? Effectiviteit van een schoolinterventie gericht op het intrapersoonlijk en interpersoonlijk domein”.

We bedanken de reviewer voor de tijd die is genomen om nogmaals naar ons manuscript te kijken en de laatste punten van feedback te geven. Om het reviewproces te faciliteren hebben wij onze reacties op de reviewer toegevoegd in de Appendix hieronder en aanpassingen in het manuscript geel gemarkeerd.

We gaan er vanuit dat deze derde revisie de laatste is en dat de redactie nu de beslissing zal nemen.

We zien er naar uit van u te horen,

Met vriendelijke groet,

Namens mijn co-auteurs Deković, Van Londen, en Reitz,

Esther Mertens

e.c.a.mertens@fsw.leidenuniv.nl

Appendix

Reviewer A:

Review Sterk als rots, flexibel als water – Revisie 2

Zoals aangegeven in mijn eerdere review kent dit onderzoek veel sterke kanten en zou het goed passen in JiO. Los van mijn waardering voor de sterke punten, zie ik ook nog punten voor verbetering.

Wij danken de reviewer voor de tijd om naar ons manuscript te kijken en zijn blij dat onze eerdere revisies verhelderend waren.

1. A) Dat R&W veel geïmplementeerd wordt op scholen zonder dat er veel bekend is over de effectiviteit ervan, is een belangrijk argument om wetenschappelijk onderzoek hiernaar te doen. Daarin volg ik de auteurs volledig. Dat neemt niet weg dat het ook belangrijk is om te weten wat de meerwaarde is van R&W ten opzichte van bestaande interventies. Dat heeft naar mijn mening niets te maken met “superioriteit”. In wetenschappelijk onderzoek is het heel gebruikelijk dat men de stand van zaken beschrijft en men vervolgens beargumenteert hoe de eigen benadering/methode/interventie daarop aansluit, zodat de lezer zicht krijgt op de meerwaarde van de nieuwe benadering. B) Het kan zijn dat ik niet duidelijk was in mijn vorige review met wat ik bedoelde met een onderbouwde theorie. De auteurs benoemen inderdaad, zij het slechts beknopt, de theorie van R&W. Dit is een specifieke interventietheorie, inderdaad belangrijk om te beschrijven. Maar deze theorie is, blijkend uit de referenties, niet peer-reviewed en niet gepubliceerd in de academische (Engelstalige) literatuur. Een interventie kan pas als beloftevol beschouwd wanneer de specifieke interventietheorie verantwoord kan worden vanuit – ik citeer - “algemeen aanvaarde en met onderzoek ondersteunde theorieën “ (zie effectladder, <https://www.nji.nl/effectieve-jeugdhulp/werken-aan-effectiviteit>). Met “algemeen aanvaard” wordt bedoeld theorieën die in de internationale onderzoeksgemeenschap aanvaard zijn en ondersteund worden door (voldoende) empirisch onderzoek. C) Wat mij (nu pas) opvalt is dat de interventie-specifieke theorie van Ykema zich vooral concentreert op jongens in de adolescentie met als doelstelling het bevorderen van mannelijkheid (<https://files.eric.ed.gov/fulltext/ED451886.pdf>). Kun jullie uitleggen hoe de theorie van Ykema zich ontwikkeld heeft tot een theorie die voor zowel meisjes als jongens geldt? Wellicht kan in de discussie hier ook nog kort aandacht aan besteed worden (bv. vervolgonderzoek naar verschillen in effectiviteit voor jongens en meisjes). In de inleiding hebben we bij de beschrijving van R&W toegevoegd wat kenmerkend is voor deze schoolinterventie en hoe deze zich daarmee onderscheidt van andere schoolinterventies.

“R&W onderscheidt zich met name van andere schoolinterventies door een actieve lesmethode waarin de nadruk ligt op activerende werkvormen in plaats van passieve verbale uitleg;” P. 6

Als onafhankelijk onderzoekers hebben wij de R&W interventie, zoals in de praktijk wordt toegepast, onderzocht op effectiviteit, wat handvatten kan geven voor de praktijk. In de discussie geven we nu aan dat de theorie van R&W slechts gedeeltelijk onderzocht is en dat meer onderzoek nodig is:

“Tenslotte, het dient opgemerkt te worden dat de theorie achter R&W slechts gedeeltelijk onderzocht is. Bijvoorbeeld, Mertens et al. (2022b) hebben laten zien dat toenames in zelfregulatie en zelfvertrouwen inderdaad gerelateerd zijn aan toenames in veerkracht en psychologisch welzijn, zoals de R&W theorie veronderstelt. Er is echter meer onderzoek nodig om de overige aspecten van deze theorie empirisch te onderbouwen.” P. 24

In eerste instantie is de interventie inderdaad ontwikkeld om ingezet te worden bij jongens. Inmiddels heeft Ykema de interventie doorontwikkeld en wordt deze al jarenlang ingezet bij zowel jongens als meisjes. Omdat in de eerdere studie naar de effectiviteit van R&W in Nederland (De Graaf et al., 2016) alleen jongens waren geïncludeerd, vonden wij het belangrijk om zowel jongens als meisjes in ons onderzoek te includeren. We hebben er bewust voor gekozen dit niet in dit manuscript te rapporteren, om de informatie die we in het manuscript geven behapbaar te houden. Het is belangrijk dat een manuscript een duidelijke focus heeft om een duidelijke en heldere boodschap te kunnen geven. We willen dat de nadruk ligt op de effectiviteit en op de verschillende implementatie strategieën. Daar bovenop nog moderatie effecten van het geslacht rapporteren valt buiten de focus van dit manuscript. In de discussie hebben we toegevoegd dat geslachtsverschillen aandacht verdienen in het toekomstig onderzoek:

“Daarnaast zou het interessant zijn voor vervolgonderzoek om meer inzicht te krijgen in verschillen in effectiviteit voor bepaalde subgroepen (bijv. jongens versus meisjes, jongeren met verschillen in persoonlijkheid: Mertens et al., 2022a).” P. 24

2. Bedankt voor het beschrijven van hoe R&W aansluit bij de leerstijl van jongeren. Dat is nu duidelijker.
3. Ik heb begrip voor het argument dat de woordlimiet dit eigenlijk niet toelaat zonder andere relevante informatie te schrappen.
4. De interventie is uitgevoerd op het niveau van de scholen. Het zou mij sterk lijken dat het aantal clusters in verhouding tot de condities niet zou uitmaken voor de poweranalyse. Hoe hebben jullie het aantal clusters en condities meegenomen in de berekening van de power? Ik ben geen statisticus, maar ik meen dat het benodigde aantal clusters samenhangt met de mate van multilevel variantie (intraclasscorrelatie, ICC) en het aantal personen per cluster? Kunnen jullie de lezer informatie geven over de power analyse? b) Een klein aantal clusters per conditie maakt ook de kans groot dat de verdeling van clustereigenschappen ongelijk is tussen condities, ondanks de randomisatie. Veel onderzoekers worstelen hiermee (vaak vanwege beperkte middelen; interventieonderzoek is duur). Dit leidt tot methodologische beperkingen die lang niet altijd te voorkomen zijn, maar ze moeten wel duidelijk benoemd worden in de limitatie sectie om de lezer hier bewust van te maken (meer dan enkel een bijzin).

De analyses zijn uitgevoerd op het niveau van de individuele leerlingen aangezien we geïnteresseerd waren in veranderingen van de leerlingen en onze variabelen allemaal op dat niveau gemeten zijn. Aggregeren van metingen naar een hoger niveau, is voor ons onderzoek conceptueel niet interessant. Bovendien is het daarbij de vraag of een vragenlijst die ontwikkeld is voor metingen op het individuele niveau op een valide manier geïnterpreteerd kan worden op een hoger niveau zoals het schoolniveau. De meetinstrumenten die wij hebben gebruikt, hebben betrekking op het welzijn van individuele leerlingen, dus bijvoorbeeld hoe een bepaalde leerling in zijn vel zit. Wanneer

je dit gaat analyseren op het niveau van de school, kijk je niet meer naar het gevoel van individuele leerlingen, maar naar het gemiddelde gevoel van een bepaalde school. Het analyseren van de data op leerling niveau sluit aan bij de gebruikte analyse techniek aangezien groeimodellen met name geschikt zijn om verschillen in ontwikkelingstrajecten te onderzoeken, dus op individueel niveau. Wij hebben er bewust voor gekozen geen multilevel model te modelleren aangezien de intraclasscorrelaties op beide niveaus (klas en school) erg laag was (Klasniveau: $M = 4,68\%$, range = 1-11%; Schoolniveau: $M = 2,29\%$, range = 0-16%) en wij geen specifieke voorspellers hadden op de verschillende niveaus waarin wij geïnteresseerd waren. Na overleg met de hoogleraar statistiek aan onze universiteit, hebben wij besloten de geclusterde structuur van onze data te modelleren door de *complex sample cluster* correctie toe te passen in ons model. Deze correctie maakt gebruik van een *sandwich estimator* waarbij er in de berekening van de standaard errors en de chi-kwadraat toets van model fit een correctie uitgevoerd wordt waarbij er rekening gehouden wordt met stratificatie, afhankelijkheid van observaties, en een ongelijke selectiekans. Deze correctie wordt aangeraden voor het analyseren van complexe vragenlijst data en groeimodellen (bijv. Muthén & Muthén, 1998-2010, p. 115).

Wanneer iemand specifiek geïnteresseerd is in de poweranalyse van onze studie, kan dit opgezocht worden in het protocol paper wat wij gepubliceerd hebben en in de online registratie van ons onderzoek. In ons manuscript verwijzen wij naar beide.

5. Bedankt voor het toevoegen van kenmerken van de steekproef van leerkrachten. Ik vind het opvallend dat er geen informatie over schoolkenmerken gegeven kan worden. Schoolkenmerken als locatie (stedelijk, rand-stedelijk, landelijk), schoolgrootte en bij voorkeur ook leerlingenpopulatie kunnen in de meeste gevallen vermeld worden zonder dat de anonimiteit in gevaar komt.

6. Waar kan ik de informatie over de selectie en samenstelling van het kernteam terugvinden in het artikel?

De informatie over de leerkrachten die de R&W lessen gaven staat op pagina 9. Om aan te geven hoe leerkrachten werden geselecteerd, hebben wij aan deze informatie toegevoegd dat dit door de scholen zelf bepaald werd.

“Scholen bepaalden zelfstandig welke leerkrachten getraind werden in het geven van R&W lessen. De R&W lessen werden in totaal verzorgd door 49 leerkrachten (trainers) tussen de 23 en 64 jaar oud ($M = 36,77$, $SD = 10,02$) met gemiddeld 12,98 jaar aan leerkracht-ervaring. De meeste trainers waren gymdocent ($n = 30$; 71%). Iets meer dan de helft van de trainers was man ($n = 27$; 56%) en bijna alle trainers hadden een Westerse achtergrond ($n = 46$; 96%).” P. 9-10

7. Bedankt voor het toevoegen van de analyses naar conditieverschillen op uitkomstmaten op T1. Ik twijfel of de participanten sectie de meeste logische plaats is voor de rapportage van deze analyses. Mijns inziens is het wat logischer om dit te vermelden aan het begin van de resultaten, maar anderzijds vind ik het geen probleem als het bij participanten zou blijven staan als de auteurs daar de voorkeur aan geven.
8. Bedankt voor het toevoegen van deze suggestie voor verder onderzoek.
9. Bedankt voor deze toevoeging.

10. Bedankt om hier aandacht aan te besteden in de discussie.

11. Bedankt voor deze toevoeging.

12. Bedankt voor deze herformulering. Zo klinkt het veel beter.

Vierde versie van het manuscript – 23-6-2023

Sterk als een rots, flexibel als water? Effectiviteit van een schoolinterventie gericht op het intrapersoonlijk en interpersoonlijk domein

Samenvatting

De huidige studie had als doel om te onderzoeken of 1) de schoolinterventie Rots en Water (R&W) effectief was in het verbeteren van competenties en voorkomen van problemen in het intra- en interpersoonlijk domein van leerlingen op het voorbereidend middelbaar beroepsonderwijs (vmbo), en of 2) interventie-effecten veranderden naarmate er meer mensen betrokken waren bij de interventie (kernteam van leerkrachten, het hele schoolteam, het hele schoolteam en ouders). We hebben een gerandomiseerde gecontroleerde studie uitgevoerd onder brugklassers ($N = 1299$; $M_{leeftijd} = 12,38$; 54% jongens) die vier keer digitale vragenlijsten hebben ingevuld. De resultaten lieten zien dat R&W een aantal interventie-specifieke uitkomsten (zelfcontrole en emotionele zelfregulatie), uitkomsten in het intrapersoonlijk domein (psychologisch welzijn, seksuele autonomie, en internaliserend gedrag) en in het interpersoonlijk domein (agressie) verbeterde, maar uitsluitend wanneer een kernteam van leerkrachten betrokken was bij de implementatie van de interventie. De (gematigde) interventie-effecten waren het grootste in het eerste jaar van de tweejarige interventie. Deze resultaten laten zien dat R&W positieve effecten kan bewerkstelligen bij vmbo-leerlingen wanneer deze geïmplementeerd wordt door een kernteam. Het lijkt dus waardevol voor scholen om te investeren in interventies met een leerstijl die aansluit bij die van de doelgroep, met een hoge kwaliteit van implementatie en een beperkt aantal betrokken personen.

Abstract

The aim of the present study was to examine 1) the effectiveness of the school-based intervention Rock and Water (R&W) in stimulating competencies and preventing problems in the intra- and interpersonal domains of prevocational students, and 2) whether intervention effects changed as more people were involved in the intervention (core team of teachers, whole school, and whole school and parents). We have conducted a Randomized Controlled Trial among 7th Grade students ($N = 1299$, $M_{age} = 12.38$, 54% boys) who completed four digital questionnaires. The results showed that R&W improved some intervention specific outcomes (self-control and emotional self-regulation), outcomes in the intrapersonal domain (psychological wellbeing, sexual autonomy, and internalizing behavior) and in the interpersonal domain (aggression), but only when solely a core team of teachers was involved in the implementation of the intervention. The (moderate) intervention effects were strongest in the first year of the two-year intervention. These results show that R&W can establish positive effects among prevocational students when implemented by a core team. Thus, it appears worthwhile for schools to invest in interventions that fit with the learning style of the

target group, with a high level of implementation quality and involving a limited number of people.

Trefwoorden: Schoolinterventie; Rots en Water; Intrapersoonlijk domein; Interpersoonlijk domein; Vmbo-leerlingen

Bijsluiter voor de praktijk

1. Vmbo-leerlingen rapporteren meer problemen in het intra- en interpersoonlijk domein dan havo- en vwo-leerlingen.
2. Wij onderzochten: 1) de effectiviteit van de schoolinterventie Rots en Water voor het stimuleren van het intra- en interpersoonlijk domein bij vmbo-leerlingen en 2) of de effectiviteit veranderde naarmate verschillende mensen betrokken waren bij de interventie (kernteam leerkrachten, hele schoolteam, hele schoolteam en ouders).
3. De interventie liet gematigde positieve effecten zien, met name op uitkomsten in het intrapersoonlijk domein, wanneer alleen een kernteam van leerkrachten betrokken was bij de uitvoering van de interventie.
4. Interventie-effecten waren het grootste in het eerste jaar van de interventie.
5. Het lijkt waardevol om te investeren in goede training voor en kwalitatief hoogstaande implementatie door een kernteam.

Aantal woorden:

5917 (hoofdtekst)

Introductie

Scholen spelen, naast andere factoren zoals gezin, leeftijdsgenoten en maatschappelijke omstandigheden, een belangrijke rol in het bevorderen van de ontwikkeling van leerlingen in het intrapersoonlijk en interpersoonlijk domein (Pellegrino & Hilton, 2012). Het *intrapersoonlijk domein* refereert naar subjectief persoonlijk functioneren (bijv. psychologisch welzijn) en is gerelateerd aan schoolresultaten en zelfregulatie. Het *interpersoonlijk domein*, daarentegen, refereert naar sociaal functioneren (bijv. agressie) en is onder andere gerelateerd aan relaties met leeftijdsgenoten en conflict oplossend vermogen (Park et al., 2017). Leerlingen kunnen in deze domeinen competenties ontwikkelen door relevante cognitieve, affectieve en sociale vaardigheden te leren (bijv. identificeren van emoties en perspectief wisselen). Wanneer het verwerven van (sommige) vaardigheden achterwege blijft, kunnen leerlingen problemen ontwikkelen (Modecki et al., 2017). Hoewel de twee domeinen elkaar beïnvloeden, worden ze als apart beschouwd. Scholen zouden bewust kunnen proberen de competenties van leerlingen op beide domeinen te stimuleren om mogelijke problemen te voorkomen (Pellegrino & Hilton, 2012).

Met name leerlingen op het voorbereidend middelbaar beroepsonderwijs (vmbo) zouden baat kunnen hebben bij stimulering van competenties en het voorkomen van problemen. Zij rapporteren, vergeleken met leerlingen op het hoger algemeen voortgezet onderwijs (havo) en voorbereidend wetenschappelijk onderwijs (vwo), over het algemeen een lager niveau van welzijn, meer gedragsproblemen, en meer problemen met leeftijdsgenoten (Stevens & De Looze, 2018). Tegelijkertijd kunnen vmbo-leerlingen een uitdagende groep zijn om te betrekken bij interventielessen op school. Over het algemeen laten deze leerlingen minder intrinsieke motivatie voor school en (verbaal) leren zien dan havo- en vwo-leerlingen (Timmermans et al., 2017). Hoewel schoolinterventies positieve effecten laten zien (e.g., Mertens et al., 2020), is het onduidelijk in hoeverre deze ook effectief zijn voor specifiek vmbo-leerlingen; vooral voor het betrekken van deze groep leerlingen bij een schoolinterventie lijkt het namelijk belangrijk een interventie te implementeren die goed aansluit bij hun interesses en leerstijl. Vmbo-leerlingen worden het beste bereikt en blijven het meest gemotiveerd en betrokken met een praktijkgerichte leer methode. Deze leer methode houdt in dat er een leeromgeving gecreëerd wordt met een praktijkgerichte insteek met activerende, stapsgewijze, gestructureerde taken; leerlingen kunnen hun nieuw geleerde vaardigheden toepassen en eigen doelen stellen. Tijdens de les worden korte momenten van instructie of reflectie afgewisseld met het toepassen van nieuwe vaardigheden en leerkrachten functioneren als mentor en rolmodel (De Bruijn et al., 2005; Smit et al., 2014).

Een universele schoolinterventie die aansluit bij deze actieve leerstijl van vmbo-leerlingen is Rots en Water (R&W; Ykema, 2002, 2018; <https://www.rotsenwater.nl>). R&W onderscheidt zich met name van andere schoolinterventies door een actieve lesmethode waarin de nadruk ligt op activerende werkvormen in plaats van passieve verbale uitleg; tijdens de R&W lessen worden klassikale moment van informatie, instructie of reflectie afgewisseld met actieve oefeningen waarin nieuwe vaardigheden geoefend kunnen worden. De R&W trainer (de leerkracht) ondersteunt leerlingen op een positieve manier en geeft, als rolmodel, tijdens de lessen het goede voorbeeld. R&W is gebaseerd op de theorie van het 'R&W huis' (Ykema, 2002, 2018). Dit huis is gebouwd op vier pilaren: zelfcontrole, zelfreflectie, zelfvertrouwen, en emotionele zelfregulatie. Deze vier competenties worden beschouwd als specifieke uitkomsten van de interventie en vormen de basisvaardigheden voor de ontwikkeling van de bredere, overkoepelende, competenties. Volgens de theorie van R&W, zou het verbeteren van deze vier competenties samengaan met verbeteringen in psychosociaal welzijn, veerkracht, en seksuele autonomie en het verminderen van internaliserend gedrag (het intrapersoonlijk domein). Door het creëren van een veilige en respectvolle omgeving en door leerlingen alternatieve gedragingen aan te bieden zou volgens de theorie van R&W externaliserend gedrag, agressie en pesten verminderen en positieve relaties tussen leerlingen toenemen (het interpersoonlijk domein).

Hoewel R&W in veel landen geïmplementeerd wordt (bijv. Frankrijk, Australië, China), en met name in Nederland vaak toegepast wordt, is er weinig bekend over de effectiviteit. Met name in Nederland is het belangrijk meer inzicht te krijgen in de effectiviteit van R&W, niet alleen omdat deze interventie vaak wordt geïmplementeerd in Nederland, maar ook omdat het ingezet wordt binnen de kaders van de Wet Veiligheid op school – een wet die scholen verplicht zich actief in te spannen om pesten tegen te gaan en een veilig sociaal schoolklimaat te creëren. In Australië hebben professionals in het werkveld verschillende kleinschalige studies uitgevoerd, vaak op basis- en middelbare scholen, die hebben aangetoond dat deelnemers zich na het afronden van R&W veerkrachtiger voelden, een meer positieve identiteit hadden, en meer actieve copingstrategieën toepasten (Ykema et al., 2006). Een onderzoek in Nederland onder (mannelijke) vmbo-leerlingen van De Graaf en collega's (2016) heeft de effectiviteit van R&W op seksuele agressie onderzocht. Na het afronden van R&W verminderden zelfgerapporteerde dwingende strategieën en verbale manipulatie, terwijl zelfregulatie toenam. Ondanks deze veelbelovende resultaten is een grootschalige studie nodig die zich richt op bredere uitkomstmaten, zowel jongens als meisjes includeert, en in het bijzonder de effectiviteit onderzoekt voor vmbo-leerlingen. Daarom was

het eerste doel van de huidige studie om te onderzoeken of R&W effectief is voor vmbo-leerlingen in het verbeteren van competenties en het voorkomen van problemen in het intrapersoonlijk en het interpersoonlijk domein. Verwacht werd dat leerlingen in de R&W conditie een significante verbetering zouden laten zien op alle gemeten interventie-specifieke en overkoepelende (intra- en interpersoonlijke) uitkomsten ten opzichte van leerlingen in de controle conditie.

Het tweede doel was om te onderzoeken of interventie-effecten afhankelijk zijn van de mate waarin verschillende systemen betrokken zijn bij de interventie. Volgens het ecologisch model van Bronfenbrenner (1979), wordt gedrag bepaald door interacties tussen verschillende systemen zoals individu-, familie-, klas-, en schoolsystemen. Het actief betrekken van gerelateerde systemen bij een interventie, zoals de school- en familiesystemen, zou mogelijk zowel generalisatie als behoud van nieuwe vaardigheden kunnen bevorderen (Mahoney et al., 2021); de inhoud van de interventie wordt dan namelijk ingebed in een groot deel van de sociale omgeving van het individu wat de mogelijke invloed ervan op het individu vergroot. Zo concludeerden Mertens en collega's (2020) in hun meta-analyse naar componenten van universele schoolinterventies dat betrokkenheid van de volledige school en/of ouders gerelateerd was aan sterkere interventie-effecten. Echter, Durlak en collega's (2011) vonden in hun meta-analyse van *social emotional learning* (SEL) interventies dat interventies waarbij de hele school en/of ouders betrokken waren niet effectiever waren dan interventies die alleen in de klas gegeven werden. Deze inconsistente bevindingen geven aan dat er meer inzicht nodig is in hoe betrokkenheid van meerdere systemen in een interventie de effectiviteit kan beïnvloeden (Kuosmanen et al., 2019).

De huidige studie onderzoekt daarom of er verschil is in de effectiviteit van drie verschillende implementatieversies van R&W: betrokkenheid van 1) een kernteam van leerkrachten, 2) een kernteam en het volledige schoolteam, en 3) een kernteam, het volledige schoolteam en ouders. We hadden geen specifieke hypothese aangezien onderzoek naar de betrokkenheid van meerdere ecologische systemen in interventies inconsistente resultaten laat zien.

Methode

Verschillende versies van R&W

We hebben drie versies van R&W onderzocht die verschillen in het aantal systemen dat betrokken was bij de interventie: 1) *Light*: alleen een kernteam van leerkrachten (die de interventielessen gaven), 2) *Standaard*: een kernteam van leerkrachten en de rest van het schoolteam, en 3) *Plus*: een kernteam, het schoolteam, en ouders. In alle drie de condities

werden de interventielessen gegeven door een kernteam van leerkrachten (R&W trainers) die de 3-daagse training succesvol hadden afgerond. Dit kernteam ontving supervisie van een persoonlijke R&W coach. In de Light conditie was alleen een kernteam getraind en op de hoogte van R&W. In de Standaard en de Plus condities had, naast het kernteam van R&W trainers, de rest van het schoolteam een 3-daagse introductiecursus gevolgd over de basisprincipes van R&W, ondersteuning van R&W trainers, en toepassing van de interventietechnieken in reguliere lessen. In de Plus conditie kregen daarnaast ouders de mogelijkheid een documentaire over R&W te bekijken, werden ze uitgenodigd voor een R&W les op school, en ontvingen ze wekelijks e-mails met daarin informatie over de R&W les van die week en werden ze aangemoedigd om met deze informatie thuis aan de slag te gaan.

Design en procedure

De effectiviteit van deze drie verschillende versies van R&W (Light, Standaard en Plus) is onderzocht in scholen met een vmbo-traject in een gerandomiseerde gecontroleerde studie met een actieve controle conditie. In de Controle conditie voerden scholen hun gebruikelijk beleid voor het verbeteren van competenties en het voorkomen van problemen bij leerlingen. Scholen werden geëxcludeerd als ze in de afgelopen twee jaar R&W hadden geïmplementeerd. Dertien scholen verspreid door Nederland werden willekeurig toegewezen aan de condities door gestratificeerde blok randomisatie met blokken van vier. Scholen werden gestratificeerd op schoolgrootte (< 100 leerlingen in de eerste klas, > 100 leerlingen in de eerste klas) om een gelijke verdeling van leerlingen over de condities te bevorderen. Eén school uit de Controle conditie viel uit na randomisatie maar voor het begin van de dataverzameling vanwege een verandering in schoolmanagement. Deze school werd vervangen door een andere geïnteresseerde school (zie Figuur 1 voor de flowchart).

[Figuur 1]

Leerlingen vulden digitale vragenlijsten in, afgenomen door getrainde onderzoeksassistenten, voor de start van de interventie (T1 = oktober/november 2017), na het afronden van het eerste jaar van de R&W lessen (T2 = maart/april 2018), voor de start van de R&W lessen in het tweede jaar van het R&W programma (T3 = oktober 2018), en na de interventie (T4 = januari 2019). Leerlingen gaven actieve, geïnformeerde toestemming voor het invullen van de vragenlijsten. Ouders gaven passieve, geïnformeerde toestemming voor deelname van hun kind. Deze studie is goedgekeurd door de Ethische Commissie [GEMASKEERD VOOR REVIEW] en geregistreerd [GEMASKEERD VOOR REVIEW] (voor protocol zie [GEMASKEERD VOOR REVIEW]).

Tabel 1

Demografische gegevens van leerlingen op de voormeting per conditie

	Light	Standaard	Plus	Controle	Verschillen op T1		
					F/χ^2	p	$\eta^2_{\text{partial}}/\phi$
<i>N</i>	373	303	249	374			
Leeftijd, <i>M</i> (<i>SD</i>)	12,33 (0,57)	12,38 (0,66)	12,34 (0,60)	12,47 (0,64)	3,89	,009	0,009
Jongens, <i>n</i> (%)	170 (48%)	161 (56%)	131 (55%)	199 (57%)	7,38	,061	0,077
Westerse achtergrond, <i>n</i> (%)	291 (82%)	115 (43%)	211 (91%)	198 (59%)	182,01	< ,001	0,392

Participanten

Op de voormeting bestond de steekproef uit 1299 eerstejaars vmbo-leerlingen ($M_{\text{leeftijd}} = 12,38$ jaar; $SD = 0,62$; 54% jongens; 69% met een Westerse achtergrond; zie Tabel 1).

Tussen de condities waren geen verschillen in sekse, maar wel in leeftijd en migratieachtergrond. Leerlingen in de Controle conditie waren ouder dan leerlingen in de Light conditie. De Controle en Standaard condities hadden ongeveer een gelijke verdeling tussen leerlingen met een Westerse en een niet-Westerse achtergrond, terwijl in de Light en Plus condities leerlingen vooral een Westerse achtergrond hadden. Daarom werd in alle analyses gecontroleerd voor leeftijd en migratieachtergrond. Een MANOVA liet zien dat er bij de start van de studie kleine, verwaarloosbare, verschillen tussen de condities waren op de uitkomstmaten ($F(39, 3591) = 2.12, p < .001, \eta^2_{\text{partial}} = .023$). De Light conditie scoorde enigszins slechter dan de Controle conditie op zelfcontrole, veerkracht, seksuele autonomie, internaliserende problemen, en externaliserend gedrag. Verder scoorde de Plus conditie iets beter op interpersoonlijke relaties in de klas dan de Light en Controle condities. In alle analyses is gecontroleerd voor de voormeting van de desbetreffende uitkomst, ongeacht of er bij aanvang significante verschillen waren tussen de condities.

Over alle meetmomenten ontbrak 15% van de data. Vergelijkingen tussen leerlingen die uitgevallen zijn ($n_{T2} = 68, n_{T3} = 83, n_{T4} = 60$) en leerlingen die de studie volledig hebben afgerond, lieten geen verschillen zien op demografische kenmerken (leeftijd, sekse verdeling en migratieachtergrond) noch op de uitkomstvariabelen.

Scholen bepaalden zelfstandig welke leerkrachten getraind werden in het geven van R&W lessen. De R&W lessen werden in totaal verzorgd door 49 leerkrachten (trainers) tussen de 23 en 64 jaar oud ($M = 36,77, SD = 10,02$) met gemiddeld 12,98 jaar aan leerkrach-tervaring. De meeste trainers waren gymdocent ($n = 30; 71\%$). Iets meer dan de helft van de

trainers was man ($n = 27$; 56%) en bijna alle trainers hadden een Westerse achtergrond ($n = 46$; 96%).

Conditie en Uitvoering

R&W

R&W (Ykema, 2002, 2018) integreert spel en opdrachten om leerlingen te begeleiden in het maken van (fysiek) contact met anderen en het ontdekken, respecteren en stellen van zowel eigen grenzen als grenzen van anderen. Middels oefeningen worden lichaamsbewustzijn (bijv. spanning in de spieren, ademhaling), non-verbale communicatie, en copingstrategieën van leerlingen versterkt. In een oefening in tweetallen moeten leerlingen bijvoorbeeld proberen de andere leerling uit balans te tikken met als doel leerlingen te laten ervaren dat het soms beter is om rigide te zijn en druk van anderen te weerstaan (als een rots) en dat het soms beter is om flexibel te zijn en open te staan voor de mening, gedachten, en gevoelens van anderen (als water).

Naast deze spellen hanteert R&W ook een gedragsmatige aanpak, zoals het gebruik van rolspellen om te oefenen met bijvoorbeeld het stellen en aangeven van grenzen, kalm reageren op provocaties, en het helpen van leerlingen die gepest worden. Tijdens en na de oefeningen geven leerlingen feedback en reflecteren ze. Trainers tonen en bekrachtigen gewenst gedrag, en creëren een veilige omgeving waarin leerlingen fouten mogen maken.

De interventie bestaat uit 22 lessen verspreid over twee schooljaren: 14 lessen in het eerste en 8 lessen in het tweede leerjaar. De lessen werden wekelijks gegeven tijdens gymlessen van 90 minuten.

Controle conditie

In de Controle conditie voerden scholen hun gebruikelijk beleid. Leerlingen hadden bijvoorbeeld een mentor of persoonlijke coach met wie ze regelmatig afspraken hadden, hun welzijn bespraken en naar wie ze toe konden voor advies. Ook was er bijvoorbeeld een projectweek over ‘anders zijn’, tekenden leerlingen een anti-pestcontract, werd pesten in de klas besproken en was er een ‘anti-pestcoördinator’.

Uitvoering interventie

De uitvoering van R&W is gemeten op twee manieren. Ten eerste, R&W trainers vulden een vragenlijst in na elke derde les in het eerste jaar en elke tweede les in het tweede jaar van de interventie. Deze vragenlijst bestond uit vragen over het verloop van de R&W lessen, zoals “Heeft u in de afgelopen drie/twee lessen van R&W alles kunnen behandelen?” (1 = *alle lessen volledig kunnen geven* tot 4 = *geen les volledig kunnen geven*) en “In hoeverre bent u afgeweken van het R&W handboek?” (1 = *bijna niet* tot 6 = *heel veel*). Ten tweede,

drie R&W experts, die alle drie de R&W training voor gevorderden hebben afgerond en meerdere jaren ervaring hebben met het implementeren van R&W, hebben 67 R&W lessen (Light = 28 lessen; Standaard = 17 lessen, Plus = 22 lessen) geobserveerd. Tijdens deze lesobservatie hebben zij een schema ingevuld, gebaseerd op Bishop en collega's (2014), waarin ze onder andere hun subjectieve beoordeling gaven over de mate waarin de handleiding werd gevolgd, aanpassingen die werden gedaan en de kwaliteit van de les.

Volgens de zelfrapportages hebben R&W trainers over het algemeen de lessen volledig kunnen geven (65%) en zijn ze niet of nauwelijks afgeweken van de handleiding (72%). De R&W experts observeerden dat de meeste lessen volledig of bijna volledig uitgevoerd waren (86%). Trainers weken niet veel af van de handleiding (91%). Eventuele aanpassingen aan de uitvoering werden over het algemeen gezien als verbeteringen (62%). Over het algemeen beoordeelden de R&W experts de lessen van goede (38%) tot zeer goede (54%) kwaliteit. De uitvoering van de R&W lessen verschilde niet tussen condities noch tussen de twee jaar van implementatie, zowel volgens de zelfrapportages als volgens de observaties. Concluderend, het grootste deel van de R&W lessen was daadwerkelijk gegeven en de uitvoering lijkt goed tot zeer goed te zijn geweest.

Om te bepalen of de Standaard en Plus versies geïmplementeerd waren zoals bedoeld, hebben R&W trainers en ouders vragenlijsten ingevuld na het eerste jaar en na het tweede jaar van de interventie. Zoals gepland waren er in de Standaard en Plus condities meer leerkrachten betrokken bij de interventie dan in de Light conditie ($F(2,24) = 5,73; p = ,009$). Ouders in de Plus conditie die de vragenlijst hebben ingevuld ($n = 47$) rapporteerden dat zij de wekelijkse informatie soms (57%) tot vaak (39%) lazen, en sommige ouders gaven aan de documentaire over R&W gekeken te hebben (15%) en deelgenomen te hebben aan een R&W les op school (23%). Verder rapporteerden ouders in de Plus conditie, vergeleken met ouders in de Standaard conditie, thuis vaker over R&W gesproken te hebben ($F(2,75) = 4,37; p = ,016$) en het vaker bij hun eigen kind toegepast te hebben ($F(2,75) = 3,44; p = ,037$). Kortom, ouders in de Plus conditie waren zoals gepland meer betrokken bij de interventie dan ouders in de Light en Standaard condities. Dus, zowel de leerkracht- als de oudercomponent blijkt geïmplementeerd te zijn.

Interventie-Specifieke Uitkomsten

Zelfcontrole

De vaardigheid van leerlingen om hun impulsen te controleren en ongepast gedrag te onderbreken werd gemeten met de korte versie van de *Self-Control Scale* (Finkenauer et al.,

2005) die bestaat uit 11 items (bijv. “Ik kan goed weerstand bieden tegen verleidingen”, $\alpha = 0,62-0,72$) beantwoord op een 5-puntsschaal ($1 = \text{helemaal niet}$ tot $5 = \text{helemaal wel}$).

Zelfreflectie

Leerlingen vulden de *Engage* subschaal in van de *Self-Reflection and Insight Scale* (Sauter et al., 2010) om de mate van reflectie op persoonlijke gedachten, gevoelens en gedrag te meten. De subschaal bestaat uit 6 items (bijv. “Ik denk vaak na over hoe ik me voel over iets”) beantwoord op een 6-puntsschaal ($1 = \text{helemaal mee oneens}$ tot $6 = \text{helemaal mee eens}$). De betrouwbaarheid was laag ($\alpha = 0,53-0,60$). Verwijdering van 3 negatief geformuleerde items (“Ik ben bijna nooit bezig met ‘zelfreflectie’”, “Ik denk niet vaak na over mijn gedachten”, en “Ik denk niet echt na over de reden waarom ik me gedraag zoals ik dat doe”) verbeterde de betrouwbaarheid ($\alpha = 0,74-0,89$).

Zelfvertrouwen

Globaal zelfvertrouwen van leerlingen is gemeten met de subschaal *Global self-perception* van de *Self-Perception Profile* (Harter, 1988). Deze subschaal bestaat uit 5 items (bijv. “Ik ben best tevreden met mezelf”) beantwoord op een 4-puntsschaal ($1 = \text{helemaal niet waar}$ tot $4 = \text{helemaal waar}$; $\alpha = 0,73-0,75$).

Emotionele zelfregulatie

Vaardigheden in het controleren van emoties en het gebruik van emotieregulatie strategieën werd gemeten met de *Difficulties in Emotion Regulation Scale* (Anderson et al., 2016). De vragenlijst bestaat uit 14 items (bijv. “Als ik van streek ben, kan ik iets bedenken om ervoor te zorgen dat ik me weer beter voel”) beantwoord op een 5-puntsschaal ($1 = \text{bijna nooit}$ tot $5 = \text{bijna altijd}$; $\alpha = 0,88-0,91$).

Overkoepelende Uitkomsten

Intrapersoonlijk domein

Psychologisch welzijn. Om de aanwezigheid van positieve emoties te meten, hebben leerlingen de subschaal *Psychological Wellbeing* ingevuld van de KIDSCREEN-27 (Ravens-Sieberer & The European KIDSCREEN Group, 2006). De subschaal bestaat uit 7 items (bijv. “In de afgelopen week, heb je lol gehad?”) beantwoord op een 5-puntsschaal ($1 = \text{nooit}$ tot $5 = \text{altijd}$; $\alpha = 0,76-0,83$).

Veerkracht. De vaardigheid om terug te veren na uitdagingen werd gemeten middels de *Connor-Davidson Resilience Scale-Short Version* (Davidson & Connor, 2017) bestaande uit 10 items (bijv. “Aanpassen aan verandering”) beantwoord op een 5-puntsschaal ($0 = \text{helemaal niet waar}$ tot $4 = \text{bijna altijd waar}$; $\alpha = 0,79-0,92$).

Seksuele autonomie. Coping vaardigheden in seksuele situaties werd gemeten met 5 items van de *Seks onder je 25^{ste}* vragenlijst (bijv. “Wanneer ik ben met iemand die ik heel erg leuk vind, voel ik me helemaal op mijn gemak”; De Graaf et al., 2005) beantwoord op een 4-puntsschaal (1 = *nooit* tot 4 = *altijd*). Betrouwbaarheid was laag op T1 ($\alpha = 0,53$) en redelijk op T2, T3, en T4 ($\alpha = 0,62-0,65$).

Internaliserend gedrag. Leerlingen vulden de internaliserende subschaal in van de korte versie van de *Youth Self Report* (Chorpita et al., 2010). De subschaal bestaat uit 6 items (bijv. “Ik maak me vaak zorgen”) beantwoord op een 3-puntsschaal (0 = *nooit* tot 2 = *vaak*; $\alpha = 0,79-0,86$).

Interpersoonlijk domein

Interpersoonlijke relaties in de klas. De manier waarop leerlingen relaties in de klas ervaren, werd gemeten met de subschalen Conflict, Cohesie en Comfort van de *Classroom Peer Context Questionnaire* (Boor-Klip et al., 2016). Deze 12 items (bijv. “In deze klas vinden kinderen elkaar aardig”) werden beantwoord op een 5-puntsschaal (1 = *helemaal niet waar* tot 5 = *helemaal waar*; $\alpha = 0,80-0,85$).

Externaliserend gedrag. Leerlingen hebben de externaliserende subschaal ingevuld van de korte versie van de *Youth Self Report* (Chorpita et al., 2010). De subschaal bestaat uit 6 items (bijv. “Ik maak veel ruzie”) beantwoord op een 3-puntsschaal (0 = *nooit* tot 2 = *vaak*; $\alpha = 0,65-0,79$).

Agressie. Agressie is gemeten met de *Reactive and Proactive Aggression Questionnaire* (Dodge & Coie, 1987). De vragenlijst bestaat uit 6 items (bijv. “Als ze mij plagen of tegen mij liegen, reageer ik kwaad”) beantwoord op een 5-puntsschaal (1 = *nooit* tot 5 = *bijna altijd*; $\alpha = 0,65-0,83$).

Pesten en gepest worden. De frequenties van pesten en ervaringen van gepest worden zijn gemeten met 2 items van de *Olweus Bully/Victim Questionnaire* (Solberg & Olweus, 2003): “Hoe vaak heb je de afgelopen twee maanden meegedaan aan het pesten van andere leerlingen op school?” en “Hoe vaak ben je de afgelopen twee maanden op school gepest?”. Leerlingen beantwoordden deze vragen op een 5-puntsschaal (1 = *nooit* tot 5 = *meerdere keren per week*).

Statistische Analyses²

Data zijn geanalyseerd op basis van een *intention-to-treat* benadering waarbij leerlingen toegewezen aan de interventie geïnccludeerd zijn in de analyse ongeacht of ze

² Uit de power analyse in het protocol paper (zie GEMASKEERD VOOR REVIEW) blijkt een minimale steekproefgrootte van 720 participanten nodig voor deze analyses.

daadwerkelijk hebben deelgenomen aan de interventie. We hebben 1-level-modellen geanalyseerd met een correctie voor participanten geclusterd in scholen middels de *complex sample cluster* in *Mplus* (versie 8.2; Muthén & Muthén, 1998-2010). Om alle participanten in de modellen te kunnen includeren, hebben we *Full Information Maximum Likelihood* (FIML) procedures gebruikt. Parameters werden verkregen met *Robust Maximum Likelihood* schattingen (MLR; Muthén & Muthén, 1998-2010).

Om de effecten van de interventie te analyseren, hebben we een serie latente groeicurve (LGC) modellen getest. LGC modellen schatten voor elke participant een individueel groeitraject gebaseerd op hun basisniveau (= intercept) en verandering over tijd (= slope). De individuele groeicurves zijn de indicatoren van latente variabelen die de groeitrajecten van de groep beschrijven (Muthén & Muthén, 1998-2010). De slope is het belangrijkste; wanneer de interventie vergeleken met de Controle conditie effectief is, verandert de interventie de slope in de gewenste richting.

Voor het bepalen van de interventie-effecten, zijn drie dummy variabelen gemaakt (Light, Standaard, en Plus conditie) met de Controle conditie als referentiegroep en toegevoegd als voorspellers van de intercept en slope. Leeftijd en migratieachtergrond van de leerlingen zijn toegevoegd als covariaten. Wanneer twee of meer interventiecondities vergeleken met de Controle conditie effectief bleken, zijn de condities onderling vergeleken met een multi-groep model door de slopes aan elkaar gelijk te stellen en vrij te schatten. De model fits van deze twee geneste modellen zijn vergeleken met de *Satorra-Bentler Scale Chi-Square test* (Satorra & Bentler, 2010). Wanneer deze significant is, past het vrij geschatte model beter bij de data en is dus de ene interventieconditie effectiever dan de andere.

We hebben effectgroottes berekend door de mate van verandering te vermenigvuldigen met de tijdsspanne van de interventie en dat vervolgens gedeeld door de standaarddeviatie van de desbetreffende uitkomst ($d = (\text{slope} * \text{tijdsspanne})/SD$; Feingold, 2013). We hebben effectgroottes berekend tussen de verschillende meetmomenten (verandering van T1 naar T2, van T2 naar T3, en van T3 naar T4) en de algemene verandering (verandering van T1 naar T4).

Resultaten

Effecten van R&W

In Tabel 2 staan de gemiddelden en standaarddeviaties van de uitkomsten per conditie op elk meetmoment. De LGC modellen hadden in het algemeen een acceptabele fit (zie voor details [GEMASKEERD VOOR REVIEW]).

De gestandaardiseerde regressiecoëfficiënten van de interventie condities op de slope zijn gerapporteerd in Tabel 3. De effectgroottes tussen de verschillende meetmomenten en de algemene effectgroottes van de interventie condities (in vergelijking met de Controle conditie) zijn gerapporteerd per uitkomst in Tabel 4.

Tabel 2

Gemiddelden en standaarddeviaties van uitkomsten per conditie per meetmoment

	Light				Standaard				Plus				Controle			
	T1	T2	T3	T4	T1	T2	T3	T4	T1	T2	T3	T4	T1	T2	T3	T4
	<i>M</i>	<i>M</i>	<i>M</i>	<i>M</i>	<i>M</i>	<i>M</i>	<i>M</i>	<i>M</i>	<i>M</i>	<i>M</i>	<i>M</i>	<i>M</i>	<i>M</i>	<i>M</i>	<i>M</i>	<i>M</i>
	(<i>SD</i>)	(<i>SD</i>)	(<i>SD</i>)	(<i>SD</i>)	(<i>SD</i>)	(<i>SD</i>)	(<i>SD</i>)	(<i>SD</i>)	(<i>SD</i>)	(<i>SD</i>)	(<i>SD</i>)	(<i>SD</i>)	(<i>SD</i>)	(<i>SD</i>)	(<i>SD</i>)	(<i>SD</i>)
<i>Interventie-specifieke uitkomsten</i>																
Zelfcontrole	3,33	3,53	3,52	3,54	3,49	3,66	3,56	3,61	3,41	3,53	3,43	3,53	3,49	3,59	3,57	3,58
	(0,58)	(0,60)	(0,61)	(0,60)	(0,57)	(0,66)	(0,65)	(0,69)	(0,60)	(0,59)	(0,61)	(0,65)	(0,56)	(0,62)	(0,65)	(0,64)
Zelfreflectie	2,92	2,83	2,80	2,76	3,02	2,83	2,76	2,90	2,90	2,96	2,85	2,81	3,09	2,95	2,98	2,86
	(1,20)	(1,32)	(1,27)	(1,36)	(1,31)	(1,31)	(1,32)	(1,42)	(1,22)	(1,33)	(1,36)	(1,52)	(1,21)	(1,44)	(1,35)	(1,43)
Zelfvertrouwen	3,23	3,28	3,24	3,28	3,28	3,33	3,29	3,29	3,25	3,33	3,18	3,12	3,31	3,29	3,22	3,22
	(0,62)	(0,66)	(0,65)	(0,62)	(0,57)	(0,60)	(0,66)	(0,62)	(0,59)	(0,57)	(0,68)	(0,70)	(0,59)	(0,64)	(0,65)	(0,65)
Emotionele zelfregulatie	3,83	3,94	3,98	3,97	3,85	3,99	3,91	3,91	3,87	3,88	3,81	3,83	3,93	3,94	3,90	3,92
	(0,77)	(0,85)	(0,76)	(0,75)	(0,69)	(0,80)	(0,82)	(0,84)	(0,71)	(0,74)	(0,80)	(0,85)	(0,71)	(0,75)	(0,76)	(0,76)
<i>Intrapersoonlijk domein</i>																
Psychologisch welzijn	3,94	3,98	3,91	3,87	4,02	4,10	4,02	3,93	3,97	4,06	3,87	3,85	4,06	4,02	3,90	3,92
	(0,62)	(0,69)	(0,71)	(0,78)	(0,61)	(0,71)	(0,73)	(0,77)	(0,62)	(0,63)	(0,72)	(0,80)	(0,59)	(0,68)	(0,75)	(0,78)
Veerkracht	2,36	2,49	2,15	2,47	2,43	2,70	2,56	2,53	2,43	2,57	2,47	2,51	2,51	2,68	2,57	2,61
	(0,65)	(0,75)	(0,87)	(0,82)	(0,67)	(0,71)	(0,70)	(0,90)	(0,64)	(0,68)	(0,71)	(0,84)	(0,65)	(0,69)	(0,74)	(0,85)
Seksuele autonomie	3,23	3,39	3,40	3,40	3,30	3,44	3,38	3,39	3,29	3,38	3,34	3,35	3,32	3,36	3,32	3,36
	(0,44)	(0,48)	(0,44)	(0,46)	(0,44)	(0,52)	(0,54)	(0,51)	(0,40)	(0,46)	(0,45)	(0,57)	(0,44)	(0,49)	(0,50)	(0,52)
Internaliserend gedrag	0,53	0,41	0,44	0,42	0,44	0,33	0,40	0,37	0,49	0,45	0,51	0,48	0,43	0,39	0,45	0,42
	(0,42)	(0,44)	(0,44)	(0,46)	(0,42)	(0,39)	(0,44)	(0,43)	(0,40)	(0,42)	(0,44)	(0,51)	(0,39)	(0,41)	(0,46)	(0,44)
<i>Interpersoonlijk domein</i>																

Interpersoonlijke relaties in de klas	3,90 (0,64)	3,81 (0,68)	3,83 (0,67)	3,73 (0,69)	3,98 (0,63)	3,87 (0,72)	3,81 (0,73)	3,80 (0,75)	4,07 (0,62)	4,00 (0,63)	3,92 (0,69)	3,84 (0,81)	3,91 (0,59)	3,78 (0,65)	3,77 (0,67)	3,75 (0,72)
Externaliserend gedrag	0,59 (0,38)	0,47 (0,39)	0,48 (0,39)	0,45 (0,41)	0,54 (0,39)	0,43 (0,38)	0,50 (0,44)	0,42 (0,42)	0,55 (0,35)	0,46 (0,36)	0,55 (0,38)	0,49 (0,45)	0,49 (0,34)	0,42 (0,37)	0,45 (0,39)	0,43 (0,39)
Agressie	1,79 (0,56)	1,71 (0,59)	1,69 (0,59)	1,66 (0,66)	1,85 (0,61)	1,84 (0,73)	1,94 (0,73)	1,83 (0,81)	1,73 (0,46)	1,72 (0,60)	1,80 (0,58)	1,74 (0,81)	1,78 (0,56)	1,76 (0,65)	1,81 (0,67)	1,85 (0,78)
Pesten	1,13 (0,48)	1,18 (0,61)	1,15 (0,55)	1,21 (0,64)	1,13 (0,41)	1,18 (0,60)	1,15 (0,58)	1,19 (0,64)	1,10 (0,41)	1,15 (0,47)	1,20 (0,68)	1,20 (0,64)	1,08 (0,36)	1,18 (0,59)	1,22 (0,71)	1,22 (0,71)
Gepest worden	1,32 (0,77)	1,43 (0,96)	1,30 (0,77)	1,31 (0,86)	1,28 (0,73)	1,35 (0,87)	1,22 (0,77)	1,22 (0,70)	1,36 (0,88)	1,37 (0,85)	1,40 (0,96)	1,40 (0,93)	1,37 (0,94)	1,56 (1,16)	1,41 (1,04)	1,34 (0,95)

Tabel 3

Interventie-effecten over tijd van de R&W condities in vergelijking met de controle conditie

	Light			Standaard			Plus		
	β	SE	p	β	SE	p	β	SE	p
<i>Interventie-specifieke uitkomsten</i>									
Zelfcontrole ¹	0,12	0,03	< ,001	0,00	0,04	,951	0,02	0,04	,532
Zelfreflectie	0,01	0,04	,803	-0,02	0,06	,761	0,04	0,03	,238
Zelfvertrouwen	0,18	0,10	,060	0,09	0,11	,394	-0,05	0,06	,327
Emotionele zelfregulatie	0,17	0,05	< ,001	0,07	0,05	,163	-0,01	0,02	,676
<i>Intrapersoonlijk domein</i>									
Psychologische welzijn	0,15	0,06	,009	0,08	0,08	,287	0,04	0,04	,228
Veerkracht	0,09	0,73	,902	0,18	0,27	,507	-0,10	0,31	,736
Seksuele autonomie ¹	0,18	0,04	< ,001	0,08	0,06	,144	0,06	0,03	,074
Internaliserend gedrag ¹	-0,13	0,04	< ,001	-0,07	0,04	,048	0,01	0,03	,812
<i>Interpersoonlijk domein</i>									
Interpersoonlijke relaties in de klas	0,05	0,03	,056	-0,05	0,07	,499	-0,04	0,04	,278
Externaliserend gedrag	-0,20	0,12	,107	-0,13	0,11	,223	-0,02	0,07	,756
Agressie	-0,19	0,09	,030	0,01	0,10	,933	-0,01	0,05	,822
Pesten	-0,12	0,06	,056	-0,12	0,09	,149	-0,05	0,09	,624
Gepest worden	-0,01	0,04	,817	-0,04	0,06	,527	0,03	0,04	,530

Noot. ¹ Variantie van de voormeting van de desbetreffende uitkomst variabele was vastgezet op nul vanwege negatieve residuele variantie van de geobserveerde variabele op T1.

Tabel 4

Effectgroottes van verandering over tijd in R&W condities in vergelijking met de controle conditie

	Light				Standaard				Plus			
	T1-T2	T2-T3	T3-T4	T1-T4	T1-T2	T2-T3	T3-T4	T1-T4	T1-T2	T2-T3	T3-T4	T1-T4
<i>Interventie-specifieke uitkomsten</i>												
Zelfcontrole ¹	0,19	0,03	0,03	0,25	0,00	0,00	0,00	0,01	0,04	0,01	0,01	0,06
Zelfreflectie	0,01	0,00	0,00	0,02	-0,02	0,00	0,00	-0,03	0,06	0,01	0,00	0,07
Zelfvertrouwen	0,04	0,15	0,04	0,22	0,02	0,08	0,02	0,12	-0,01	-0,05	-0,01	-0,08
Emotionele zelfregulatie	0,17	0,09	0,04	0,29	0,08	0,04	0,02	0,13	-0,01	0,00	0,00	-0,01
<i>Intrapersoonlijk domein</i>												
Psychologisch welzijn	0,07	0,18	-0,01	0,24	0,04	0,10	-0,01	0,14	0,02	0,06	0,00	0,08
Veerkracht	0,04	-0,03	0,02	0,03	0,09	-0,07	0,04	0,06	-0,05	0,04	-0,02	-0,04
Seksuele autonomie ¹	0,32	0,06	0,01	0,38	0,15	0,03	0,00	0,18	0,12	0,02	0,00	0,15
Internaliserend gedrag ¹	0,19	0,04	0,03	0,26	0,11	0,03	0,02	0,15	-0,01	0,00	0,00	-0,01
<i>Interpersoonlijk domein</i>												
Interpersoonlijke relaties in de klas	0,05	0,03	0,02	0,10	-0,05	-0,03	-0,02	-0,10	-0,04	-0,02	-0,02	-0,08
Externaliserend gedrag	0,13	-0,05	0,03	0,11	0,09	-0,03	0,02	0,08	0,02	-0,01	0,00	0,02
Agressie	0,18	0,10	0,03	0,31	-0,01	-0,01	0,00	-0,02	0,01	0,01	0,00	0,02
Pesten	0,10	0,08	0,06	0,24	0,11	0,09	0,07	0,28	0,05	0,04	0,03	0,11
Gepest worden	0,01	0,01	0,00	0,01	0,04	0,02	0,00	0,06	-0,03	-0,02	0,00	-0,04

Noot. ¹ Variantie van de voormeting van de desbetreffende uitkomst variabele was vastgezet op nul vanwege negatieve residuele variantie van de geobserveerde variabele op T1; Effectgroottes (*d*) zijn dusdanig berekend dat positieve effectgroottes een verandering in de gewenste richting voor R&W aangeven.

Interventie-Specifieke Uitkomsten

Leerlingen in de Light conditie toonden, vergeleken met de Controle conditie, gunstigere veranderingstrajecten voor zelfcontrole ($d = 0,25$) en emotionele zelfregulatie ($d = 0,29$; zie Figuur 2). Zij hadden een sterkere stijging in zelfcontrole (slope = 1,75) dan leerlingen in de Controle conditie (slope = 1,44). Met betrekking tot emotionele zelfregulatie lieten leerlingen in de Light conditie een lichte verbetering over tijd zien (slope = 0,24), terwijl leerlingen in de Controle conditie een lichte afname lieten zien (slope = -0,10). Op beide uitkomsten verbeterden leerlingen met name in het eerste jaar van de interventie. Verder laten de resultaten een relevant ($d = 0,22$), maar niet significant ($p = ,060$), verschil zien tussen de veranderingstrajecten van zelfvertrouwen in de Light conditie en Controle conditie. Leerlingen in de Light conditie bleven relatief stabiel (slope = -1,68), terwijl leerlingen in de Controle conditie een kleine afname in zelfvertrouwen rapporteerden (slope = -2,58). Dit interventie-effect was het sterkst tussen T2 en T3. We hebben geen interventie-effecten gevonden voor zelfreflectie noch voor leerlingen in de Standaard en Plus condities. Over het algemeen leek de Light conditie de meest gunstige veranderingstrajecten te bewerkstelligen op de interventie specifieke uitkomsten, met name in het begin van de interventie.

[Figuur 2]

Overkoepelende Uitkomsten

Intrapersoonlijk domein. Ook voor psychologisch welzijn, seksuele autonomie, en internaliserend gedrag hadden leerlingen in de Light conditie positievere veranderingstrajecten dan leerlingen in de Controle conditie (range $d = 0,26-0,38$; zie Figuur 3). Zij bleven stabiel in hun psychologisch welzijn (slope = 0,08), terwijl leerlingen in de Controle conditie verslechterden (slope = -0,25). Het interventie-effect was het sterkste tussen T2 en T3. Op seksuele autonomie verbeterden leerlingen in de Light conditie sterker (slope = 0,73) dan leerlingen in de Controle conditie (slope = 0,32), vooral in het eerste jaar van de interventie. Wat betreft internaliserend gedrag rapporteerden leerlingen in de Light en Standaard condities een sterkere afname (slope_{Light} = -0,60; slope_{Standaard} = -0,53) dan leerlingen in de Controle conditie (slope = -0,41), met name in het eerste jaar. Er waren geen interventie-effecten op veerkracht noch voor leerlingen in de Plus conditie.

[Figuur 3]

Wanneer we de leerlingen in de Light conditie vergeleken met de leerlingen in de Standaard conditie op verandering in internaliserend gedrag, paste het vrij geschatte model significant beter op de data, $\Delta\chi^2_{SB}(1) = 6,60$; $p = ,010$, dan het model met gelijkgestelde slopes. Internaliserend gedrag nam dus sterker af bij leerlingen in de Light conditie (slope = -

0,60) dan bij leerlingen in de Standaard conditie (slope = -0,53). In het intrapersoonlijk domein toonden leerlingen in de Light conditie over het algemeen dus de positiefste veranderingen over tijd, waarbij de interventie de grootste impact had aan het begin van het programma.

Interpersoonlijk domein. Leerlingen in de Light conditie hadden tevens een voordeliger veranderingstraject voor agressie vergeleken met de Controle conditie ($d = 0,31$; zie Figuur 3), vooral in het eerste jaar. Zij bleven over tijd relatief stabiel op agressie (slope = 0,09), terwijl leerlingen in de Controle conditie een toename lieten zien (slope = 0,56). Verder lieten de resultaten een relevant ($d = 0,24$), hoewel niet significant ($p = ,056$), verschil zien tussen veranderingstrajecten in pesten in de Light conditie ten opzichte van de Controle conditie. Leerlingen in de Light conditie vertoonden een minder sterke toename in pesten (slope = 2,34) dan leerlingen in de Controle conditie (slope = 2,80; zie Figuur 3). Het interventie-effect was het sterkste in het eerste jaar. Er zijn geen interventie-effecten gevonden op de andere uitkomsten noch voor leerlingen in de Standaard en Plus condities. Ook in het interpersoonlijk domein hadden leerlingen in de Light conditie over het algemeen betere veranderingstrajecten met, wederom, de sterkste interventie-effecten in het eerste jaar van de interventie.

Discussie

Het doel van de huidige studie was om 1) te onderzoeken of R&W een positief effect heeft op vmbo-leerlingen en 2) om te bepalen of verschillende versies van deze interventie verschillende effecten zouden bewerkstelligen. De interventie bleek gematigd effectief in het stimuleren van interventie-specifieke uitkomsten (zelfcontrole, emotionele zelfregulatie, en een trend voor zelfvertrouwen) en in het bevorderen van sommige aspecten van het intra- en interpersoonlijk domein, maar uitsluitend wanneer een kernteam betrokken was bij de implementatie van de interventie en vooral in het eerste jaar. R&W Light was met name effectief in het stimuleren van competenties en het voorkomen van problemen in het intrapersoonlijk domein (psychologisch welzijn, seksuele autonomie, en internaliserend gedrag). In het interpersoonlijk domein leek R&W Light een buffer te zijn tegen toenames in agressie en pesten. Interventie-effecten waren over het algemeen gematigd (effecten van 0,24-0,38) en vergelijkbaar met effecten van andere universele schoolinterventies met soortgelijke doelen. Zo rapporteerden Durlak en collega's (2011) in hun meta-analyse effectgroottes van 0,23 voor positieve houdingen naar jezelf en anderen, 0,24 voor positief sociaal gedrag, 0,22 voor gedragsproblemen, en 0,24 voor emotionele problemen. Mertens en collega's (2020)

vonden in hun meta-analyse effectgroottes van 0,19 en 0,15 op respectievelijk het intrapersoonlijk en interpersoonlijk domein.

Met betrekking tot de interventie-specifieke uitkomsten vonden we op drie van de vier basiscompetenties, die van belang zijn volgens de R&W theorie, positieve interventie-effecten (zelfcontrole, emotionele zelfregulatie, en een trend voor zelfvertrouwen) wat de potentie van de interventie aangeeft. Echter, er is geen effect gevonden op de mate van zelfreflectie van leerlingen. Ook eerdere interventies zijn ineffectief gebleken in het verbeteren van zelfreflectie bij vmbo-leerlingen. Ter Vrugte en collega's (2015), bijvoorbeeld, onderzochten de effectiviteit van een spel om de rekenvaardigheid van vmbo-leerlingen te verbeteren. Dit spel bevatte een reflectiestimulerend component, maar ze vonden hier geen effect op. Mogelijk is de vroege adolescentie niet de beste ontwikkelingsfase om zelfreflectie te meten omdat de metacognitieve vaardigheden, die adolescenten in staat stellen om betekenisvol te reflecteren op gedachten, gevoelens en gedrag (Sauter et al., 2010), nog volop in ontwikkeling zijn. Desalniettemin, aangezien zelfreflectie een potentieel effectief component in schoolinterventies op het voortgezet onderwijs is (Mertens et al., 2020), zou toekomstig onderzoek moeten nagaan wanneer en hoe zelfreflectie het beste gestimuleerd kan worden bij vmbo-leerlingen.

Wat betreft de overkoepelende uitkomsten vonden we de meeste effecten in het intrapersoonlijk domein. Slechts één interventie-effect (agressie en een trend voor pesten) werd gevonden in het interpersoonlijk domein. Mogelijk komt dit doordat in het intrapersoonlijk domein een meer praktische aanpak gebruikt wordt, terwijl een meer verbale aanpak gebruikt wordt bij het interpersoonlijk domein. In de handleiding van de interventie (Ykema, 2002, 2018) wordt de nadruk in de oefeningen en spellen gelegd op gevoelens, emoties, en attitudes van leerlingen (dus het intrapersoonlijk domein; "Was je in balans tijdens het spel?"). De attitudes en gedragingen van leerlingen in relatie met anderen (dus het interpersoonlijk domein) worden voornamelijk aangekaart tijdens (verbale) rolspellen en discussies. Zo bespreken leerlingen bijvoorbeeld wat pesten is en wat leerlingen er tegen kunnen doen (bijv. "Hoe voelde de pester zich toen zij samen 'stop' zeiden?"). Deze meer verbale aanpak sluit minder goed aan bij een praktische leer methode waardoor vmbo-leerlingen mogelijk niet optimaal profiteren van de interventie op het interpersoonlijk domein.

De interventie-effecten waren het sterkste tijdens het eerste gedeelte van de interventie en namen af tot onbeduidende effecten in het tweede jaar; mogelijk is het dus voldoende om alleen het eerste jaar van de interventie te implementeren. Een afname in effecten in relatief lange interventies is eerder gevonden in meta-analyses (Bakermans-Kranenburg et al., 2003;

Cuijpers, 2002; De Mooij et al., 2020). Deze meta-analyses suggereren dat korte-termijn interventies met een beperkte hoeveelheid sessies de voorkeur hebben. Onderzoek heeft ook aangetoond dat participanten die profiteren van een interventie vaak vroeg in de interventie verbetering laten zien, ongeacht de tijdsspanne van de interventie (Lutz et al., 2014; Tadić et al., 2010). Deze afname in interventie-effecten zou mogelijk gerelateerd kunnen zijn aan een afname in de motivatie van leerlingen (Philips & Wennberg, 2014). Toekomstig onderzoek zou kunnen onderzoeken wat de effectiviteit van R&W is indien het tweede jaar van de interventie niet gegeven wordt, of wanneer alleen enkele boostersessies gegeven worden in dat jaar. Alleen het eerste jaar van de interventie implementeren zou wellicht de belasting voor leerlingen kunnen verlichten zonder de effectiviteit van de interventie in gevaar te brengen.

Verder lieten de resultaten zien dat de Light conditie (kernteam) het meest effectief was, wat suggereert dat voor sommige interventies geldt “*less is more*”. Dit is in lijn met de resultaten van de meta-analyse van Durlak en collega’s (2011) die aantoonde dat het positieve effect van het betrekken van meerdere systemen in interventies niet altijd gevonden wordt. In een interventie met een smalle ecologische focus (alleen kernteam) zijn slechts een paar personen betrokken; zij zijn dus als enige verantwoordelijk voor het correct en effectief toepassen van de interventie, wat de implementatie zou kunnen versterken (Durlak, 2016). In een interventie met een brede ecologische focus (bijv. toevoegen van een leerkracht- en/of oudercomponent) zijn veel mensen betrokken die de verantwoordelijkheid voor implementatie kunnen verdelen. Deze gedeelde verantwoordelijkheid zou mogelijk een “bijstandereffect” kunnen uitlokken, verminderde verantwoordelijkheidsgevoelens van individuen wanneer meerdere mensen aanwezig zijn (Fischer et al., 2011). Toekomstig onderzoek zou zich moeten richten op de mogelijk invloed van bijstandereffecten wanneer de verantwoordelijkheid voor implementatie verdeeld is onder een relatief grote groep mensen.

De resultaten van de huidige studie moeten geïnterpreteerd worden in het licht van zijn sterke en zwakke punten. Sterke punten van de studie zijn de meerdere meetmomenten, de drie condities die verschilden in de ecologische focus, en de grote steekproef. Een zwak punt is het gebrek aan informatie over de inhoud van de betrokkenheid van het schoolteam en de oudercomponent. Hoewel we weten dat deze schoolteam- en oudercomponenten geïmplementeerd zijn, weten we relatief weinig over *hoe* interventietechnieken toegepast zijn door schoolteam en ouders. Ten tweede, de R&W trainers gaven de interventie voor de eerste keer, na drie trainingdagen. Meer ervaren trainers zouden mogelijk meer verandering kunnen bewerkstelligen. Ten derde, de uitkomsten zijn enkel gemeten door middel van zelfrapportage

van de leerlingen. Daarnaast waren er aan het begin van de studie kleine verschillen tussen de condities op uitkomstmaten wat effect gehad zou kunnen hebben op de resultaten, ondanks dat we hier in de analyses voor gecontroleerd hebben, en hebben we SES niet onderzocht. Ten vierde, we hebben effecten meteen na de interventie onderzocht. Toekomstig onderzoek zou ook lange-termijn effecten van de interventie moeten onderzoeken. Tot slot, in deze studie is R&W onderzocht onder een specifieke doelgroep, namelijk vmbo-leerlingen, en bij een beperkt aantal scholen verspreid over Nederland. Hoewel het belangrijk is om inzicht te krijgen in welke interventies effectief zijn voor deze doelgroep, kunnen de bevindingen niet zonder verder onderzoek gegeneraliseerd worden naar andere groepen (leerlingen) en alle vmbo-leerlingen van Nederland. Daarnaast zou het interessant zijn voor vervolgonderzoek om meer inzicht te krijgen in verschillen in effectiviteit voor bepaalde subgroepen (bijv. jongens versus meisjes, jongeren met verschillen in persoonlijkheid: Mertens et al., 2022a). Tenslotte, het dient opgemerkt te worden dat de theorie achter R&W slechts gedeeltelijk onderzocht is. Bijvoorbeeld, Mertens et al. (2022b) hebben laten zien dat toenames in zelfregulatie en zelfvertrouwen inderdaad gerelateerd zijn aan toenames in veerkracht en psychologisch welzijn, zoals de R&W theorie veronderstelt. Er is echter meer onderzoek nodig om de overige aspecten van deze theorie empirisch te onderbouwen.

Conclusie

Met name vmbo-leerlingen lijken behoefte te hebben aan een effectieve interventie die hun competenties kan stimuleren en de ontwikkeling van problemen kan voorkomen in zowel het intra- als het interpersoonlijk domein (Stevens & De Looze, 2018). R&W kan een positieve invloed hebben op vmbo-leerlingen, met name in het verbeteren van het intrapersonlijk domein en met name in het eerste jaar van de interventie. In het interpersoonlijk domein fungeert de interventie mogelijk als een buffer tegen toenames in agressie en pesten. Verder is gebleken dat effecten niet noodzakelijkerwijs toenemen bij een langere interventie of door meer mensen erbij te betrekken (dus door een bredere ecologische implementatiestrategie).

Onze bevindingen hebben belangrijke implicaties voor onderzoek en de praktijk. Zo laat onze studie zien dat verschillende implementatiestrategieën van dezelfde interventie effect kan hebben op de effectiviteit ervan. Dit onderstreept het belang om bij de evaluatie van een interventie niet alleen het programma zelf maar ook de implementatiestrategie te onderzoeken. Voor de praktijk laat onze studie zien dat interventies met een actieve en praktijkgerichte leerstijl positieve effecten kunnen hebben op vmbo-leerlingen, een kortere interventie periode mogelijk voldoende is om verandering te bewerkstelligen en het betrekken

van meer leerkrachten en ouders niet altijd een positief effect heeft. Voor scholen zou het daarom waardevol kunnen zijn om te investeren in interventies met een leerstijl die aansluit bij die van de doelgroep, een goede kwaliteit training voor een kernteam van het schoolteam en een hoge kwaliteit van implementatie met een smalle ecologische focus.

Trial Registratie

Nederlands Trial Register nummer NL6371 (oud nummer: NTR 6554), geregistreerd op 3 juli 2017.

Verklaring van Concurrerende Belangen

Geen.

Dankbetuigingen

Deze studie is gefinancierd door een beurs van ZonMw, nummer 531001106

Auteurschap (CRediT)

EM: Conceptualisatie, methodologie, formele analyses, dataverzameling, schrijven – conceptversie, schrijven – revisies, projectmanagement; **MD:** Conceptualisatie, methodologie, schrijven – revisies, supervisie; **MvL:** Conceptualisatie, methodologie, schrijven – revisies, supervisie; **ER:** Conceptualisatie, methodologie, schrijven – revisies, supervisie, subsidie verworven.

Korte Verklaring over de Data en Analyses

De analyse codes en output zijn vrij toegankelijk via Open Science Framework ([GEMASKEERD VOOR REVIEW]). De data zijn wegens gevoeligheid niet vrij toegankelijk, maar op aanvraag beschikbaar bij de eerste auteur.

Literatuur

- Anderson, L. M., Reilly, E. E., Gorrell, S., Schaumberg, K., & Anderson, D. A. (2016). Gender-based differential item function for the difficulties in emotion regulation scale. *Personality and Individual Differences, 92*, 87–91.
- Bakermans-Kranenburg, M. J., IJzendoorn, M. H. van, & Juffer, F. (2003). Less is more: Meta-analyses of sensitivity and attachment interventions in early childhood. *Psychological Bulletin, 129*, 195–215.
- Bishop, D. C., Pankratz, M. M., Hansen, W. B., Albritton, J., Albritton, L., & Strack, J. (2014). Measuring fidelity and adaptation: Reliability of an instrument for school-based prevention programs. *Issues and Solutions in Risky Behavior Prevention Research, 37*, 231-257.
- Boor-Klip, H. J., Segers, E., Hendrickx, M. M. H. G., & Cillessen, A. H. N. (2016). Development and psychometric properties of the classroom peer context questionnaire. *Social Development, 25*, 370–389.
- Bronfenbrenner, U. (1979). *The ecology of human development*. Harvard University Press.
- Bruijn, D. de, Overmaat, M., Glaudé, M., Heemskerk, I., Leeman, Y., Roeleveld, J., & Venne, L. van de (2005). Krachtige leeromgevingen in het middelbaar beroepsonderwijs: Vormgeving en effecten. *Pedagogische Studiën, 82*, 77–95.
- Chorpita, B. F., Reise, S., Weisz, J. R., Grubbs, K., Becker, K. D., Krull, J.L., & The Research Network on Youth Mental Health (2010). Evaluation of the brief problem checklist: Child and caregiver interviews to measure clinical progress. *Journal of Consulting and Clinical Psychology, 78*, 526–536.
- Cuijpers, P. (2002). Effective ingredients of school-based drug prevention programs. A systematic review. *Addictive Behaviors, 27*, 1009–1023.
- Davidson, J. R. T., & Connor, K. M. (2017). *Connor-Davidson Resilience Scale (CD-RISC) manual*. Unpublished and partly accessible at www.cd-risc.com.
- Dodge, K. A., & Coie, J. D. (1987). Social-information-processing factors in reactive and proactive aggression in children's peer groups. *Journal of Personality and Social Psychology, 53*, 1146–1158.
- Durlak, J. A. (2016). Programme implementation in social and emotional learning: Basic issues and research findings. *Cambridge Journal of Education, 4*(3), 333–345.
<https://doi.org/10.1080/030574X.2016.1142504>

- Durlak, J. A., Weissberg, R. P., Dymnicki, A. B., Taylor, R. D., & Schellinger, K. B. (2011). The impact of enhancing students' social and emotional learning: A meta-analysis of school-based universal interventions. *Child Development, 82*, 405–432.
- Feingold, A. (2013). A regression framework for effect size assessments in longitudinal modeling of group differences. *Review of General Psychology, 17*, 111–121.
- Finkenauer, C., Engels, R. C. M. E., Baumeister, R. F. (2005). Parenting behaviour and adolescent behavioural and emotional problems: The role of self-control. *International Journal of Behavioral Development, 29*, 58–69.
- Fischer, P., Krueger, J. I., Greitemeyer, T., Vogrincic, C., Kastenmüller, A., Frey, D., Heene, M., Wicher, M., & Kainbacher, M. (2011). The bystander-effect: A meta-analytic review on bystander intervention in dangerous and non-dangerous emergencies. *Psychological Bulletin, 137*, 517–537.
- Graaf, H. de, Meijer, S., Poelman, J., & Vanwesenbeeck, I. (2005). *Seks onder je 25e*. Uitgeverij Eburon.
- Graaf, I. de, Haas, S. de, Zaagsma, M., & Wijssen, C. (2016). Effects of rock and water: An intervention to prevent sexual aggression. *Journal of Sexual Aggression, 22*, 4–19.
- Harter, S. (1988). *Manual for the Self-Perception Profile for Adolescents*. University of Denver.
- Kuosmanen, T., Clarke, A., & Barry, M. M. (2019). Promoting adolescents' mental health and wellbeing: Evidence synthesis. *Journal of Public Mental Health, 18*(1), 73–83. <https://doi.org/10.1108/JPMH-07-2018-0036>
- Lutz, W., Hofmann, S. G., Rubel, J., Boswell, J. F., Shear, M. K., Gorman, J. M., Woods, S. W., & Barlow, D. H. (2014). Patterns of early change and their relationship to outcome and early treatment termination in patients with panic disorder. *Journal of Consulting and Clinical Psychology, 82*, 287–297.
- Mahoney, J. L., Weissberg, R. P., Greenberg, M. T., Dusenbury, L., Jagers, R. J., Niemi, K., Schlinger, M., Schlund, J., Shriver, T. P., VanAusdal, K., & Yoder, N. (2021). Systemic social and emotional learning: Promoting educational success for all preschool to high school students. *American Psychologist, 76*(7), 1128–1142. <https://doi.org/10.1037/amp0000701>
- Mertens, E. C. A., Deković, M., Leijten, P., Londen, M. van, & Reitz, E. (2020). Components of school-based interventions stimulating students' intrapersonal and interpersonal domains: A meta-analysis. *Clinical Child and Family Psychology Review, 23*, 605–631. <https://doi.org/10.1007/s10567-020-00328-y>

- Mertens, E. C. A., Deković, M., Londen, M. van, & Reitz, E. (2022a). Personality as a moderator of intervention effects: Examining differential susceptibility. *Personality and Individual Differences, 186 Part A*. Article number 111323.
<https://doi.org/10.1016/j.paid.2021.111323>
- Mertens, E. C. A., Deković, M., Londen, M. van, & Reitz, E. (2022b). Parallel changes in positive youth development and self-awareness: The role of emotional self-regulation, self-esteem, and self-reflection. *Prevention Science, 23*, 502-512.
<https://doi.org/10.1007/s11121-022-01345-9>
- Modecki, K. L., Zimmer-Gembeck, M. J., & Guerrà, N. (2017). Emotion regulation, coping, and decision making: Three linked skills for preventing externalizing problems in adolescence. *Child Development, 88*, 417-426.
- Mooij, B. de, Fekkes, M., Scholte, R. H. J., & Overbeek, G. (2020). Effective components of social skills training programs for children and adolescents in nonclinical samples: A multilevel meta-analysis. *Clinical Child and Family Psychology Review, 23*, 250-264.
- Muthén, L. K., & Muthén, B. O. (1998-2010). *Mplus user's guide* (6th ed.). Muthén & Muthén.
- Park, D., Tsukayama, E., Goodwin, G., Patrick, S., & Duckworth, A. L. (2017). A tripartite taxonomy of character: Evidence for intrapersonal, interpersonal, and intellectual competencies in children. *Contemporary Educational Psychology, 48*, 16-27.
- Pellegrino, J. W., & Hilton, M. L. (2012). *Education for life and work: Developing transferable knowledge and skills in the 21st century*. National Academy of Sciences.
- Philips, B., & Wennberg, P. (2014). The importance of therapy motivation for patients with substance use disorders. *Psychotherapy, 51*, 555-562.
- Ravens-Sieberer, U., & The European KIDSCREEN Group (2006). *The KIDSCREEN questionnaires. Quality of life questionnaires for children and adolescents – Handbook*. Pabst Science Publisher.
- Satorra, A., & Bentler, P. M. (2010). Ensuring positiveness of the scaled difference chi-square test statistic. *Psychometrika, 75*, 243-248.
- Sauter, F. M., Heyne, D., Blöte, A. W., Widenfelt, B. M. van, & Westenberg, P. M. (2010). Assessing therapy-relevant cognitive capacities in young people: Development and psychometric evaluation of the self-reflection and insight scale for youth. *Behavioral and Cognitive Psychotherapy, 38*, 303-317.
- Smit, K., Brabander, C. J. de, & Martens, R. L. (2014). Student-centred and teacher-centred learning environment in pre-vocational secondary education: Psychological needs, and

- motivation. *Scandinavian Journal of Educational Research*, 58(6), 695–712.
<https://doi.org/10.1080/00313831.2013.821090>
- Solberg, M. E., & Olweus, D. (2003). Prevalence estimation of school bullying with the Olweus bully/victim questionnaire. *Aggressive Behavior*, 29, 239–268.
- Stevens, G., & Looze, M. de (2018). Welbevinden en psychische problemen. In M. Boer & G. Stevens (Eds.), *HBSC 2017: Gezondheid, welzijn en opvoeding van jongeren in Nederland* (pp. 63–75). Ridderprint BV.
- Tadić, A., Helmreich, I., Mergl, R., Hautzinger, M., Kohnen, R., Henkel, V., & Hegerl, U. (2010). Early improvement is a predictor of treatment outcome in patients with mild major, minor or subsyndromal depression. *Journal of Affective Disorders*, 120, 86–93.
- Timmermans, A., Naaijer, H., Keuning, J., & Zijssling, D. (2017). *Cohortonderzoek COOL5-18 basisrapport meting VO-3 in 2014*. GION (onderwijs/onderzoek).
- Vrugte, J. ter, Jong, T. de, Wouters, P., Vandercruysse, S., Elen, J., & Oostendorp, H. van (2015). When a game supports prevocational math education but integrated reflection does not. *Journal of Computer Assisted Learning*, 31, 462–480.
- Ykema, F. (2002). *Het Rots en Water perspectief. Een psychofysische training voor jongens. Basisboek*. Uitgeverij SWP.
- Ykema, F. (2018). *Rots en Water praktijkboek. Psychofysische training voor jongens én meisjes*. Uitgeverij SWP.
- Ykema, F., Hartman, D., & Imms, W. (2006). *Bringing it together. Includes 22 case studies of Rock & Water in practice in various settings*. Callaghan: University of Newcastle, Family action Centre and Gadaku institute.

Vierde reactie reviewer 1– 15-8-2023

Juli 2023

Graag wil ik de auteurs bedanken voor hun antwoorden en verduidelijkingen. De auteurs hebben belangrijke informatie toegevoegd aan het artikel. Er staat ook informatie in de rebuttal voor de reviewers die niet opgenomen is in het artikel, maar naar mijn mening ook voor de lezer relevant is. Daarom nog enkele voorstellen voor korte toevoegingen:

-Zou er kort vermeld kunnen worden in het artikel dat R&W aanvankelijk voor jongens ontwikkeld is, maar doorontwikkeld is voor jongens en meisjes en nu al jaren ingezet wordt bij zowel jongens als meisjes.

-Ik stel voor dat jullie in de discussie nog een opmerking toevoegen dat vervolgonderzoek aandacht kan besteden aan de vraag hoe de interventie-specifieke theorie van R&W zich verhoudt tot algemeen aanvaarde theorieën over de psychosociale ontwikkeling van adolescenten.

-Kunnen jullie in het artikel vermelden met welke interventies (of voorbeelden van) jullie R&W vergelijken qua doelstellingen en methodiek? Met andere woorden: kunnen jullie “andere schoolinterventies” specificeren op p.6? (“R&W onderscheidt zich met name van *andere schoolinterventies* door een actieve lesmethode waarin de nadruk ligt op activerende werkvormen in plaats van passieve verbale uitleg;” p. 6). Zo kan de lezer de vergelijking ook zelf maken.

-Wat betreft de analyses: Het was geenszins mijn bedoeling om te suggereren dat de analyses anders zouden moeten of dat er effecten op schoolniveau onderzocht moeten worden. Zouden jullie de ICCs op klas- en schoolniveau die jullie rapporteren in de rebuttal ook willen rapporteren in het artikel? Dat zou jullie beslissing om 1-level modellen te onderzoeken verder ondersteunen.

-Power: Ik kan de pre-registratie en het protocol niet raadplegen vanwege anonimisering (waarvoor uiteraard alle begrip). Ik stel voor dat jullie kort het benodigde aantal clusters benoemen in voetnoot 1: Uit de power analyse in het protocol paper (zie GEMASKEERD VOOR REVIEW) blijkt een minimale steekproefgrootte van 720 participanten en x clusters (scholen) nodig voor deze analyses.

-Op mijn opmerking over kenmerken van de steekproef op schoolniveau (bv. leerlingenpopulatiekenmerken, schoolgrootte,...) wordt niet gereageerd. Ik veronderstel dat jullie goede argumenten hebben om geen schoolkenmerken te rapporteren. Daarom zou ik willen voorstellen dat jullie in de discussie kort toevoegen dat vervolgonderzoek kan nagaan of de interventie effecten ook beïnvloed worden door schoolkenmerken. Op die manier worden schoolkenmerken toch onder de aandacht van de lezer gebracht.

Vierde reactiebrief van auteur – 28-8-2023

Utrecht, 28 Augustus 2023

Onderwerp: Vierde revisie manuscript

Geachte Prof. dr. Boendermaker en meneer Veenstra,

Bijgevoegd vindt u de vierde revisie van ons manuscript: “Stevig als een rots, flexibel als water? Effectiviteit van een schoolinterventie gericht op het intrapersonlijk en interpersoonlijk domein”.

Wij hebben de laatste punten van feedback van Reviewer A verwerkt. Om het reviewproces te faciliteren hebben wij onze reacties op de reviewer toegevoegd in de Appendix hieronder en aanpassingen in het manuscript geel gemarkeerd.

Wij zijn verheugd te horen dat deze revisie geaccepteerd gaat worden voor publicatie en bedanken u voor uw tijd.

Met vriendelijke groet,

Namens mijn co-auteurs Deković, Van Londen, en Reitz,

Esther Mertens

e.c.a.mertens@fsw.leidenuniv.nl

Appendix - Reviewer A

Graag wil ik de auteurs bedanken voor hun antwoorden en verduidelijkingen. De auteurs hebben belangrijke informatie toegevoegd aan het artikel. Er staat ook informatie in de rebuttal voor de reviewers die niet opgenomen is in het artikel, maar naar mijn mening ook voor de lezer relevant is. Daarom nog enkele voorstellen voor korte toevoegingen:

-Zou er kort vermeld kunnen worden in het artikel dat R&W aanvankelijk voor jongens ontwikkeld is, maar doorontwikkeld is voor jongens en meisjes en nu al jaren ingezet wordt bij zowel jongens als meisjes.

We hebben dit in de inleiding toegevoegd.

“Oorspronkelijk is deze interventie ontwikkeld om ingezet te worden bij jongens, maar wordt na doorontwikkeling inmiddels al jarenlang ingezet bij zowel jongens als meisjes.” p. 6

-Ik stel voor dat jullie in de discussie nog een opmerking toevoegen dat vervolgonderzoek aandacht kan besteden aan de vraag hoe de interventie-specifieke theorie van R&W zich verhoudt tot algemeen aanvaarde theorieën over de psychosociale ontwikkeling van adolescenten.

We hebben dit in de discussie toegevoegd.

“Er is echter meer onderzoek nodig om de overige aspecten van deze theorie empirisch te onderbouwen en na te gaan hoe de R&W theorie zich verhoudt tot andere theorieën over de psychosociale ontwikkeling van adolescenten.” p. 24

-Kunnen jullie in het artikel vermelden met welke interventies (of voorbeelden van) jullie R&W vergelijken qua doelstellingen en methodiek? Met andere woorden: kunnen jullie “andere schoolinterventies” specificeren op p.6? (“R&W onderscheidt zich met name van andere schoolinterventies door een actieve lesmethode waarin de nadruk ligt op activerende werkvormen in plaats van passieve verbale uitleg;” p. 6). Zo kan de lezer de vergelijking ook zelf maken.

We hebben voorbeelden van andere interventies toegevoegd in de introductie.

“R&W onderscheidt zich van andere schoolinterventies (zoals Kanjertraining, Levensvaardigheden, Happyles) door een meer actieve lesmethode te gebruiken waarin de nadruk ligt op activerende werkvormen in plaats van passieve verbale uitleg;” p. 6

-Wat betreft de analyses: Het was geenszins mijn bedoeling om te suggereren dat de analyses anders zouden moeten of dat er effecten op schoolniveau onderzocht moeten worden. Zouden jullie de ICCs op klas- en schoolniveau die jullie rapporteren in de rebuttal ook willen rapporteren in het artikel? Dat zou jullie beslissing om 1-level modellen te onderzoeken verder ondersteunen.

In ons analyseplan beschrijven wij dat we 1-level-modellen analyseren waarbij we corrigeren voor clustering van participanten op schoolniveau middels de complex sample cluster. Ondanks dat de ICCs op klas- en schoolniveau relatief laag was, nemen we dit (deels) wel mee in onze analyses. Naar ons idee zal het toevoegen van de ICCs op klas- en schoolniveau met name de illusie wekken dat wij de clustering van onze data volledig buiten beschouwing hebben gelaten. Om verwarring over onze analyses te voorkomen, hebben wij besloten de ICCs niet expliciet te vermelden in het manuscript (geïnteresseerden kunnen de exacte ICCs

opzoeken via [Open Science Framework](#) waar onze analyse codes en output vrij toegankelijk zijn).

-Power: Ik kan de pre-registratie en het protocol niet raadplegen vanwege anonimisering (waarvoor uiteraard alle begrip). Ik stel voor dat jullie kort het benodigde aantal clusters benoemen in voetnoot 1: Uit de power analyse in het protocol paper (zie GEMASKEERD VOOR REVIEW) blijkt een minimale steekproefgrootte van 720 participanten en x clusters (scholen) nodig voor deze analyses.

Omdat onze analyses uitgevoerd zijn op het niveau van leerlingen, is ook onze poweranalyses uitgevoerd op het niveau van leerlingen in plaats van op schoolniveau. Het is daarom niet mogelijk om de gevraagde informatie toe te voegen.

-Op mijn opmerking over kenmerken van de steekproef op schoolniveau (bv. leerlingenpopulatiekenmerken, schoolgrootte,...) wordt niet gereageerd. Ik veronderstel dat jullie goede argumenten hebben om geen schoolkenmerken te rapporteren. Daarom zou ik willen voorstellen dat jullie in de discussie kort toevoegen dat vervolgonderzoek kan nagaan of de interventie effecten ook beïnvloed worden door schoolkenmerken. Op die manier worden schoolkenmerken toch onder de aandacht van de lezer gebracht.

Wij hebben dit toegevoegd in de discussie.

“Daarnaast zou het interessant zijn voor vervolgonderzoek om meer inzicht te krijgen in verschillen in effectiviteit voor bepaalde subgroepen (bijv. jongens versus meisjes, jongeren met verschillen in persoonlijkheid: Mertens et al., 2022a) of schoolkenmerken (bijv. locatie en schoolgrootte).” p. 24

Vijfde versie van het manuscript – 28-8-2023

Sterk als een rots, flexibel als water? Effectiviteit van een schoolinterventie gericht op het intrapersoonlijk en interpersoonlijk domein

Samenvatting

De huidige studie had als doel om te onderzoeken of 1) de schoolinterventie Rots en Water (R&W) effectief was in het verbeteren van competenties en voorkomen van problemen in het intra- en interpersoonlijk domein van leerlingen op het voorbereidend middelbaar beroepsonderwijs (vmbo), en of 2) interventie-effecten veranderden naarmate er meer mensen betrokken waren bij de interventie (kernteam van leerkrachten, het hele schoolteam, het hele schoolteam en ouders). We hebben een gerandomiseerde gecontroleerde studie uitgevoerd onder brugklassers ($N = 1299$; $M_{leeftijd} = 12,38$; 54% jongens) die vier keer digitale vragenlijsten hebben ingevuld. De resultaten lieten zien dat R&W een aantal interventie-specifieke uitkomsten (zelfcontrole en emotionele zelfregulatie), uitkomsten in het intrapersoonlijk domein (psychologisch welzijn, seksuele autonomie, en internaliserend gedrag) en in het interpersoonlijk domein (agressie) verbeterde, maar uitsluitend wanneer een kernteam van leerkrachten betrokken was bij de implementatie van de interventie. De (gematigde) interventie-effecten waren het grootste in het eerste jaar van de tweejarige interventie. Deze resultaten laten zien dat R&W positieve effecten kan bewerkstelligen bij vmbo-leerlingen wanneer deze geïmplementeerd wordt door een kernteam. Het lijkt dus waardevol voor scholen om te investeren in interventies met een leerstijl die aansluit bij die van de doelgroep, met een hoge kwaliteit van implementatie en een beperkt aantal betrokken personen.

Abstract

The aim of the present study was to examine 1) the effectiveness of the school-based intervention Rock and Water (R&W) in stimulating competencies and preventing problems in the intra- and interpersonal domains of prevocational students, and 2) whether intervention effects changed as more people were involved in the intervention (core team of teachers, whole school, and whole school and parents). We have conducted a Randomized Controlled Trial among 7th Grade students ($N = 1299$, $M_{age} = 12.38$, 54% boys) who completed four digital questionnaires. The results showed that R&W improved some intervention specific outcomes (self-control and emotional self-regulation), outcomes in the intrapersonal domain (psychological wellbeing, sexual autonomy, and internalizing behavior) and in the interpersonal domain (aggression), but only when solely a core team of teachers was involved in the implementation of the intervention. The (moderate) intervention effects were strongest in the first year of the two-year intervention. These results show that R&W can establish positive effects among prevocational students when implemented by a core team. Thus, it appears worthwhile for schools to invest in interventions that fit with the learning style of the

target group, with a high level of implementation quality and involving a limited number of people.

Trefwoorden: Schoolinterventie; Rots en Water; Intrapersoonlijk domein; Interpersoonlijk domein; Vmbo-leerlingen

Bijsluiter voor de praktijk

1. Vmbo-leerlingen rapporteren meer problemen in het intra- en interpersoonlijk domein dan havo- en vwo-leerlingen.
2. Wij onderzochten: 1) de effectiviteit van de schoolinterventie Rots en Water voor het stimuleren van het intra- en interpersoonlijk domein bij vmbo-leerlingen en 2) of de effectiviteit veranderde naarmate verschillende mensen betrokken waren bij de interventie (kernteam leerkrachten, hele schoolteam, hele schoolteam en ouders).
3. De interventie liet gematigde positieve effecten zien, met name op uitkomsten in het intrapersoonlijk domein, wanneer alleen een kernteam van leerkrachten betrokken was bij de uitvoering van de interventie.
4. Interventie-effecten waren het grootste in het eerste jaar van de interventie.
5. Het lijkt waardevol om te investeren in goede training voor en kwalitatief hoogstaande implementatie door een kernteam.

Aantal woorden:

5977 (hoofdtekst)

Introductie

Scholen spelen, naast andere factoren zoals gezin, leeftijdsgenoten en maatschappelijke omstandigheden, een belangrijke rol in het bevorderen van de ontwikkeling van leerlingen in het intrapersoonlijk en interpersoonlijk domein (Pellegrino & Hilton, 2012). Het *intrapersoonlijk domein* refereert naar subjectief persoonlijk functioneren (bijv. psychologisch welzijn) en is gerelateerd aan schoolresultaten en zelfregulatie. Het *interpersoonlijk domein*, daarentegen, refereert naar sociaal functioneren (bijv. agressie) en is onder andere gerelateerd aan relaties met leeftijdsgenoten en conflict oplossend vermogen (Park et al., 2017). Leerlingen kunnen in deze domeinen competenties ontwikkelen door relevante cognitieve, affectieve en sociale vaardigheden te leren (bijv. identificeren van emoties en perspectief wisselen). Wanneer het verwerven van (sommige) vaardigheden achterwege blijft, kunnen leerlingen problemen ontwikkelen (Modecki et al., 2017). Hoewel de twee domeinen elkaar beïnvloeden, worden ze als apart beschouwd. Scholen zouden bewust kunnen proberen de competenties van leerlingen op beide domeinen te stimuleren om mogelijke problemen te voorkomen (Pellegrino & Hilton, 2012).

Met name leerlingen op het voorbereidend middelbaar beroepsonderwijs (vmbo) zouden baat kunnen hebben bij stimulering van competenties en het voorkomen van problemen. Zij rapporteren, vergeleken met leerlingen op het hoger algemeen voortgezet onderwijs (havo) en voorbereidend wetenschappelijk onderwijs (vwo), over het algemeen een lager niveau van welzijn, meer gedragsproblemen, en meer problemen met leeftijdsgenoten (Stevens & De Looze, 2018). Tegelijkertijd kunnen vmbo-leerlingen een uitdagende groep zijn om te betrekken bij interventielessen op school. Over het algemeen laten deze leerlingen minder intrinsieke motivatie voor school en (verbaal) leren zien dan havo- en vwo-leerlingen (Timmermans et al., 2017). Hoewel schoolinterventies positieve effecten laten zien (e.g., Mertens et al., 2020), is het onduidelijk in hoeverre deze ook effectief zijn voor specifiek vmbo-leerlingen; vooral voor het betrekken van deze groep leerlingen bij een schoolinterventie lijkt het namelijk belangrijk een interventie te implementeren die goed aansluit bij hun interesses en leerstijl. Vmbo-leerlingen worden het beste bereikt en blijven het meest gemotiveerd en betrokken met een praktijkgerichte leer methode. Deze leer methode houdt in dat er een leeromgeving gecreëerd wordt met een praktijkgerichte insteek met activerende, stapsgewijze, gestructureerde taken; leerlingen kunnen hun nieuw geleerde vaardigheden toepassen en eigen doelen stellen. Tijdens de les worden korte momenten van instructie of reflectie afgewisseld met het toepassen van nieuwe vaardigheden en leerkrachten functioneren als mentor en rolmodel (De Bruijn et al., 2005; Smit et al., 2014).

Een universele schoolinterventie die aansluit bij deze actieve leerstijl van vmbo-leerlingen is Rots en Water (R&W; Ykema, 2002, 2018; <https://www.rotsenwater.nl>). Oorspronkelijk is deze interventie ontwikkeld om ingezet te worden bij jongens, maar wordt na doorontwikkeling inmiddels al jarenlang ingezet bij zowel jongens als meisjes. R&W onderscheidt zich van andere schoolinterventies (zoals Kanjertraining, Levensvaardigheden, Happyles) door een meer actieve lesmethode te gebruiken waarin de nadruk ligt op activerende werkvormen in plaats van passieve verbale uitleg; tijdens de R&W lessen worden klassikale moment van informatie, instructie of reflectie afgewisseld met actieve oefeningen waarin nieuwe vaardigheden geoefend kunnen worden. De R&W trainer (de leerkracht) ondersteunt leerlingen op een positieve manier en geeft, als rolmodel, tijdens de lessen het goede voorbeeld. R&W is gebaseerd op de theorie van het 'R&W huis' (Ykema, 2002, 2018). Dit huis is gebouwd op vier pilaren: zelfcontrole, zelfreflectie, zelfvertrouwen, en emotionele zelfregulatie. Deze vier competenties worden beschouwd als specifieke uitkomsten van de interventie en vormen de basisvaardigheden voor de ontwikkeling van de bredere, overkoepelende, competenties. Volgens de theorie van R&W, zou het verbeteren van deze vier competenties samengaan met verbeteringen in psychosociaal welzijn, veerkracht, en seksuele autonomie en het verminderen van internaliserend gedrag (het intrapersoonlijk domein). Door het creëren van een veilige en respectvolle omgeving en door leerlingen alternatieve gedragingen aan te bieden zou volgens de theorie van R&W externaliserend gedrag, agressie en pesten verminderen en positieve relaties tussen leerlingen toenemen (het interpersoonlijk domein).

Hoewel R&W in veel landen geïmplementeerd wordt (bijv. Frankrijk, Australië, China), en met name in Nederland vaak toegepast wordt, is er weinig bekend over de effectiviteit. Met name in Nederland is het belangrijk meer inzicht te krijgen in de effectiviteit van R&W, niet alleen omdat deze interventie vaak wordt geïmplementeerd in Nederland, maar ook omdat het ingezet wordt binnen de kaders van de Wet Veiligheid op school – een wet die scholen verplicht zich actief in te spannen om pesten tegen te gaan en een veilig sociaal schoolklimaat te creëren. In Australië hebben professionals in het werkveld verschillende kleinschalige studies uitgevoerd, vaak op basis- en middelbare scholen, die hebben aangetoond dat deelnemers zich na het afronden van R&W veerkrachtiger voelden, een meer positieve identiteit hadden, en meer actieve copingstrategieën toepasten (Ykema et al., 2006). Een onderzoek in Nederland onder (mannelijke) vmbo-leerlingen van De Graaf en collega's (2016) heeft de effectiviteit van R&W op seksuele agressie onderzocht. Na het afronden van R&W verminderden zelfgerapporteerde dwingende strategieën en verbale

manipulatie, terwijl zelfregulatie toenam. Ondanks deze veelbelovende resultaten is een grootschalige studie nodig die zich richt op bredere uitkomstmaten, zowel jongens als meisjes includeert, en in het bijzonder de effectiviteit onderzoekt voor vmbo-leerlingen. Daarom was het eerste doel van de huidige studie om te onderzoeken of R&W effectief is voor vmbo-leerlingen in het verbeteren van competenties en het voorkomen van problemen in het intrapersoonlijk en het interpersoonlijk domein. Verwacht werd dat leerlingen in de R&W conditie een significante verbetering zouden laten zien op alle gemeten interventie-specifieke en overkoepelende (intra- en interpersoonlijke) uitkomsten ten opzichte van leerlingen in de controle conditie.

Het tweede doel was om te onderzoeken of interventie-effecten afhankelijk zijn van de mate waarin verschillende systemen betrokken zijn bij de interventie. Volgens het ecologisch model van Bronfenbrenner (1979), wordt gedrag bepaald door interacties tussen verschillende systemen zoals individu-, familie-, klas-, en schoolsystemen. Het actief betrekken van gerelateerde systemen bij een interventie, zoals de school- en familiesystemen, zou mogelijk zowel generalisatie als behoud van nieuwe vaardigheden kunnen bevorderen (Mahoney et al., 2021); de inhoud van de interventie wordt dan namelijk ingebed in een groot deel van de sociale omgeving van het individu wat de mogelijke invloed ervan op het individu vergroot. Zo concludeerden Mertens en collega's (2020) in hun meta-analyse naar componenten van universele schoolinterventies dat betrokkenheid van de volledige school en/of ouders gerelateerd was aan sterkere interventie-effecten. Echter, Durlak en collega's (2011) vonden in hun meta-analyse van *social emotional learning* (SEL) interventies dat interventies waarbij de hele school en/of ouders betrokken waren niet effectiever waren dan interventies die alleen in de klas gegeven werden. Deze inconsistente bevindingen geven aan dat er meer inzicht nodig is in hoe betrokkenheid van meerdere systemen in een interventie de effectiviteit kan beïnvloeden (Kuosmanen et al., 2019).

De huidige studie onderzocht daarom of er verschil is in de effectiviteit van drie verschillende implementatieversies van R&W: betrokkenheid van 1) een kernteam van leerkrachten, 2) een kernteam en het volledige schoolteam, en 3) een kernteam, het volledige schoolteam en ouders. We hadden geen specifieke hypothese aangezien onderzoek naar de betrokkenheid van meerdere ecologische systemen in interventies inconsistente resultaten laat zien.

Methode

Verschillende versies van R&W

We hebben drie versies van R&W onderzocht die verschilden in het aantal systemen dat betrokken was bij de interventie: 1) *Light*: alleen een kernteam van leerkrachten (die de interventielessen gaven), 2) *Standaard*: een kernteam van leerkrachten en de rest van het schoolteam, en 3) *Plus*: een kernteam, het schoolteam, en ouders. In alle drie de condities werden de interventielessen gegeven door een kernteam van leerkrachten (R&W trainers) die de 3-daagse training succesvol hadden afgerond. Dit kernteam ontving supervisie van een persoonlijke R&W coach. In de Light conditie was alleen een kernteam getraind en op de hoogte van R&W. In de Standaard en de Plus condities had, naast het kernteam van R&W trainers, de rest van het schoolteam een 3-daagse introductie cursus gevolgd over de basisprincipes van R&W, ondersteuning van R&W trainers, en toepassing van de interventietechnieken in reguliere lessen. In de Plus conditie kregen daarnaast ouders de mogelijkheid een documentaire over R&W te bekijken, werden ze uitgenodigd voor een R&W les op school, en ontvingen ze wekelijks e-mails met daarin informatie over de R&W les van die week en werden ze aangemoedigd om met deze informatie thuis aan de slag te gaan.

Design en procedure

De effectiviteit van deze drie verschillende versies van R&W (Light, Standaard en Plus) is onderzocht in scholen met een vmbo-traject in een gerandomiseerde gecontroleerde studie met een actieve controle conditie. In de Controle conditie voerden scholen hun gebruikelijk beleid voor het verbeteren van competenties en het voorkomen van problemen bij leerlingen. Scholen werden geëxcludeerd als ze in de afgelopen twee jaar R&W hadden geïmplementeerd. Dertien scholen verspreid door Nederland werden willekeurig toegewezen aan de condities door gestratificeerde blok randomisatie met blokken van vier. Scholen werden gestratificeerd op schoolgrootte (< 100 leerlingen in de eerste klas, > 100 leerlingen in de eerste klas) om een gelijke verdeling van leerlingen over de condities te bevorderen. Eén school uit de Controle conditie viel uit na randomisatie maar voor het begin van de dataverzameling vanwege een verandering in schoolmanagement. Deze school werd vervangen door een andere geïnteresseerde school (zie Figuur 1 voor de flowchart).

[Figuur 1]

Leerlingen vulden digitale vragenlijsten in, afgenomen door getrainde onderzoeksassistenten, voor de start van de interventie (T1 = oktober/november 2017), na het afronden van het eerste jaar van de R&W lessen (T2 = maart/april 2018), voor de start van de R&W lessen in het tweede jaar van het R&W programma (T3 = oktober 2018), en na de interventie (T4 = januari 2019). Leerlingen gaven actieve, geïnformeerde toestemming voor

het invullen van de vragenlijsten. Ouders gaven passieve, geïnformeerde toestemming voor deelname van hun kind. Deze studie is goedgekeurd door de Ethische Commissie [GEMASKEERD VOOR REVIEW] en geregistreerd [GEMASKEERD VOOR REVIEW] (voor protocol zie [GEMASKEERD VOOR REVIEW]).

Participanten

Op de voormeting bestond de steekproef uit 1299 eerstejaars vmbo-leerlingen ($M_{leeftijd} = 12,38$ jaar; $SD = 0,62$; 54% jongens; 69% met een Westerse achtergrond; zie Tabel 1). Tussen de condities waren geen verschillen in sekse, maar wel in leeftijd en migratieachtergrond. Leerlingen in de Controle conditie waren ouder dan leerlingen in de Light conditie. De Controle en Standaard condities hadden ongeveer een gelijke verdeling tussen leerlingen met een Westerse en een niet-Westerse achtergrond, terwijl in de Light en Plus condities leerlingen vooral een Westerse achtergrond hadden. Daarom werd in alle analyses gecontroleerd voor leeftijd en migratieachtergrond. Een MANOVA liet zien dat er bij de start van de studie kleine, verwaarloosbare, verschillen tussen de condities waren op de uitkomstmaten ($F(39, 3591) = 2.12, p < .001, \eta^2_{\text{partial}} = .023$). De Light conditie scoorde enigszins slechter dan de Controle conditie op zelfcontrole, veerkracht, seksuele autonomie, internaliserende problemen, en externaliserend gedrag. Verder scoorde de Plus conditie iets beter op interpersoonlijke relaties in de klas dan de Light en Controle condities. In alle analyses is gecontroleerd voor de voormeting van de desbetreffende uitkomst, ongeacht of er bij aanvang significante verschillen waren tussen de condities.

Over alle meetmomenten ontbrak 15% van de data. Vergelijkingen tussen leerlingen die uitgevallen zijn ($n_{T2} = 68, n_{T3} = 83, n_{T4} = 60$) en leerlingen die de studie volledig hebben afgerond, lieten geen verschillen zien op demografische kenmerken (leeftijd, sekse verdeling en migratieachtergrond) noch op de uitkomstvariabelen.

Tabel 1

Demografische gegevens van leerlingen op de voormeting per conditie

	Light	Standaard	Plus	Controle	Verschillen op T1		
					F/χ^2	p	$\eta^2_{\text{partial}}/\phi$
N	373	303	249	374			
Leeftijd, M (SD)	12,33 (0,57)	12,38 (0,66)	12,34 (0,60)	12,47 (0,64)	3,89	,009	0,009
Jongens, n (%)	170 (48%)	161 (56%)	131 (55%)	199 (57%)	7,38	,061	0,077

Westerse achtergrond, <i>n</i> (%)	291 (82%)	115 (43%)	211 (91%)	198 (59%)	182,01	< ,001	0,392
------------------------------------	-----------	-----------	-----------	-----------	--------	--------	-------

Scholen bepaalden zelfstandig welke leerkrachten getraind werden in het geven van R&W lessen. De R&W lessen werden in totaal verzorgd door 49 leerkrachten (trainers) tussen de 23 en 64 jaar oud ($M = 36,77$, $SD = 10,02$) met gemiddeld 12,98 jaar aan leerkracht-ervaring. De meeste trainers waren gymdocent ($n = 30$; 71%). Iets meer dan de helft van de trainers was man ($n = 27$; 56%) en bijna alle trainers hadden een Westerse achtergrond ($n = 46$; 96%).

Conditie en Uitvoering

R&W

R&W (Ykema, 2002, 2018) integreert spel en opdrachten om leerlingen te begeleiden in het maken van (fysiek) contact met anderen en het ontdekken, respecteren en stellen van zowel eigen grenzen als grenzen van anderen. Middels oefeningen worden lichaamsbewustzijn (bijv. spanning in de spieren, ademhaling), non-verbale communicatie, en copingstrategieën van leerlingen versterkt. In een oefening in tweetallen moeten leerlingen bijvoorbeeld proberen de andere leerling uit balans te tikken met als doel leerlingen te laten ervaren dat het soms beter is om rigide te zijn en druk van anderen te weerstaan (als een rots) en dat het soms beter is om flexibel te zijn en open te staan voor de mening, gedachten, en gevoelens van anderen (als water).

Naast deze spellen hanteert R&W ook een gedragsmatige aanpak, zoals het gebruik van rolspellen om te oefenen met bijvoorbeeld het stellen en aangeven van grenzen, kalm reageren op provocaties, en het helpen van leerlingen die gepest worden. Tijdens en na de oefeningen geven leerlingen feedback en reflecteren ze. Trainers tonen en bekrachtigen gewenst gedrag, en creëren een veilige omgeving waarin leerlingen fouten mogen maken.

De interventie bestaat uit 22 lessen verspreid over twee schooljaren: 14 lessen in het eerste en 8 lessen in het tweede leerjaar. De lessen werden wekelijks gegeven tijdens gymlessen van 90 minuten.

Controle conditie

In de Controle conditie voerden scholen hun gebruikelijk beleid. Leerlingen hadden bijvoorbeeld een mentor of persoonlijke coach met wie ze regelmatig afspraken hadden, hun welzijn bespraken en naar wie ze toe konden voor advies. Ook was er bijvoorbeeld een

projectweek over ‘anders zijn’, tekenden leerlingen een anti-pestcontract, werd pesten in de klas besproken en was er een ‘anti-pestcoördinator’.

Uitvoering interventie

De uitvoering van R&W is gemeten op twee manieren. Ten eerste, R&W trainers vulden een vragenlijst in na elke derde les in het eerste jaar en elke tweede les in het tweede jaar van de interventie. Deze vragenlijst bestond uit vragen over het verloop van de R&W lessen, zoals “Heeft u in de afgelopen drie/twee lessen van R&W alles kunnen behandelen?” (1 = *alle lessen volledig kunnen geven* tot 4 = *geen les volledig kunnen geven*) en “In hoeverre bent u afgeweken van het R&W handboek?” (1 = *bijna niet* tot 6 = *heel veel*). Ten tweede, drie R&W experts, die alle drie de R&W training voor gevorderden hebben afgerond en meerdere jaren ervaring hebben met het implementeren van R&W, hebben 67 R&W lessen (Light = 28 lessen; Standaard = 17 lessen, Plus = 22 lessen) geobserveerd. Tijdens deze lesobservatie hebben zij een schema ingevuld, gebaseerd op Bishop en collega’s (2014), waarin ze onder andere hun subjectieve beoordeling gaven over de mate waarin de handleiding werd gevolgd, aanpassingen die werden gedaan en de kwaliteit van de les.

Volgens de zelfrapportages hebben R&W trainers over het algemeen de lessen volledig kunnen geven (65%) en zijn ze niet of nauwelijks afgeweken van de handleiding (72%). De R&W experts observeerden dat de meeste lessen volledig of bijna volledig uitgevoerd waren (86%). Trainers weken niet veel af van de handleiding (91%). Eventuele aanpassingen aan de uitvoering werden over het algemeen gezien als verbeteringen (62%). Over het algemeen beoordeelden de R&W experts de lessen van goede (38%) tot zeer goede (54%) kwaliteit. De uitvoering van de R&W lessen verschilde niet tussen condities noch tussen de twee jaar van implementatie, zowel volgens de zelfrapportages als volgens de observaties. Concluderend, het grootste deel van de R&W lessen was daadwerkelijk gegeven en de uitvoering lijkt goed tot zeer goed te zijn geweest.

Om te bepalen of de Standaard en Plus versies geïmplementeerd waren zoals bedoeld, hebben R&W trainers en ouders vragenlijsten ingevuld na het eerste jaar en na het tweede jaar van de interventie. Zoals gepland waren er in de Standaard en Plus condities meer leerkrachten betrokken bij de interventie dan in de Light conditie ($F(2,24) = 5,73; p = ,009$). Ouders in de Plus conditie die de vragenlijst hebben ingevuld ($n = 47$) rapporteerden dat zij de wekelijkse informatie soms (57%) tot vaak (39%) lazen, en sommige ouders gaven aan de documentaire over R&W gekeken te hebben (15%) en deelgenomen te hebben aan een R&W les op school (23%). Verder rapporteerden ouders in de Plus conditie, vergeleken met ouders in de Standaard conditie, thuis vaker over R&W gesproken te hebben ($F(2,75) = 4.37; p =$

,016) en het vaker bij hun eigen kind toegepast te hebben ($F(2,75) = 3,44; p = ,037$). Kortom, ouders in de Plus conditie waren zoals gepland meer betrokken bij de interventie dan ouders in de Light en Standaard condities. Dus, zowel de leerkracht- als de oudercomponent blijkt geïmplementeerd te zijn.

Interventie-Specifieke Uitkomsten

Zelfcontrole

De vaardigheid van leerlingen om hun impulsen te controleren en ongepast gedrag te onderbreken werd gemeten met de korte versie van de *Self-Control Scale* (Finkenauer et al., 2005) die bestaat uit 11 items (bijv. “Ik kan goed weerstand bieden tegen verleidingen”, $\alpha = 0,62-0,72$) beantwoord op een 5-puntsschaal ($1 = \text{helemaal niet}$ tot $5 = \text{helemaal wel}$).

Zelfreflectie

Leerlingen vulden de *Engage* subschaal in van de *Self-Reflection and Insight Scale* (Sauter et al., 2010) om de mate van reflectie op persoonlijke gedachten, gevoelens en gedrag te meten. De subschaal bestaat uit 6 items (bijv. “Ik denk vaak na over hoe ik me voel over iets”) beantwoord op een 6-puntsschaal ($1 = \text{helemaal mee oneens}$ tot $6 = \text{helemaal mee eens}$). De betrouwbaarheid was laag ($\alpha = 0,53-0,60$). Verwijdering van 3 negatief geformuleerde items (“Ik ben bijna nooit bezig met ‘zelfreflectie’”, “Ik denk niet vaak na over mijn gedachten”, en “Ik denk niet echt na over de reden waarom ik me gedraag zoals ik dat doe”) verbeterde de betrouwbaarheid ($\alpha = 0,74-0,89$).

Zelfvertrouwen

Globaal zelfvertrouwen van leerlingen is gemeten met de subschaal *Global self-perception* van de *Self-Perception Profile* (Harter, 1988). Deze subschaal bestaat uit 5 items (bijv. “Ik ben best tevreden met mezelf”) beantwoord op een 4-puntsschaal ($1 = \text{helemaal niet waar}$ tot $4 = \text{helemaal waar}$; $\alpha = 0,73-0,75$).

Emotionele zelfregulatie

Vaardigheden in het controleren van emoties en het gebruik van emotieregulatie strategieën werd gemeten met de *Difficulties in Emotion Regulation Scale* (Anderson et al., 2016). De vragenlijst bestaat uit 14 items (bijv. “Als ik van streek ben, kan ik iets bedenken om ervoor te zorgen dat ik me weer beter voel”) beantwoord op een 5-puntsschaal ($1 = \text{bijna nooit}$ tot $5 = \text{bijna altijd}$; $\alpha = 0,88-0,91$).

Overkoepelende Uitkomsten

Intrapersoonlijk domein

Psychologisch welzijn. Om de aanwezigheid van positieve emoties te meten, hebben leerlingen de subschaal *Psychological Wellbeing* ingevuld van de KIDSCREEN-27 (Ravens-

Sieberer & The European KIDSCREEN Group, 2006). De subschaal bestaat uit 7 items (bijv. “In de afgelopen week, heb je lol gehad?”) beantwoord op een 5-puntsschaal ($1 = \text{nooit}$ tot $5 = \text{altijd}$; $\alpha = 0,76-0,83$).

Veerkracht. De vaardigheid om terug te veren na uitdagingen werd gemeten middels de *Connor-Davidson Resilience Scale-Short Version* (Davidson & Connor, 2017) bestaande uit 10 items (bijv. “Aanpassen aan verandering”) beantwoord op een 5-puntsschaal ($0 = \text{helemaal niet waar}$ tot $4 = \text{bijna altijd waar}$; $\alpha = 0,79-0,92$).

Seksuele autonomie. Coping vaardigheden in seksuele situaties werd gemeten met 5 items van de *Seks onder je 25^{ste} vragenlijst* (bijv. “Wanneer ik ben met iemand die ik heel erg leuk vind, voel ik me helemaal op mijn gemak”; De Graaf et al., 2005) beantwoord op een 4-puntsschaal ($1 = \text{nooit}$ tot $4 = \text{altijd}$). Betrouwbaarheid was laag op T1 ($\alpha = 0,53$) en redelijk op T2, T3, en T4 ($\alpha = 0,62-0,65$).

Internaliserend gedrag. Leerlingen vulden de internaliserende subschaal in van de korte versie van de *Youth Self Report* (Chorpita et al., 2010). De subschaal bestaat uit 6 items (bijv. “Ik maak me vaak zorgen”) beantwoord op een 3-puntsschaal ($0 = \text{nooit}$ tot $2 = \text{vaak}$; $\alpha = 0,79-0,86$).

Interpersoonlijk domein

Interpersoonlijke relaties in de klas. De manier waarop leerlingen relaties in de klas ervaren, werd gemeten met de subschalen Conflict, Cohesie en Comfort van de *Classroom Peer Context Questionnaire* (Boor-Klip et al., 2016). Deze 12 items (bijv. “In deze klas vinden kinderen elkaar aardig”) werden beantwoord op een 5-puntsschaal ($1 = \text{helemaal niet waar}$ tot $5 = \text{helemaal waar}$; $\alpha = 0,80-0,85$).

Externaliserend gedrag. Leerlingen hebben de externaliserende subschaal ingevuld van de korte versie van de *Youth Self Report* (Chorpita et al., 2010). De subschaal bestaat uit 6 items (bijv. “Ik maak veel ruzie”) beantwoord op een 3-puntsschaal ($0 = \text{nooit}$ tot $2 = \text{vaak}$; $\alpha = 0,65-0,79$).

Agressie. Agressie is gemeten met de *Reactive and Proactive Aggression Questionnaire* (Dodge & Coie, 1987). De vragenlijst bestaat uit 6 items (bijv. “Als ze mij plagen of tegen mij liegen, reageer ik kwaad”) beantwoord op een 5-puntsschaal ($1 = \text{nooit}$ tot $5 = \text{bijna altijd}$; $\alpha = 0,65-0,83$).

Pesten en gepest worden. De frequenties van pesten en ervaringen van gepest worden zijn gemeten met 2 items van de *Olweus Bully/Victim Questionnaire* (Solberg & Olweus, 2003): “Hoe vaak heb je de afgelopen twee maanden meegedaan aan het pesten van andere leerlingen op school?” en “Hoe vaak ben je de afgelopen twee maanden op school gepest?”.

Leerlingen beantwoorden deze vragen op een 5-puntsschaal ($1 = \text{nooit}$ tot $5 = \text{meerdere keren per week}$).

Statistische Analyses³

Data zijn geanalyseerd op basis van een *intention-to-treat* benadering waarbij leerlingen toegewezen aan de interventie geïnccludeerd zijn in de analyse ongeacht of ze daadwerkelijk hebben deelgenomen aan de interventie. We hebben 1-level-modellen geanalyseerd met een correctie voor participanten geclusterd in scholen middels de *complex sample cluster* in *Mplus* (versie 8.2; Muthén & Muthén, 1998-2010). Om alle participanten in de modellen te kunnen includeren, hebben we *Full Information Maximum Likelihood* (FIML) procedures gebruikt. Parameters werden verkregen met *Robust Maximum Likelihood* schattingen (MLR; Muthén & Muthén, 1998-2010).

Om de effecten van de interventie te analyseren, hebben we een serie latente groeicurve (LGC) modellen getest. LGC modellen schatten voor elke participant een individueel groeitraject gebaseerd op hun basisniveau (= intercept) en verandering over tijd (= slope). De individuele groeicurves zijn de indicatoren van latente variabelen die de groeitrajecten van de groep beschrijven (Muthén & Muthén, 1998-2010). De slope is het belangrijkste; wanneer de interventie vergeleken met de Controle conditie effectief is, verandert de interventie de slope in de gewenste richting.

Voor het bepalen van de interventie-effecten, zijn drie dummy variabelen gemaakt (Light, Standaard, en Plus conditie) met de Controle conditie als referentiegroep en toegevoegd als voorspellers van de intercept en slope. Leeftijd en migratieachtergrond van de leerlingen zijn toegevoegd als covariaten. Wanneer twee of meer interventiecondities vergeleken met de Controle conditie effectief bleken, zijn de condities onderling vergeleken met een multi-groep model door de slopes aan elkaar gelijk te stellen en vrij te schatten. De model fits van deze twee geneste modellen zijn vergeleken met de *Satorra-Bentler Scale Chi-Square test* (Satorra & Bentler, 2010). Wanneer deze significant is, past het vrij geschatte model beter bij de data en is dus de ene interventieconditie effectiever dan de andere.

We hebben effectgroottes berekend door de mate van verandering te vermenigvuldigen met de tijdsspanne van de interventie en dat vervolgens gedeeld door de standaarddeviatie van de desbetreffende uitkomst ($d = (\text{slope} * \text{tijdsspanne})/SD$; Feingold, 2013). We hebben effectgroottes berekend tussen de verschillende meetmomenten

³ Uit de power analyse in het protocol paper (zie GEMASKEERD VOOR REVIEW) blijkt een minimale steekproefgrootte van 720 participanten nodig voor deze analyses.

(verandering van T1 naar T2, van T2 naar T3, en van T3 naar T4) en de algemene verandering (verandering van T1 naar T4).

Resultaten

Effecten van R&W

In Tabel 2 staan de gemiddelden en standaarddeviaties van de uitkomsten per conditie op elk meetmoment. De LGC modellen hadden in het algemeen een acceptabele fit (zie voor details [GEMASKEERD VOOR REVIEW]).

De gestandaardiseerde regressiecoëfficiënten van de interventie condities op de slope zijn gerapporteerd in Tabel 3. De effectgroottes tussen de verschillende meetmomenten en de algemene effectgroottes van de interventie condities (in vergelijking met de Controle conditie) zijn gerapporteerd per uitkomst in Tabel 4.

Interpersoonlijke relaties in de klas	3,90 (0,64)	3,81 (0,68)	3,83 (0,67)	3,73 (0,69)	3,98 (0,63)	3,87 (0,72)	3,81 (0,73)	3,80 (0,75)	4,07 (0,62)	4,00 (0,63)	3,92 (0,69)	3,84 (0,81)	3,91 (0,59)	3,78 (0,65)	3,77 (0,67)	3,75 (0,72)
Externaliserend gedrag	0,59 (0,38)	0,47 (0,39)	0,48 (0,39)	0,45 (0,41)	0,54 (0,39)	0,43 (0,38)	0,50 (0,44)	0,42 (0,42)	0,55 (0,35)	0,46 (0,36)	0,55 (0,38)	0,49 (0,45)	0,49 (0,34)	0,42 (0,37)	0,45 (0,39)	0,43 (0,39)
Agressie	1,79 (0,56)	1,71 (0,59)	1,69 (0,59)	1,66 (0,66)	1,85 (0,61)	1,84 (0,73)	1,94 (0,73)	1,83 (0,81)	1,73 (0,46)	1,72 (0,60)	1,80 (0,58)	1,74 (0,81)	1,78 (0,56)	1,76 (0,65)	1,81 (0,67)	1,85 (0,78)
Pesten	1,13 (0,48)	1,18 (0,61)	1,15 (0,55)	1,21 (0,64)	1,13 (0,41)	1,18 (0,60)	1,15 (0,58)	1,19 (0,64)	1,10 (0,41)	1,15 (0,47)	1,20 (0,68)	1,20 (0,64)	1,08 (0,36)	1,18 (0,59)	1,22 (0,71)	1,22 (0,71)
Gepest worden	1,32 (0,77)	1,43 (0,96)	1,30 (0,77)	1,31 (0,86)	1,28 (0,73)	1,35 (0,87)	1,22 (0,77)	1,22 (0,70)	1,36 (0,88)	1,37 (0,85)	1,40 (0,96)	1,40 (0,93)	1,37 (0,94)	1,56 (1,16)	1,41 (1,04)	1,34 (0,95)

Tabel 3

Interventie-effecten over tijd van de R&W condities in vergelijking met de controle conditie

	Light			Standaard			Plus		
	β	SE	p	β	SE	p	β	SE	p
<i>Interventie-specifieke uitkomsten</i>									
Zelfcontrole ¹	0,12	0,03	< ,001	0,00	0,04	,951	0,02	0,04	,532
Zelfreflectie	0,01	0,04	,803	-0,02	0,06	,761	0,04	0,03	,238
Zelfvertrouwen	0,18	0,10	,060	0,09	0,11	,394	-0,05	0,06	,327
Emotionele zelfregulatie	0,17	0,05	< ,001	0,07	0,05	,163	-0,01	0,02	,676
<i>Intrapersoonlijk domein</i>									
Psychologische welzijn	0,15	0,06	,009	0,08	0,08	,287	0,04	0,04	,228
Veerkracht	0,09	0,73	,902	0,18	0,27	,507	-0,10	0,31	,736
Seksuele autonomie ¹	0,18	0,04	< ,001	0,08	0,06	,144	0,06	0,03	,074
Internaliserend gedrag ¹	-0,13	0,04	< ,001	-0,07	0,04	,048	0,01	0,03	,812
<i>Interpersoonlijk domein</i>									
Interpersoonlijke relaties in de klas	0,05	0,03	,056	-0,05	0,07	,499	-0,04	0,04	,278
Externaliserend gedrag	-0,20	0,12	,107	-0,13	0,11	,223	-0,02	0,07	,756
Agressie	-0,19	0,09	,030	0,01	0,10	,933	-0,01	0,05	,822
Pesten	-0,12	0,06	,056	-0,12	0,09	,149	-0,05	0,09	,624
Gepest worden	-0,01	0,04	,817	-0,04	0,06	,527	0,03	0,04	,530

Noot. ¹ Variantie van de voormeting van de desbetreffende uitkomst variabele was vastgezet op nul vanwege negatieve residuele variantie van de geobserveerde variabele op T1.

Tabel 4

Effectgroottes van verandering over tijd in R&W condities in vergelijking met de controle conditie

	Light				Standaard				Plus			
	T1-T2	T2-T3	T3-T4	T1-T4	T1-T2	T2-T3	T3-T4	T1-T4	T1-T2	T2-T3	T3-T4	T1-T4
<i>Interventie-specifieke uitkomsten</i>												
Zelfcontrole ¹	0,19	0,03	0,03	0,25	0,00	0,00	0,00	0,01	0,04	0,01	0,01	0,06
Zelfreflectie	0,01	0,00	0,00	0,02	-0,02	0,00	0,00	-0,03	0,06	0,01	0,00	0,07
Zelfvertrouwen	0,04	0,15	0,04	0,22	0,02	0,08	0,02	0,12	-0,01	-0,05	-0,01	-0,08
Emotionele zelfregulatie	0,17	0,09	0,04	0,29	0,08	0,04	0,02	0,13	-0,01	0,00	0,00	-0,01
<i>Intrapersoonlijk domein</i>												
Psychologisch welzijn	0,07	0,18	-0,01	0,24	0,04	0,10	-0,01	0,14	0,02	0,06	0,00	0,08
Veerkracht	0,04	-0,03	0,02	0,03	0,09	-0,07	0,04	0,06	-0,05	0,04	-0,02	-0,04
Seksuele autonomie ¹	0,32	0,06	0,01	0,38	0,15	0,03	0,00	0,18	0,12	0,02	0,00	0,15
Internaliserend gedrag ¹	0,19	0,04	0,03	0,26	0,11	0,03	0,02	0,15	-0,01	0,00	0,00	-0,01
<i>Interpersoonlijk domein</i>												
Interpersoonlijke relaties in de klas	0,05	0,03	0,02	0,10	-0,05	-0,03	-0,02	-0,10	-0,04	-0,02	-0,02	-0,08
Externaliserend gedrag	0,13	-0,05	0,03	0,11	0,09	-0,03	0,02	0,08	0,02	-0,01	0,00	0,02
Agressie	0,18	0,10	0,03	0,31	-0,01	-0,01	0,00	-0,02	0,01	0,01	0,00	0,02
Pesten	0,10	0,08	0,06	0,24	0,11	0,09	0,07	0,28	0,05	0,04	0,03	0,11
Gepest worden	0,01	0,01	0,00	0,01	0,04	0,02	0,00	0,06	-0,03	-0,02	0,00	-0,04

Noot. ¹ Variantie van de voormeting van de desbetreffende uitkomst variabele was vastgezet op nul vanwege negatieve residuele variantie van de geobserveerde variabele op T1; Effectgroottes (*d*) zijn dusdanig berekend dat positieve effectgroottes een verandering in de gewenste richting voor R&W aangeven.

Interventie-Specifieke Uitkomsten

Leerlingen in de Light conditie toonden, vergeleken met de Controle conditie, gunstigere veranderingstrajecten voor zelfcontrole ($d = 0,25$) en emotionele zelfregulatie ($d = 0,29$; zie Figuur 2). Zij hadden een sterkere stijging in zelfcontrole (slope = 1,75) dan leerlingen in de Controle conditie (slope = 1,44). Met betrekking tot emotionele zelfregulatie lieten leerlingen in de Light conditie een lichte verbetering over tijd zien (slope = 0,24), terwijl leerlingen in de Controle conditie een lichte afname lieten zien (slope = -0,10). Op beide uitkomsten verbeterden leerlingen met name in het eerste jaar van de interventie. Verder laten de resultaten een relevant ($d = 0,22$), maar niet significant ($p = ,060$), verschil zien tussen de veranderingstrajecten van zelfvertrouwen in de Light conditie en Controle conditie. Leerlingen in de Light conditie bleven relatief stabiel (slope = -1,68), terwijl leerlingen in de Controle conditie een kleine afname in zelfvertrouwen rapporteerden (slope = -2,58). Dit interventie-effect was het sterkst tussen T2 en T3. We hebben geen interventie-effecten gevonden voor zelfreflectie noch voor leerlingen in de Standaard en Plus condities. Over het algemeen leek de Light conditie de meest gunstige veranderingstrajecten te bewerkstelligen op de interventie specifieke uitkomsten, met name in het begin van de interventie.

[Figuur 2]

Overkoepelende Uitkomsten

Intrapersoonlijk domein. Ook voor psychologisch welzijn, seksuele autonomie, en internaliserend gedrag hadden leerlingen in de Light conditie positievere veranderingstrajecten dan leerlingen in de Controle conditie (range $d = 0,26-0,38$; zie Figuur 3). Zij bleven stabiel in hun psychologisch welzijn (slope = 0,08), terwijl leerlingen in de Controle conditie verslechterden (slope = -0,25). Het interventie-effect was het sterkste tussen T2 en T3. Op seksuele autonomie verbeterden leerlingen in de Light conditie sterker (slope = 0,73) dan leerlingen in de Controle conditie (slope = 0,32), vooral in het eerste jaar van de interventie. Wat betreft internaliserend gedrag rapporteerden leerlingen in de Light en Standaard condities een sterkere afname (slope_{Light} = -0,60; slope_{Standaard} = -0,53) dan leerlingen in de Controle conditie (slope = -0,41), met name in het eerste jaar. Er waren geen interventie-effecten op veerkracht noch voor leerlingen in de Plus conditie.

[Figuur 3]

Wanneer we de leerlingen in de Light conditie vergeleken met de leerlingen in de Standaard conditie op verandering in internaliserend gedrag, paste het vrij geschatte model significant beter op de data, $\Delta\chi^2_{SB}(1) = 6,60$; $p = ,010$, dan het model met gelijkgestelde

slopes. Internaliserend gedrag nam dus sterker af bij leerlingen in de Light conditie (slope = -0,60) dan bij leerlingen in de Standaard conditie (slope = -0,53). In het intrapersoonlijk domein toonden leerlingen in de Light conditie over het algemeen dus de positiefste veranderingen over tijd, waarbij de interventie de grootste impact had aan het begin van het programma.

Interpersoonlijk domein. Leerlingen in de Light conditie hadden tevens een voordeliger veranderingstraject voor agressie vergeleken met de Controle conditie ($d = 0,31$; zie Figuur 3), vooral in het eerste jaar. Zij bleven over tijd relatief stabiel op agressie (slope = 0,09), terwijl leerlingen in de Controle conditie een toename lieten zien (slope = 0,56). Verder lieten de resultaten een relevant ($d = 0,24$), hoewel niet significant ($p = ,056$), verschil zien tussen veranderingstrajecten in pesten in de Light conditie ten opzichte van de Controle conditie. Leerlingen in de Light conditie vertoonden een minder sterke toename in pesten (slope = 2,34) dan leerlingen in de Controle conditie (slope = 2,80; zie Figuur 3). Het interventie-effect was het sterkste in het eerste jaar. Er zijn geen interventie-effecten gevonden op de andere uitkomsten noch voor leerlingen in de Standaard en Plus condities. Ook in het interpersoonlijk domein hadden leerlingen in de Light conditie over het algemeen betere veranderingstrajecten met, wederom, de sterkste interventie-effecten in het eerste jaar van de interventie.

Discussie

Het doel van de huidige studie was om 1) te onderzoeken of R&W een positief effect heeft op vmbo-leerlingen en 2) om te bepalen of verschillende versies van deze interventie verschillende effecten zouden bewerkstelligen. De interventie bleek gematigd effectief in het stimuleren van interventie-specifieke uitkomsten (zelfcontrole, emotionele zelfregulatie, en een trend voor zelfvertrouwen) en in het bevorderen van sommige aspecten van het intra- en interpersoonlijk domein, maar uitsluitend wanneer een kernteam betrokken was bij de implementatie van de interventie en vooral in het eerste jaar. R&W Light was met name effectief in het stimuleren van competenties en het voorkomen van problemen in het intrapersoonlijk domein (psychologisch welzijn, seksuele autonomie, en internaliserend gedrag). In het interpersoonlijk domein leek R&W Light een buffer te zijn tegen toenames in agressie en pesten. Interventie-effecten waren over het algemeen gematigd (effecten van 0,24-0,38) en vergelijkbaar met effecten van andere universele schoolinterventies met soortgelijke doelen. Zo rapporteerden Durlak en collega's (2011) in hun meta-analyse effectgroottes van 0,23 voor positieve houdingen naar jezelf en anderen, 0,24 voor positief sociaal gedrag, 0,22 voor gedragsproblemen, en 0,24 voor emotionele problemen. Mertens en

collega's (2020) vonden in hun meta-analyse effectgroottes van 0,19 en 0,15 op respectievelijk het intrapersoonlijk en interpersoonlijk domein.

Met betrekking tot de interventie-specifieke uitkomsten vonden we op drie van de vier basiscompetenties, die van belang zijn volgens de R&W theorie, positieve interventie-effecten (zelfcontrole, emotionele zelfregulatie, en een trend voor zelfvertrouwen) wat de potentie van de interventie aangeeft. Echter, er is geen effect gevonden op de mate van zelfreflectie van leerlingen. Ook eerdere interventies zijn ineffectief gebleken in het verbeteren van zelfreflectie bij vmbo-leerlingen. Ter Vrugte en collega's (2015), bijvoorbeeld, onderzochten de effectiviteit van een spel om de rekenvaardigheid van vmbo-leerlingen te verbeteren. Dit spel bevatte een reflectiestimulerend component, maar ze vonden hier geen effect op. Mogelijk is de vroege adolescentie niet de beste ontwikkelingsfase om zelfreflectie te meten omdat de metacognitieve vaardigheden, die adolescenten in staat stellen om betekenisvol te reflecteren op gedachten, gevoelens en gedrag (Sauter et al., 2010), nog volop in ontwikkeling zijn. Desalniettemin, aangezien zelfreflectie een potentieel effectief component in schoolinterventies op het voortgezet onderwijs is (Mertens et al., 2020), zou toekomstig onderzoek moeten nagaan wanneer en hoe zelfreflectie het beste gestimuleerd kan worden bij vmbo-leerlingen.

Wat betreft de overkoepelende uitkomsten vonden we de meeste effecten in het intrapersoonlijk domein. Slechts één interventie-effect (agressie en een trend voor pesten) werd gevonden in het interpersoonlijk domein. Mogelijk komt dit doordat in het intrapersoonlijk domein een meer praktische aanpak gebruikt wordt, terwijl een meer verbale aanpak gebruikt wordt bij het interpersoonlijk domein. In de handleiding van de interventie (Ykema, 2002, 2018) wordt de nadruk in de oefeningen en spellen gelegd op gevoelens, emoties, en attitudes van leerlingen (dus het intrapersoonlijk domein; "Was je in balans tijdens het spel?"). De attitudes en gedragingen van leerlingen in relatie met anderen (dus het interpersoonlijk domein) worden voornamelijk aangekaart tijdens (verbale) rolspellen en discussies. Zo bespreken leerlingen bijvoorbeeld wat pesten is en wat leerlingen er tegen kunnen doen (bijv. "Hoe voelde de pester zich toen zij samen 'stop' zeiden?"). Deze meer verbale aanpak sluit minder goed aan bij een praktische leer methode waardoor vmbo-leerlingen mogelijk niet optimaal profiteren van de interventie op het interpersoonlijk domein.

De interventie-effecten waren het sterkste tijdens het eerste gedeelte van de interventie en namen af tot onbeduidende effecten in het tweede jaar; mogelijk is het dus voldoende om alleen het eerste jaar van de interventie te implementeren. Een afname in

effecten in relatief lange interventies is eerder gevonden in meta-analyses (Bakermans-Kranenburg et al., 2003; Cuijpers, 2002; De Mooij et al., 2020). Deze meta-analyses suggereren dat korte-termijn interventies met een beperkte hoeveelheid sessies de voorkeur hebben. Onderzoek heeft ook aangetoond dat participanten die profiteren van een interventie vaak vroeg in de interventie verbetering laten zien, ongeacht de tijdsspanne van de interventie (Lutz et al., 2014; Tadić et al., 2010). Deze afname in interventie-effecten zou mogelijk gerelateerd kunnen zijn aan een afname in de motivatie van leerlingen (Philips & Wennberg, 2014). Toekomstig onderzoek zou kunnen onderzoeken wat de effectiviteit van R&W is indien het tweede jaar van de interventie niet gegeven wordt, of wanneer alleen enkele boostersessies gegeven worden in dat jaar. Alleen het eerste jaar van de interventie implementeren zou wellicht de belasting voor leerlingen kunnen verlichten zonder de effectiviteit van de interventie in gevaar te brengen.

Verder lieten de resultaten zien dat de Light conditie (kernteam) het meest effectief was, wat suggereert dat voor sommige interventies geldt “*less is more*”. Dit is in lijn met de resultaten van de meta-analyse van Durlak en collega’s (2011) die aantoonde dat het positieve effect van het betrekken van meerdere systemen in interventies niet altijd gevonden wordt. In een interventie met een smalle ecologische focus (alleen kernteam) zijn slechts een paar personen betrokken; zij zijn dus als enige verantwoordelijk voor het correct en effectief toepassen van de interventie, wat de implementatie zou kunnen versterken (Durlak, 2016). In een interventie met een brede ecologische focus (bijv. toevoegen van een leerkracht- en/of oudercomponent) zijn veel mensen betrokken die de verantwoordelijkheid voor implementatie kunnen verdelen. Deze gedeelde verantwoordelijkheid zou mogelijk een “bijstandereffect” kunnen uitlokken, verminderde verantwoordelijkheidsgevoelens van individuen wanneer meerdere mensen aanwezig zijn (Fischer et al., 2011). Toekomstig onderzoek zou zich moeten richten op de mogelijk invloed van bijstandereffecten wanneer de verantwoordelijkheid voor implementatie verdeeld is onder een relatief grote groep mensen.

De resultaten van de huidige studie moeten geïnterpreteerd worden in het licht van zijn sterke en zwakke punten. Sterke punten van de studie zijn de meerdere meetmomenten, de drie condities die verschilden in de ecologische focus, en de grote steekproef. Een zwak punt is het gebrek aan informatie over de inhoud van de betrokkenheid van het schoolteam en de oudercomponent. Hoewel we weten dat deze schoolteam- en oudercomponenten geïmplementeerd zijn, weten we relatief weinig over *hoe* interventietechnieken toegepast zijn door schoolteam en ouders. Ten tweede, de R&W trainers gaven de interventie voor de eerste

keer, na drie trainingsdagen. Meer ervaren trainers zouden mogelijk meer verandering kunnen bewerkstelligen. Ten derde, de uitkomsten zijn enkel gemeten door middel van zelfrapportage van de leerlingen. Daarnaast waren er aan het begin van de studie kleine verschillen tussen de condities op uitkomstmaten wat effect gehad zou kunnen hebben op de resultaten, ondanks dat we hier in de analyses voor gecontroleerd hebben, en hebben we SES niet onderzocht. Ten vierde, we hebben effecten meteen na de interventie onderzocht. Toekomstig onderzoek zou ook lange-termijn effecten van de interventie moeten onderzoeken. Tot slot, in deze studie is R&W onderzocht onder een specifieke doelgroep, namelijk vmbo-leerlingen, en bij een beperkt aantal scholen verspreid over Nederland. Hoewel het belangrijk is om inzicht te krijgen in welke interventies effectief zijn voor deze doelgroep, kunnen de bevindingen niet zonder verder onderzoek gegeneraliseerd worden naar andere groepen (leerlingen) en alle vmbo-leerlingen van Nederland. Daarnaast zou het interessant zijn voor vervolgonderzoek om meer inzicht te krijgen in verschillen in effectiviteit voor bepaalde subgroepen (bijv. jongens versus meisjes, jongeren met verschillen in persoonlijkheid: Mertens et al., 2022a) of schoolkenmerken (bijv. locatie en schoolgrootte). Tenslotte, het dient opgemerkt te worden dat de theorie achter R&W slechts gedeeltelijk onderzocht is. Bijvoorbeeld, Mertens et al. (2022b) hebben laten zien dat toenames in zelfregulatie en zelfvertrouwen inderdaad gerelateerd zijn aan toenames in veerkracht en psychologisch welzijn, zoals de R&W theorie veronderstelt. Er is echter meer onderzoek nodig om de overige aspecten van deze theorie empirisch te onderbouwen en na te gaan hoe de R&W theorie zich verhoudt tot andere theorieën over de psychosociale ontwikkeling van adolescenten.

Conclusie

Met name vmbo-leerlingen lijken behoefte te hebben aan een effectieve interventie die hun competenties kan stimuleren en de ontwikkeling van problemen kan voorkomen in zowel het intra- als het interpersoonlijk domein (Stevens & De Looze, 2018). R&W kan een positieve invloed hebben op vmbo-leerlingen, met name in het verbeteren van het intrapersoonlijk domein en met name in het eerste jaar van de interventie. In het interpersoonlijk domein fungeert de interventie mogelijk als een buffer tegen toenames in agressie en pesten. Verder is gebleken dat effecten niet noodzakelijkerwijs toenemen bij een langere interventie of door meer mensen erbij te betrekken (dus door een bredere ecologische implementatiestrategie).

Onze bevindingen hebben belangrijke implicaties voor onderzoek en de praktijk. Zo laat onze studie zien dat verschillende implementatiestrategieën van dezelfde interventie

effect kan hebben op de effectiviteit ervan. Dit onderstreept het belang om bij de evaluatie van een interventie niet alleen het programma zelf maar ook de implementatiestrategie te onderzoeken. Voor de praktijk laat onze studie zien dat interventies met een actieve en praktijkgerichte leerstijl positieve effecten kunnen hebben op vmbo-leerlingen, een kortere interventie periode mogelijk voldoende is om verandering te bewerkstelligen en het betrekken van meer leerkrachten en ouders niet altijd een positief effect heeft. Voor scholen zou het daarom waardevol kunnen zijn om te investeren in interventies met een leerstijl die aansluit bij die van de doelgroep, een goede kwaliteit training voor een kernteam van het schoolteam en een hoge kwaliteit van implementatie met een smalle ecologische focus.

Trial Registratie

Nederlands Trial Register nummer NL6371 (oud nummer: NTR 6554), geregistreerd op 3 juli 2017.

Verklaring van Concurrerende Belangen

Geen.

Dankbetuigingen

Deze studie is gefinancierd door een beurs van ZonMw, nummer 531001106

Auteurschap (CRediT)

EM: Conceptualisatie, methodologie, formele analyses, dataverzameling, schrijven – conceptversie, schrijven – revisies, projectmanagement; **MD:** Conceptualisatie, methodologie, schrijven – revisies, supervisie; **MvL:** Conceptualisatie, methodologie, schrijven – revisies, supervisie; **ER:** Conceptualisatie, methodologie, schrijven – revisies, supervisie, subsidie verworven.

Korte Verklaring over de Data en Analyses

De analyse codes en output zijn vrij toegankelijk via Open Science Framework ([GEMASKEERD VOOR REVIEW]). De data zijn wegens gevoeligheid niet vrij toegankelijk, maar op aanvraag beschikbaar bij de eerste auteur.

Literatuur

- Anderson, L. M., Reilly, E. E., Gorrell, S., Schaumberg, K., & Anderson, D. A. (2016). Gender-based differential item function for the difficulties in emotion regulation scale. *Personality and Individual Differences, 92*, 87–91.
- Bakermans-Kranenburg, M. J., IJzendoorn, M. H. van, & Juffer, F. (2003). Less is more: Meta-analyses of sensitivity and attachment interventions in early childhood. *Psychological Bulletin, 129*, 195–215.
- Bishop, D. C., Pankratz, M. M., Hansen, W. B., Albritton, J., Albritton, L., & Strack, J. (2014). Measuring fidelity and adaptation: Reliability of an instrument for school-based prevention programs. *Issues and Solutions in Risky Behavior Prevention Research, 37*, 231–257.
- Boor-Klip, H. J., Segers, E., Hendrickx, M. M. H. G., & Cillessen, A. H. N. (2016). Development and psychometric properties of the classroom peer context questionnaire. *Social Development, 25*, 370–389.
- Bronfenbrenner, U. (1979). *The ecology of human development*. Harvard University Press.
- Bruijn, D. de, Overmaat, M., Glaudé, M., Heemskerk, I., Leeman, Y., Roeleveld, J., & Venne, L. van de (2005). Krachtige leeromgevingen in het middelbaar beroepsonderwijs: Vormgeving en effecten. *Pedagogische Studiën, 82*, 77–95.
- Chorpita, B. F., Reise, S., Weisz, J. R., Grubbs, K., Becker, K. D., Krull, J.L., & The Research Network on Youth Mental Health (2010). Evaluation of the brief problem checklist: Child and caregiver interviews to measure clinical progress. *Journal of Consulting and Clinical Psychology, 78*, 526–536.
- Cuijpers, P. (2002). Effective ingredients of school-based drug prevention programs. A systematic review. *Addictive Behaviors, 27*, 1009–1023.
- Davidson, J. R. T., & Connor, K. M. (2017). *Connor-Davidson Resilience Scale (CD-RISC) manual*. Unpublished and partly accessible at www.cd-risc.com.
- Dodge, K. A., & Coie, J. D. (1987). Social-information-processing factors in reactive and proactive aggression in children's peer groups. *Journal of Personality and Social Psychology, 53*, 1146–1158.
- Durlak, J. A. (2016). Programme implementation in social and emotional learning: Basic issues and research findings. *Cambridge Journal of Education, 4*(3), 333–345.
<https://doi.org/10.1080/030574X.2016.1142504>

- Durlak, J. A., Weissberg, R. P., Dymnicki, A. B., Taylor, R. D., & Schellinger, K. B. (2011). The impact of enhancing students' social and emotional learning: A meta-analysis of school-based universal interventions. *Child Development, 82*, 405–432.
- Feingold, A. (2013). A regression framework for effect size assessments in longitudinal modeling of group differences. *Review of General Psychology, 17*, 111–121.
- Finkenauer, C., Engels, R. C. M. E., Baumeister, R. F. (2005). Parenting behaviour and adolescent behavioural and emotional problems: The role of self-control. *International Journal of Behavioral Development, 29*, 58–69.
- Fischer, P., Krueger, J. I., Greitemeyer, T., Vogrincic, C., Kastenmüller, A., Frey, D., Heene, M., Wicher, M., & Kainbacher, M. (2011). The bystander-effect: A meta-analytic review on bystander intervention in dangerous and non-dangerous emergencies. *Psychological Bulletin, 137*, 517–537.
- Graaf, H. de, Meijer, S., Poelman, J., & Vanwesenbeeck, I. (2005). *Seks onder je 25e*. Uitgeverij Eburon.
- Graaf, I. de, Haas, S. de, Zaagsma, M., & Wijssen, C. (2016). Effects of rock and water: An intervention to prevent sexual aggression. *Journal of Sexual Aggression, 22*, 4–19.
- Harter, S. (1988). *Manual for the Self-Perception Profile for Adolescents*. University of Denver.
- Kuosmanen, T., Clarke, A., & Barry, M. M. (2019). Promoting adolescents' mental health and wellbeing: Evidence synthesis. *Journal of Public Mental Health, 18*(1), 73–83. <https://doi.org/10.1108/JPMH-07-2018-0036>
- Lutz, W., Hofmann, S. G., Rubel, J., Boswell, J. F., Shear, M. K., Gorman, J. M., Woods, S. W., & Barlow, D. H. (2014). Patterns of early change and their relationship to outcome and early treatment termination in patients with panic disorder. *Journal of Consulting and Clinical Psychology, 82*, 287–297.
- Mahoney, J. L., Weissberg, R. P., Greenberg, M. T., Dusenbury, L., Jagers, R. J., Niemi, K., Schlinger, M., Schlund, J., Shriver, T. P., VanAusdal, K., & Yoder, N. (2021). Systemic social and emotional learning: Promoting educational success for all preschool to high school students. *American Psychologist, 76*(7), 1128–1142. <https://doi.org/10.1037/amp0000701>
- Mertens, E. C. A., Deković, M., Leijten, P., Londen, M. van, & Reitz, E. (2020). Components of school-based interventions stimulating students' intrapersonal and interpersonal domains: A meta-analysis. *Clinical Child and Family Psychology Review, 23*, 605–631. <https://doi.org/10.1007/s10567-020-00328-y>

- Mertens, E. C. A., Deković, M., Londen, M. van, & Reitz, E. (2022a). Personality as a moderator of intervention effects: Examining differential susceptibility. *Personality and Individual Differences, 186 Part A*. Article number 111323.
<https://doi.org/10.1016.j.paid.2021.111323>
- Mertens, E. C. A., Deković, M., Londen, M. van, & Reitz, E. (2022b). Parallel changes in positive youth development and self-awareness: The role of emotional self-regulation, self-esteem, and self-reflection. *Prevention Science, 23*, 502-512.
<https://doi.org/10.1007/s11121-022-01345-9>
- Modecki, K. L., Zimmer-Gembeck, M. J., & Guerrà, N. (2017). Emotion regulation, coping, and decision making: Three linked skills for preventing externalizing problems in adolescence. *Child Development, 88*, 417–426.
- Mooij, B. de, Fekkes, M., Scholte, R. H. J., & Overbeek, G. (2020). Effective components of social skills training programs for children and adolescents in nonclinical samples: A multilevel meta-analysis. *Clinical Child and Family Psychology Review, 23*, 250–264.
- Muthén, L. K., & Muthén, B. O. (1998–2010). *Mplus user's guide* (6th ed.). Muthén & Muthén.
- Park, D., Tsukayama, E., Goodwin, G., Patrick, S., & Duckworth, A. L. (2017). A tripartite taxonomy of character: Evidence for intrapersonal, interpersonal, and intellectual competencies in children. *Contemporary Educational Psychology, 48*, 16–27.
- Pellegrino, J. W., & Hilton, M. L. (2012). *Education for life and work: Developing transferable knowledge and skills in the 21st century*. National Academy of Sciences.
- Philips, B., & Wennberg, P. (2014). The importance of therapy motivation for patients with substance use disorders. *Psychotherapy, 51*, 555–562.
- Ravens-Sieberer, U., & The European KIDSCREEN Group (2006). *The KIDSCREEN questionnaires. Quality of life questionnaires for children and adolescents – Handbook*. Pabst Science Publisher.
- Satorra, A., & Bentler, P. M. (2010). Ensuring positiveness of the scaled difference chi-square test statistic. *Psychometrika, 75*, 243–248.
- Sauter, F. M., Heyne, D., Blöte, A. W., Widenfelt, B. M. van, & Westenberg, P. M. (2010). Assessing therapy-relevant cognitive capacities in young people: Development and psychometric evaluation of the self-reflection and insight scale for youth. *Behavioral and Cognitive Psychotherapy, 38*, 303–317.
- Smit, K., Brabander, C. J. de, & Martens, R. L. (2014). Student-centred and teacher-centred learning environment in pre-vocational secondary education: Psychological needs,

- and motivation. *Scandinavian Journal of Educational Research*, 58(6), 695–712.
<https://doi.org/10.1080/00313831.2013.821090>
- Solberg, M. E., & Olweus, D. (2003). Prevalence estimation of school bullying with the Olweus bully/victim questionnaire. *Aggressive Behavior*, 29, 239–268.
- Stevens, G., & Looze, M. de (2018). Welbevinden en psychische problemen. In M. Boer & G. Stevens (Eds.), *HBSC 2017: Gezondheid, welzijn en opvoeding van jongeren in Nederland* (pp. 63–75). Ridderprint BV.
- Tadić, A., Helmreich, I., Mergl, R., Hautzinger, M., Kohnen, R., Henkel, V., & Hegerl, U. (2010). Early improvement is a predictor of treatment outcome in patients with mild major, minor or subsyndromal depression. *Journal of Affective Disorders*, 120, 86–93.
- Timmermans, A., Naaijer, H., Keuning, J., & Zijlsling, D. (2017). *Cohortonderzoek COOL5-18 basisrapport meting VO-3 in 2014*. GION (onderwijs/onderzoek).
- Vrugte, J. ter, Jong, T. de, Wouters, P., Vandercruyssen, S., Elen, J., & Oostendorp, H. van (2015). When a game supports prevocational math education but integrated reflection does not. *Journal of Computer Assisted Learning*, 31, 462–480.
- Ykema, F. (2002). *Het Rots en Water perspectief. Een psychofysische training voor jongens. Basisboek*. Uitgeverij SWP.
- Ykema, F. (2018). *Rots en Water praktijkboek. Psychofysische training voor jongens én meisjes*. Uitgeverij SWP.
- Ykema, F., Hartman, D., & Imms, W. (2006). *Bringing it together. Includes 22 case studies of Rock & Water in practice in various settings*. Callaghan: University of Newcastle, Family action Centre and Gadaku institute.

Vijfde reactie reviewer 1– 13-9-2023

Aanvaard manuscript