

Peer-reviewverslag ‘Stevig als een rots, flexibel als water? Effectiviteit van een schoolinterventie gericht op het intrapersoonlijk en interpersoonlijk domein’

Originele manuscript en reviews

23-11-2022	Originele manuscript	Pagina 2
17-1-2023	Reactie reviewer 1	Pagina 29
7-2-2023	Reactie reviewer 2	Pagina 31
7-2-2023	Reactie reviewer 3	Pagina 32

Tweede versie

14-3-2023	Reactiebrief van auteur	Pagina 35
14-3-2023	Tweede versie van het manuscript	Pagina 48
19-4-2023	Tweede reactie reviewer 1	Pagina 76
15-3-2023	Tweede reactie reviewer 2	Pagina 79
14-3-2023	Tweede reactie reviewer 3	Pagina 80

Derde versie

15-5-2023	Tweede reactiebrief van auteur	Pagina 81
15-5-2023	Derde versie van het manuscript	Pagina 71
14-6-2023	Derde reactie reviewer 1	Pagina 116

Vierde versie

23-6-2023	Derde reactiebrief van auteur	Pagina 119
23-6-2023	Vierde versie van het manuscript	Pagina 124
15-8-2023	Vierde reactie reviewer 1	Pagina 154

Vijfde versie

28-8-2023	Vierde reactiebrief van auteur	Pagina 155
28-8-2023	Vijfde versie van het manuscript	Pagina 159
13-9-2023	Vijfde reactie reviewer 1	Pagina 189

Publicatie

Acceptatie: 13-09-2023

Online: 22-11-2023

Originele manuscript – 23-11-2022

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 vmbo-leerlingen, 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 proximale 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. Deze resultaten laten zien dat R&W positieve effecten kan bewerkstelligen bij vmbo-leerlingen wanneer deze geïmplementeerd wordt door een kernteam, hoewel gepoogd moet worden de effectiviteit van de interventie te vergroten.

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 proximal 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. These results show that R&W can establish positive effects among prevocational students when implemented by a core team, though efforts should be made to increase its effectiveness.

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.
5. Het lijkt waardevol om te investeren in goede training voor en kwalitatief hoogstaande implementatie door een kernteam.

Aantal woorden:

4998 (hoofdtekst)

Introductie

Scholen spelen 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 en zouden scholen bewust de competenties van leerlingen moeten stimuleren op beide domeinen om mogelijke problemen te voorkomen (Pellegrino & Hilton, 2012).

Met name vmbo-leerlingen zouden baat kunnen hebben van stimulering van competenties en het voorkomen van problemen. Zij rapporteren, vergeleken met havo- en vwo-leerlingen, over het algemeen een lager niveau van welzijn, meer gedragsproblemen, en meer problemen met leeftijdsgenoten (Stevens & De Looze, 2018). Vmbo-leerlingen worden het beste bereikt en blijven het meest gemotiveerd en betrokken wanneer een praktijkgerichte leer methode wordt gehanteerd. Deze leer methode houdt in dat er een leeromgeving gecreëerd wordt met een praktijkgerichte insteek, met activerende, stapsgewijze, gestructureerde taken waarbij leerlingen hun nieuw geleerde vaardigheden kunnen toepassen en hun eigen doelen kunnen stellen, waarbij korte momenten van instructie of reflectie worden 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 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 proximale uitkomsten van de interventie en vormen de basisvaardigheden voor de ontwikkeling van de bredere 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 het bieden van alternatieve gedragingen aan leerlingen 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. Nederland, Australië, China), is er weinig bekend over de effectiviteit. Verschillende kleinschalige studies hebben aangetoond dat na het afronden van R&W deelnemers zich veerkrachtiger voelden, een meer positieve identiteit hadden, en meer actieve coping strategieën toepasten (Ykema et al., 2006). De Graaf en collega's (2016) hebben de effectiviteit van R&W op seksuele agressie onderzocht bij jongens op het vmbo. Na het afronden van R&W verminderden zelf gerapporteerde 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 proximale en distale (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 socio-ecologisch model van Bronfenbrenner (1979), wordt gedrag bepaald door interacties tussen verschillende systemen zoals individu-, familie-, klas-, en schoolsystemen. Mogelijk bevordert het actief betrekken van gerelateerde systemen zowel generalisatie als behoud van nieuwe vaardigheden (Mahoney et al., 2021). [Gemaskeerd] (2020) concludeerden 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 implementatie versies 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 socio-ecologische systemen in interventies inconsistente resultaten laat zien.

Methoden

Design en procedure

De effectiviteit van R&W is onderzocht in scholen met een vmbo-traject in een gerandomiseerde gecontroleerde studie met vier condities. In de *Light* conditie was alleen een kernteam van leerkrachten betrokken bij de interventie. In de *Standaard* conditie was naast het kernteam het hele schoolteam betrokken. In de *Plus* conditie was een ouder component toegevoegd aan de Standaard 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).

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	0,009	0,009
Jongens, <i>n</i> (%)	170 (48%)	161 (56%)	131 (55%)	199 (57%)	7,38	0,061	0,077
Westerse achtergrond, <i>n</i> (%)	291 (82%)	115 (43%)	211 (91%)	198 (59%)	182,01	< 0,001	0,392

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.

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.

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 coping strategieën 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 hoe ze leerlingen kunnen helpen die gepest worden. Tijdens en na

de oefeningen geven leerlingen feedback en reflecteren ze. Trainers (voornamelijk gymleerkrachten) 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.

Verskillende versies van R&W

In de Light conditie werd de interventie geïmplementeerd door een kernteam van leerkrachten (R&W trainers) die de 3-daagse training succesvol hadden afgerond. In de Standaard en de Plus condities heeft, naast dit kernteam van R&W trainers, de rest van het schoolteam een 3-daagse introductie cursus ontvangen over de basisprincipes van R&W, ondersteuning van R&W trainers, en toepassing van de interventietechnieken in reguliere lessen. In de Plus conditie kregen ouders daarnaast 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. In alle drie de condities ontvingen de R&W trainers supervisie van hun persoonlijke R&W coach.

Controle conditie

In de Controle conditie voerden scholen hun gebruikelijk beleid. Leerlingen hadden bijvoorbeeld 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: 1) zelfrapportage van R&W trainers na elke derde les in het eerste jaar en elke tweede les in het tweede jaar van de interventie en 2) observaties van 67 R&W lessen (Light = 28 lessen; Standaard = 17 lessen, Plus = 22 lessen) door drie R&W experts.

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 vonden de R&W experts de lessen van goede (38%) tot zeer goede (54%)

kwaliteit. De uitvoering van de R&W lessen verschilden 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 was goed tot zeer goed.

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 ($F(2,24) = 5,73$; $p = 0,009$). 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 ($F(2,75) = 4,37$; $p = 0,016$) en het vaker bij hun eigen kind toegepast te hebben ($F(2,75) = 3,44$; $p = 0,037$). 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 ouder component blijken geïmplementeerd te zijn.

Proximale 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 = *helemaal niet* tot 5 = *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*; $\alpha = 0,88-0,91$).

Distale 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 = *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 beantwoorden deze vragen op een 5-puntsschaal (1 = *nooit* tot 5 = *meerdere keren per week*).

Statistische Analyses

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. Participanten waren geclusterd in scholen waarvoor de *complex sample cluster* in *Mplus* (versie 8.2; Muthén & Muthén, 1998-2010) is toegepast. 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 interventie condities 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 werden 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 interventie conditie effectiever dan de andere.

We hebben effect groottes berekend door de mate van verandering te vermenigvuldigen met de tijdsspanne van de interventie en dat vervolgens gedeeld door de standaard deviatie van de desbetreffende uitkomst ($d = (\text{slope} * \text{tijdsspanne})/SD$; Feingold, 2013). We hebben effect groottes 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 standaard deviaties 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 effect groottes tussen de verschillende meetmomenten en de algemene effect groottes van de interventie condities (in vergelijking met de Controle conditie) zijn gerapporteerd per uitkomst in Tabel 4.

Interpersoonlijke	3,90	3,81	3,83	3,73	3,98	3,87	3,81	3,80	4,07	4,00	3,92	3,84	3,91	3,78	3,77	3,75
relaties in de klas	(0,64)	(0,68)	(0,67)	(0,69)	(0,63)	(0,72)	(0,73)	(0,75)	(0,62)	(0,63)	(0,69)	(0,81)	(0,59)	(0,65)	(0,67)	(0,72)
Externaliserend	0,59	0,47	0,48	0,45	0,54	0,43	0,50	0,42	0,55	0,46	0,55	0,49	0,49	0,42	0,45	0,43
gedrag	(0,38)	(0,39)	(0,39)	(0,41)	(0,39)	(0,38)	(0,44)	(0,42)	(0,35)	(0,36)	(0,38)	(0,45)	(0,34)	(0,37)	(0,39)	(0,39)
Agressie	1,79	1,71	1,69	1,66	1,85	1,84	1,94	1,83	1,73	1,72	1,80	1,74	1,78	1,76	1,81	1,85
	(0,56)	(0,59)	(0,59)	(0,66)	(0,61)	(0,73)	(0,73)	(0,81)	(0,46)	(0,60)	(0,58)	(0,81)	(0,56)	(0,65)	(0,67)	(0,78)
Pesten	1,13	1,18	1,15	1,21	1,13	1,18	1,15	1,19	1,10	1,15	1,20	1,20	1,08	1,18	1,22	1,22
	(0,48)	(0,61)	(0,55)	(0,64)	(0,41)	(0,60)	(0,58)	(0,64)	(0,41)	(0,47)	(0,68)	(0,64)	(0,36)	(0,59)	(0,71)	(0,71)
Gepest worden	1,32	1,43	1,30	1,31	1,28	1,35	1,22	1,22	1,36	1,37	1,40	1,40	1,37	1,56	1,41	1,34
	(0,77)	(0,96)	(0,77)	(0,86)	(0,73)	(0,87)	(0,77)	(0,70)	(0,88)	(0,85)	(0,96)	(0,93)	(0,94)	(1,16)	(1,04)	(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
Proximale uitkomsten									
Zelfcontrole ¹	0,12	0,03	< 0,001	0,00	0,04	0,951	0,02	0,04	0,532
Zelfreflectie	0,01	0,04	0,803	-0,02	0,06	0,761	0,04	0,03	0,238
Zelfvertrouwen	0,18	0,10	0,060	0,09	0,11	0,394	-0,05	0,06	0,327
Emotionele zelfregulatie	0,17	0,05	< 0,001	0,07	0,05	0,163	-0,01	0,02	0,676
Intrapersoonlijk domein									
Psychologische welzijn	0,15	0,06	0,009	0,08	0,08	0,287	0,04	0,04	0,228
Veerkracht	0,09	0,73	0,902	0,18	0,27	0,507	-0,10	0,31	0,736
Seksuele autonomie ¹	0,18	0,04	< 0,001	0,08	0,06	0,144	0,06	0,03	0,074
Internaliserend gedrag ¹	-0,13	0,04	< 0,001	-0,07	0,04	0,048	0,01	0,03	0,812
Interpersoonlijk domein									
Interpersoonlijke relaties in de klas	0,05	0,03	0,056	-0,05	0,07	0,499	-0,04	0,04	0,278
Externaliserend gedrag	-0,20	0,12	0,107	-0,13	0,11	0,223	-0,02	0,07	0,756
Agressie	-0,19	0,09	0,030	0,01	0,10	0,933	-0,01	0,05	0,822
Pesten	-0,12	0,06	0,056	-0,12	0,09	0,149	-0,05	0,09	0,624
Gepest worden	-0,01	0,04	0,817	-0,04	0,06	0,527	0,03	0,04	0,530

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

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

	Light				Standaard				Plus			
	T1-T2	T2-T3	T3-T4	Overall	T1-T2	T2-T3	T3-T4	Overall	T1-T2	T2-T3	T3-T4	Overall
Proximale uitkomsten												
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
Intrapersoonlijk domein												
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
Interpersoonlijk domein												
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; Effect groottes (*d*) zijn dusdanig berekend dat positieve effect groottes een verandering in de gewenste richting voor R&W aangeven.

Proximale Uitkomsten

Leerlingen in de Light conditie toonden gunstigere veranderingstrajecten voor zelfcontrole ($d = 0,25$) en emotionele zelfregulatie ($d = 0,29$) vergeleken met leerlingen in de Controle conditie (zie Figuur 2). Leerlingen in de Light conditie 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 ($p = 0,060$), verschil zien tussen de veranderingstrajecten van zelfvertrouwen in de Light conditie en in de 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 proximale uitkomsten, met name in het begin van de interventie.

[Figuur 2]

Distale Uitkomsten

Intrapersoonlijk domein. Ook voor psychologisch welzijn, seksuele autonomie, en internaliserend gedrag hadden leerlingen in de Light conditie positievere veranderingstrajecten in vergelijking met leerlingen in de Controle conditie (zie Figuur 3; range $d = 0,26-0,38$). In de Light conditie bleven leerlingen 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, $\Delta\chi^2_{SB}(1) = 6,60$; $p = 0,010$, 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 leerlingen in de Controle conditie (zie Figuur 3; $d = 0,31$), vooral in het eerste jaar. Leerlingen in de Light conditie 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 ($p = 0,056$), 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 proximale 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 te fungeren als een buffer 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 effect groottes 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. [Gemaskeerd] (2020) vonden in hun meta-analyse effect groottes van 0,19 en 0,15 op respectievelijk het intrapersoonlijk en interpersoonlijk domein.

Met betrekking tot de proximale 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 reflectie stimulerend 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 in de eerste fase van ontwikkeling zijn. Desalniettemin, aangezien zelfreflectie een potentieel effectief component in schoolinterventies op het voortgezet onderwijs is ([gemaskeerd], 2020), zou toekomstig onderzoek moeten nagaan wanneer en hoe zelfreflectie het beste gestimuleerd kan worden bij vmbo-leerlingen.

Wat betreft de distale 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. Een mogelijke verklaring voor de grotere impact in het intrapersoonlijk domein is dat daarbij 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 leermethode 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, waardoor het mogelijk voldoende is 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 kortetermijn interventies met een bescheiden 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). Mogelijk zou de afname in interventie effecten gerelateerd kunnen zijn aan een afname in de motivatie van leerlingen (Philips & Wennberg, 2014). Alleen het eerste jaar van de interventie implementeren zou dus 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 socio-ecologische focus (alleen kernteam) zijn slechts een paar personen betrokken; zij zijn dus als enige verantwoordelijk voor het correct en effectief implementeren van de interventie, wat de implementatie zou kunnen versterken (Durlak, 2016). In een interventie met een brede socio-ecologische focus (bijv. toevoegen van een leerkracht en/of ouder component) zijn veel mensen betrokken die de verantwoordelijkheid voor implementatie kunnen verdelen. Deze gedeelde verantwoordelijkheid zou mogelijk een “bijstanderseffect” 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 bijstanderseffecten 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 socio-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 ouder component. 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. Tot slot, we hebben effecten meteen na de interventie onderzocht.

Toekomstig onderzoek zou ook lange-termijn effecten van de interventie moeten onderzoeken.

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. 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. 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 bij een interventie met een bredere socio-ecologische implementatiestrategie.

Onze bevindingen hebben belangrijke implicaties voor de praktijk: een kortere interventie periode is mogelijk voldoende om verandering te bewerkstelligen en het betrekken van meer leerkrachten en ouders heeft niet altijd een positief effect. Het zou voor scholen daarom waardevoller kunnen zijn om te investeren in een goede kwaliteit training voor een kernteam van het schoolteam en een hoge kwaliteit van implementatie met een smalle socio-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

[Gemaskeerd] (2020).

- 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.
- 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>
- Modecki, K. L., Zimmer-Gembeck, M. J., & Guerra, 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.
- 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.

Reactie reviewer 1 – 17-1-2023

Ik heb het ingediende artikel over de interventie Rots en Water met interesse gelezen. Het onderwerp is relevant voor JiO. Het onderzoek kent veel sterke punten, zoals een grote steekproef, meerdere meetmomenten, verschillende condities met random toewijzing, een brede waaier aan uitkomstenmaten (met positieve en negatieve indicatoren van welbevinden), aandacht voor implementatiekwaliteit, aandacht voor missende waarden, controlevariabelen, en clusterbias in de analyses. De resultaten zijn relevant en de conclusies die getrokken worden helder en correct.

Niettegenstaande deze sterke punten zijn er ook enkele serieuze tekortkomingen. De theoretische onderbouwing van de interventie is zeer mager. Er wordt niet verwezen naar evidence-based theorie. Ook ontbreekt een review van bestaande evidence-based interventies voor het verbeteren van het welbevinden of psychosociaal functioneren van adolescenten, waardoor het niet duidelijk is wat R&W kan bijdragen aan de bestaande literatuur. Er wordt wel verwezen naar een aantal eerdere kleinschalige studies naar R&W, maar de lezer krijgt daar geen informatie over. Enkel een studie naar seksuele agressie bij jongens wordt besproken, maar dit lijkt niet aan te sluiten bij de primaire doelstelling van R&W. Wat betreft de metingen van de uitkomsten kan opgemerkt worden dat de validiteit van de instrumenten niet wordt besproken (niet in het algemeen en ook niet specifiek voor een vmbo populatie). Daarnaast wordt er enkel gebruik gemaakt van leerlingrapportage op vragenlijsten, wat een risico vormt voor single-method, single-informant bias.

Andere punten van aandacht:

- De keuze voor een populatie VMBO leerlingen en de specifieke problematiek van deze leerlingen krijgt relatief weinig aandacht in de inleiding. Een bespreking van meer empirisch onderzoek zou de argumentatie voor deze populatie kunnen versterken.
- Een power analyse ontbreekt: hoeveel clusters zijn er nodig om kleine effecten te kunnen detecteren (rekening houdend met het aantal condities)?
- Er is geen informatie over de steekproef leerkrachten die R&W implementeerden
- Uit hoeveel leerkrachten bestond een “kernteam”? Hoe werden taken verdeeld? Werden sessies door meerdere leerkrachten gegeven?
- Er is niet onderzocht of er op de voormeting (T1) verschillen waren tussen condities in de uitkomstmaten
- Er is niet onderzocht of subpopulaties van leerlingen anders reageerden op de interventie (bv. leerlingen met een verhoogd risico op psychosociale problemen). Hoewel dit geen doel was van de studie, had het wel interessante inzichten kunnen opleveren (of een piste voor vervolgonderzoek).
- Het is niet duidelijk of er een 2-level model onderzocht is met demografische gegevens op leerlingniveau en conditie op schoolniveau.
- Zou het kunnen dat ouderbetrokkenheid samenhangt met demografische gegevens en/of psychosociaal functioneren van de leerlingen? Mogelijk werpt dat extra licht op de resultaten.
- De generalisatie van de gegevens naar andere populaties mag meer aandacht krijgen in de discussie

- De lijnen in de grafieken zijn moeilijk van elkaar te onderscheiden

Reactie reviewer 2 – 7-2-2023

Onderwerp:

- Het onderwerp is relevant voor het tijdschrift.

Inhoud:

In de samenvatting kan toegevoegd worden hoe R&W verbeterd kan worden

- De inleiding ligt de variabelen toe en de onderzoeksvraag wordt beantwoord.
- Hypothesen worden niet toegelicht, met name de veronderstelling waarom betrokkenheid van ouders invloed zal hebben.

Methode:

- In methode worden de termen licht, standaard en plus geïntroduceerd maar later pas toegelicht. Voorkeur omdat eerder te doen. Ook lijkt er een relatie te zijn tussen uitvoering van kernteam, ouders, uitgebreid schoolteam maar dat blijft in methode nog onduidelijk.
- Waardevol dat er een missing data analyse is gedaan.
- Onduidelijk is het verschil tussen R&W trainers en R&W experts. Er worden uitspraken gedaan over de kwaliteit van de uitvoering maar dit komt suggestief over. Beter uitleggen verhoogd de transparantie.
- Nauwkeurige uitwerking van de data-analyse

Discussie:

- De discussie biedt een overzicht van de bevindingen en de bevindingen worden verbonden aan bestaande literatuur.
- De implicaties van het onderzoek voor onderzoek en praktijk zijn wat mager en kunnen uitgebreid worden.

Opbouw:

- Heldere opbouw. De leesbaarheid kan verhoogd worden door minder technische kopjes te kiezen (bijv. Proximaal, distaal etc. Deze termen zijn bij het lezerspubliek wellicht onbekend).

Schrijfstijl:

- Het artikel is prettig te lezen. De schrijfstijl kan wat minder technische en meer aangepast worden aan het publiek van professionals in de praktijk.

Reactie reviewer 3 – 7-2-2023

Het artikel rapporteert over een interventiestudie, waarbij de effecten van het programma Rots & Water zijn nagegaan op competenties en problemen bij leerlingen van gemiddeld 12.38 jaar. Dit is een interessante studie, met veel sterktes. Het design is een gerandomiseerde gecontroleerde studie met vier condities: light-conditie (enkel kernteam leerkrachten), standaardconditie (kernteam + hele schoolteam), plusconditie (standaardconditie + oudercomponent) en controleconditie (gebruikelijk beleid voor preventie van problemen bij leerlingen en verbeteren van competenties). De randomisatie gebeurde op schoolniveau. Er wordt een overzichtelijke flowchart gegeven met de verschillende fasen in het onderzoek (toewijzing van condities, verloop onderzoek over verschillende meetmomenten) en hun respectievelijke deelnemers (scholen, leerlingen) per conditie en redenen voor uitval.

Er is een drop-out analyse uitgevoerd. Er wordt gerapporteerd over treatment fidelity. Data zijn geanalyseerd volgens de intention-to-treat benadering. In de analyses wordt gecontroleerd voor leeftijd en migratieachtergrond omdat er op dat vlak verschillen zijn tussen condities

De studie werd goedgekeurd door een Ethische Commissie en is geregistreerd bij het Nederlands Trial Register.

De discussie biedt aannemelijke verklaringen voor de gevonden resultaten en heeft oog voor de beperkingen van de studie. De studie genereert zinvolle conclusies voor de praktijk.

Abstract

Graag 'vmbo-leerlingen' voluit schrijven of omschrijven zoals in de Engelse samenvatting, voor Vlaamse lezers.

De zin 'De gematigde interventie-effecten waren het grootste in het eerste jaar': hier is niet duidelijk of de effecten over een specifiek meetmoment gaan, dan wel over een bepaald leerjaar in het middelbaar onderwijs.

Introductie

Na het lezen van de eerste paragraaf, maakte ik mij de bedenking dat de nadruk sterk ligt op het verwerven van competenties of vaardigheden als middel om tot psychologisch welzijn te komen (intrapersoonlijk domein). Dat past uiteraard binnen de idee dat de interventie Rots & Water inzet op competenties. Maar daardoor gaat de nuance verloren dat welzijn niet alleen samenhangt met (of afhankelijk is) van persoonlijke competenties, maar ook met omstandigheden waar de school niet (altijd) kan op inspelen (context van gezin, vrije tijd, maatschappelijke omstandigheden, ...). Scholen 'kunnen' iets doen, maar zijn niet alleen verantwoordelijk. Misschien kan die nuance nog toegevoegd worden?

Kan in de inleiding nog iets gezegd worden over de verwachtingen naar verandering op kortere (na 1 jaar) en langere termijn (na 2 jaar), bijvoorbeeld op basis van eerder onderzoek?

Methode

Kan toegevoegd worden hoe nagegaan is of de uitvoering van R&W gebeurd is volgens de handleiding? Is de zelfrapportage aan de hand van checklists gebeurd? Was er overeenkomst tussen de zelfrapportages van de R&W trainers en de observaties van de R&W experts?

Resultaten

Tabel 4: 'Overall' verduidelijken in de noot waar dit effect betrekking op heeft.

Discussie

Omdat de effecten in het tweede jaar van de interventie onbeduidend zijn, wordt geconcludeerd dat het mogelijk voldoende is om alleen het eerste jaar van de interventie te implementeren. Zou het voor de auteurs ook een optie zijn om te werken met één of meerdere boostersessies?

Een andere beperking van het onderzoek is dat enkel zelfrapportering door leerlingen is gebruikt om uitkomsten te meten.

Conclusie

De eerste zin van de 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' is een aanname die gebaseerd is op eerder onderzoek, en niet op de resultaten van huidige studie. Best referentie Stevens & De Looze toevoegen.

Taal en vorm

Meerdere woorden in de tekst zijn samenstellingen en moeten aaneengeschreven worden, in plaats van los van elkaar. Bijvoorbeeld: oudercomponent, standaardconditie, copingstrategieën, effectgrootte, standaarddeviatie

<https://taaladvies.net/aaneenschrijven-drieledige-samenstelling-of-woordgroep-algemeen/>

In sommige gevallen moet een koppelteken gebruikt worden: <https://www.vlaanderen.be/taaladvies/taaladviezen/koppelteken-of-trema-bij-klinkerbotsing>

<https://woordenlijst.org/leidraad/6>

Bijvoorbeeld: interventie-effect

Gebruik overal punten voor decimale getallen in plaats van komma's.

Bij p-waardes kan de 0 voor het decimale punt weggelaten worden (APA: "Do not use a zero before a decimal when the statistic cannot be greater than 1 (proportion, correlation, level of statistical significance)").

Reactiebrief van auteur – 14-3-2023

Utrecht, 14 maart 2023

Onderwerp: Revisie manuscript

Geachte Prof. dr. Boendermaker en meneer Veenstra,

Bijgevoegd vindt u de 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 reviewers voor hun complimenten en de constructieve feedback.

Wij geloven dat de revisies ons manuscript verhelderd en verbeterd hebben. Zo geven we in de inleiding nu meer informatie over de toedracht en het belang van onze studie. Verder lichten we in de methode sectie de verschillende versies van R&W die we hebben onderzocht eerder toe en bespreken we de manier waarop de uitvoering van de interventie is gemeten uitgebreider. Daarnaast gaan we in de discussie nu meer in op de beperkingen en implicaties van ons onderzoek. Om het reviewproces te faciliteren hebben wij onze reacties op de reviewers toegevoegd in de Appendix hieronder.

Bedankt dat u ons de mogelijkheid geeft een revisie van ons manuscript in te dienen. We zien er naar uit van u te horen,

Met vriendelijke groet,

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

Esther Mertens

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

Appendix – Reacties op Reviewers

Reviewer 1:

Onderwerp:

- Het onderwerp is relevant voor het tijdschrift.

Inhoud:

- In de samenvatting kan toegevoegd worden hoe R&W verbeterd kan worden

Het is op basis van de resultaten van ons onderzoek lastig om concrete verbeterpunten van de interventie voor te stellen, omdat we daar geen onderzoek naar hebben gedaan. We kunnen wel aanbevelingen doen voor het type interventie dat scholen in kunnen zetten en de implementatie methode. In de samenvatting sluiten we daarom nu af met aanbevelingen voor implementatie.

"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.." p. 2

"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." p. 2-3

- De inleiding ligt de variabelen toe en de onderzoeksvraag wordt beantwoord.

- Hypothesen worden niet toegelicht, met name de veronderstelling waarom betrokkenheid van ouders invloed zal hebben.

Bij het beschrijven van het tweede doel van onze studie, halen wij het ecologisch model van Bronfenbrenner aan. Volgens dit model interacteren verschillende systemen rondom het individu met elkaar en met het individu wat vervolgens diens gedrag zou beïnvloeden. Wanneer dit in een interventie context wordt geplaatst, suggereert dat model dat een interventie effectiever zou kunnen worden wanneer meer systemen betrokken zijn bij de interventie; meer systemen passen de interventie inhoud toe, waardoor de interventie meer ingebed is in de sociale omgeving van een individu en de invloed ervan op een individu groter zou kunnen worden. Echter, onderzoek naar het effect op de effectiviteit van het betrekken van meerdere systemen bij een interventie laat geen consistente bevindingen zien. Daarom hebben wij met betrekking tot dit doel van onze studie geen specifieke hypothesen opgesteld. Om deze beredenering duidelijker naar voren te laten komen in onze introductie, benoemen we bij de uitleg van het model van Bronfenbrenner nu expliciet het school- en familiesysteem als voorbeeld van gerelateerde systemen en benoemen we expliciet dat het betrekken van meerdere systemen dus betekent dat de inhoud van de interventie meer ingebed wordt in de sociale omgeving van de leerling. We hopen dat met deze aanvullende uitleg ook de rol van ouders binnen schoolinterventies voldoende naar voren komt.

"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).” p. 7

Methode:

- In methode worden de termen light, standaard en plus geïntroduceerd maar later pas toegelicht. Voorkeur omdat eerder te doen. Ook lijkt er een relatie te zijn tussen uitvoering van kernteam, ouders, uitgebreid schoolteam maar dat blijft in methode nog onduidelijk.

Zoals aangeraden door de reviewer hebben we de uitleg van de condities, waarvoor we de termen light, standaard en plus gebruiken, vooraan in de methode sectie gezet. Ook hebben we de uitleg van de interventie condities uitgebreid waardoor de verschillen in de uitvoering van R&W duidelijker naar voren komen.

“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 dit 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.” p. 7-8

- Waardevol dat er een missing data analyse is gedaan.
- Onduidelijk is het verschil tussen R&W trainers en R&W experts. Er worden uitspraken gedaan over de kwaliteit van de uitvoering maar dit komt suggestief over. Beter uitleggen verhoogd de transparantie.

We hebben informatie toegevoegd over wie de R&W experts zijn en de manier waarop de uitvoering van de interventie is onderzocht. We gaan hierbij meer in op de inhoud van de vragenlijsten ingevuld door de R&W trainers en op de observatieschema's ingevuld door de R&W experts (zie ook reviewer II). Daarnaast hebben we, bij het rapporteren van de bevindingen, de formulering af en toe aangepast zodat het duidelijker naar voren komt dat het subjectieve beoordelingen waren.

“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 lessenvolledig 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.” p. 10

“Over het algemeen beoordeelden de R&W experts de lessen van goede (38%) tot zeer goede (54%) kwaliteit.” p. 10

“Concluderend, het grootste deel van de R&W lessen was daadwerkelijk gegeven en de uitvoering lijkt goed tot zeer goed te zijn geweest.” p. 11

- Nauwkeurige uitwerking van de data-analyse

Discussie:

- De discussie biedt een overzicht van de bevindingen en de bevindingen worden verbonden aan bestaande literatuur.

- De implicaties van het onderzoek voor onderzoek en praktijk zijn wat mager en kunnen uitgebreid worden.

We hebben de implicaties van onze studie voor onderzoek en praktijk verder uitgewerkt. We benadrukken nu dat toekomstig onderzoek zowel een interventieprogramma als de implementatiestrategie moet onderzoeken en dat scholen het beste een interventie kunnen toepassen met een leermethode die aansluit bij de leerstijl van leerlingen die ze ermee willen bereiken.

“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.” p. 23

Opbouw:

- Heldere opbouw. De leesbaarheid kan verhoogd worden door minder technische kopjes te kiezen (bijv. Proximaal, distaal etc. Deze termen zijn bij het lezerspubliek wellicht onbekend).

We hebben deze termen aangepast naar “interventie specifieke uitkomsten” en “overkoepelende uitkomsten”, zowel in de tekst als in de kopjes.

E.g.,

“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

“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.” p. 6

Schrijfstijl:

- Het artikel is prettig te lezen. De schrijfstijl kan wat minder technische en meer aangepast worden aan het publiek van professionals in de praktijk.
We hebben het manuscript doorlopen en formuleringen aangepast om deze minder technisch te maken.

*Reviewer II***Review: Stevig als een rots, flexibel als water? Effectiviteit van een schoolinterventie gericht op het intrapersoonlijk en interpersoonlijk domein"**

Het artikel rapporteert over een interventiestudie, waarbij de effecten van het programma Rots & Water zijn nagegaan op competenties en problemen bij leerlingen van gemiddeld 12.38 jaar. Dit is een interessante studie, met veel sterktes. Het design is een gerandomiseerde gecontroleerde studie met vier condities: light-conditie (enkel kernteam leerkrachten), standaardconditie (kernteam + hele schoolteam), plusconditie (standaardconditie + oudercomponent) en controleconditie (gebruikelijk beleid voor preventie van problemen bij leerlingen en verbeteren van competenties). De randomisatie gebeurde op schoolniveau. Er wordt een overzichtelijke flowchart gegeven met de verschillende fasen in het onderzoek (toewijzing van condities, verloop onderzoek over verschillende meetmomenten) en hun respectievelijke deelnemers (scholen, leerlingen) per conditie en redenen voor uitval. Er is een drop-out analyse uitgevoerd. Er wordt gerapporteerd over treatment fidelity. Data zijn geanalyseerd volgens de intention-to-treat benadering. In de analyses wordt gecontroleerd voor leeftijd en migratieachtergrond omdat er op dat vlak verschillen zijn tussen condities. De studie werd goedgekeurd door een Ethische Commissie en is geregistreerd bij het Nederlands Trial Register.

De discussie biedt aannemelijke verklaringen voor de gevonden resultaten en heeft oog voor de beperkingen van de studie. De studie genereert zinvolle conclusies voor de praktijk.

Abstract

Graag 'vmbo-leerlingen' voluit schrijven of omschrijven zoals in de Engelse samenvatting, voor Vlaamse lezers.

We danken de reviewer voor het onder de aandacht brengen dat mogelijk niet alle lezers op de hoogte zijn van het onderwijssysteem in Nederland. We hebben in de abstract en in de inleiding de eerste keer dat de afkortingen van de verschillende onderwijstrajecten volledig uitgeschreven en de afkorting gegeven.

" (...) interpersoonlijk domein van leerlingen op het voorbereidend middelbaar beroepsonderwijs (vmbo), en of (...) " p. 2

"Met name leerlingen op het voorbereidend middelbaar beroepsonderwijs (vmbo) zouden baat kunnen hebben van 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, (...) " p. 5

De zin 'De gematigde interventie-effecten waren het grootste in het eerste jaar': hier is niet duidelijk of de effecten over een specifiek meetmoment gaan, dan wel over een bepaald leerjaar in het middelbaar onderwijs.

Excuses voor de onduidelijkheid. We hebben dit verduidelijkt in zowel de Nederlandse als de Engelse abstract door expliciet te benoemen dat het hier gaat over het eerste jaar van de tweejarige interventie. Dit hebben we ter verduidelijking ook toegevoegd in de bijsluiter voor de praktijk en in de conclusie van de discussie sectie.

"De (gematigde) interventie-effecten waren het grootste in het eerste jaar van de tweejarige interventie." p. 2

“The (moderate) intervention effects were strongest in the first year of the two-year intervention.” p. 2

“4. Interventie-effecten waren het grootste in het eerste jaar van de interventie.” p. 4

“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.” p. 23

Introductie

Na het lezen van de eerste paragraaf, maakte ik mij de bedenking dat de nadruk sterk ligt op het verwerven van competenties of vaardigheden als middel om tot psychologisch welzijn te komen (intrapersoonlijk domein). Dat past uiteraard binnen de idee dat de interventie Rots & Water inzet op competenties. Maar daardoor gaat de nuance verloren dat welzijn niet alleen samenhangt met (of afhankelijk is) van persoonlijke competenties, maar ook met omstandigheden waar de school niet (altijd) kan op inspelen (context van gezin, vrije tijd, maatschappelijke omstandigheden, ...). Scholen ‘kunnen’ iets doen, maar zijn niet alleen verantwoordelijk. Misschien kan die nuance nog toegevoegd worden?

We danken de reviewer voor deze suggestie en hebben dit toegevoegd.

“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).” p. 5

“Scholen zouden bewust kunnen proberen de competenties van leerlingen op beide domeinen te stimuleren om mogelijke problemen te voorkomen (Pellegrino & Hilton, 2012).” p. 5

Kan in de inleiding nog iets gezegd worden over de verwachtingen naar verandering op kortere (na 1 jaar) en langere termijn (na 2 jaar), bijvoorbeeld op basis van eerder onderzoek? De reviewer haalt een interessant punt aan. Helaas is het niet mogelijk om op basis van eerder onderzoek iets te zeggen over de verwachtingen van veranderingen op kortere en langere termijn. Eerder onderzoek heeft zich vooral gericht op interventie-effecten op de korte termijn, bijvoorbeeld na implementatie van 10 lessen. Bovendien is het meeste onderzoek uitgevoerd in Australië met een focus op jongens. De resultaten van deze onderzoeken kunnen niet zonder meer vertaald worden naar de Nederlandse samenleving, en in het bijzonder vmbo-leerlingen. Het grootschalige onderzoek van De Graaf en collega's (2016) is wel uitgevoerd in Nederland bij vmbo-leerlingen (allemaal jongens) en heeft een follow-up meting 5 maanden na het afronden van de R&W lessen. Echter, deze follow-up periode valt op de tijdlijn van ons onderzoek nog steeds binnen verandering in het eerste jaar van de interventie. Doordat we geen verwachtingen kunnen uitspreken over veranderingen op de kortere en langere termijn op basis van eerder onderzoek, hebben wij besloten geen specifieke hypothesen hierover op te stellen.

Methode

Kan toegevoegd worden hoe nagegaan is of de uitvoering van R&W gebeurd is volgens de handleiding? Is de zelfrapportage aan de hand van checklists gebeurd? Was er overeenkomst tussen de zelfrapportages van de R&W trainers en de observaties van de R&W experts?

We hebben meer informatie toegevoegd over de manier waarop we de uitvoering van R&W hebben gemeten. In de methode sectie gaan we nu meer in op de zelfrapportage door expliciet te benoemen dat dit vragenlijsten waren en twee voorbeeld items toe te voegen inclusief de antwoordcategorieën. Ook geven we meer informatie over de R&W experts en de observatieschema's die zij hebben ingevuld. We hebben de zelfrapportages van de R&W trainers niet direct vergeleken met de observaties van de R&W experts, omdat dit een vertekend beeld zal geven. De zelfrapportages hebben namelijk betrekking op meerdere R&W lessen, terwijl de observaties over één specifieke les gaan. Desalniettemin zien we bij beide methodes een vergelijkbaar beeld met betrekking tot de uitvoering van R&W, waarbij de R&W trainers in de zelfrapportage mogelijk iets strenger waren dan de observaties van de R&W experts, wat voor ons een bevestiging is dat beide methodes een betrouwbare weergave zijn van de uitvoering van R&W.

“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 lessenvolledig 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.” p. 10

Resultaten

Tabel 4: ‘Overall’ verduidelijken in de noot waar dit effect betrekking op heeft.

We hebben in Tabel 4 “Overall” aangepast naar “T1-T4”, zodat dit in lijn is met de manier waarop we de andere effecten van de verschillende meetmomenten hebben weergegeven.

Discussie

Omdat de effecten in het tweede jaar van de interventie onbeduidend zijn, wordt geconcludeerd dat het mogelijk voldoende is om alleen het eerste jaar van de interventie te implementeren. Zou het voor de auteurs ook een optie zijn om te werken met één of meerdere boostersessies?

De reviewer haalt hier een relevant punt aan. Mogelijk zouden boostersessies in het tweede jaar een positief effect hebben, maar hier kunnen we geen stellige uitspraken over doen. Onderzoek naar de effectiviteit van boostersessies laat namelijk geen eenduidige resultaten zien; boostersessies dragen niet altijd bij aan interventie-effecten (bijv. Neil & Christensen, 2009; Lochman et al., 2014). Om de mogelijkheid boostersessies te implementeren in het tweede jaar wel aan te kunnen stippen, hebben we dit als suggestie voor toekomstig onderzoek toegevoegd in de discussie sectie.

“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 dus wellicht de belasting voor leerlingen kunnen verlichten zonder de effectiviteit van de interventie in gevaar te brengen.” p. 22

Een andere beperking van het onderzoek is dat enkel zelfrapportering door leerlingen is gebruikt om uitkomsten te meten.

We hebben de suggestie van de reviewer overgenomen en in de beperkingen opgenomen dat de uitkomsten alleen door middel van zelfrapportage zijn gemeten.

“Ten derde, de uitkomsten zijn enkel gemeten door middel van zelfrapportage van de leerlingen.” p. 22-23

Conclusie

De eerste zin van de 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’ is een aanname die gebaseerd is op eerder onderzoek, en niet op de resultaten van huidige studie. Best referentie Stevens & De Looze toevoegen.

We hebben de referentie toegevoegd aan de desbetreffende zin.

Taal en vorm

Meerdere woorden in de tekst zijn samenstellingen en moeten aaneengeschreven worden, in plaats van los van elkaar. Bijvoorbeeld: oudercomponent, standaardconditie, copingstrategieën, effectgrootte, standaarddeviatie

<https://taaladvies.net/aaneenschrijven-drieledige-samenstelling-of-woordgroep-algemeen/>

In sommige gevallen moet een koppelteken gebruikt

worden: <https://www.vlaanderen.be/taaladvies/taaladviezen/koppelteken-of-trema-bij-klinkerbotsing>

<https://woordenlijst.org/leidraad/6>

Bijvoorbeeld: interventie-effect

Wij danken de reviewer voor de oplettendheid op de schrijfwijze van samengestelde woorden.

We hebben de aanpassingen van de reviewer doorgevoerd in de tekst en de volledige tekst nogmaals doorlopen met extra aandacht voor samenstellingen.

Met betrekking tot “Standaard conditie” hebben wij ervoor gekozen deze als twee losse woorden te schrijven om zo te benadrukken dat “Standaard” de naam is die wij hebben gegeven om de desbetreffende conditie aan te duiden en zodat de schrijfwijze van deze conditie in lijn is met de andere drie condities (Light, Plus, en Controle).

Gebruik overal punten voor decimale getallen in plaats van komma’s.

Omdat het manuscript in het Nederlands geschreven is, hebben wij ervoor gekozen om decimalen met een ‘komma’ aan te duiden. In het Nederlands worden decimalen aangegeven met komma’s en duizendtallen met een punt (dus tegenovergesteld aan de Engelse schrijfwijze).

Bij p-waarden kan de 0 voor het decimale punt weggelaten worden (APA: “Do not use a zero before a decimal when the statistic cannot be greater than 1 (proportion, correlation, level of statistical significance”).

In lijn met de aanbeveling van de reviewer, hebben we de “0” voor de p-waarden weggehaald.

Reviewer III

Ik heb het ingediende artikel over de interventie Rots en Water met interesse gelezen. Het onderwerp is relevant voor JiO. Het onderzoek kent veel sterke punten, zoals een grote steekproef, meerdere meetmomenten, verschillende condities met random toewijzing, een brede waaier aan uitkomstenmaten (met positieve en negatieve indicatoren van welbevinden), aandacht voor implementatiekwaliteit, aandacht voor missende waarden, controlevariabelen, en clusterbias in de analyses. De resultaten zijn relevant en de conclusies die getrokken worden helder en correct. Er zijn echter ook enkele belangrijke vragen/aandachtspunten:

- 1) In de inleiding wordt niet duidelijk waarom onderzoek naar de interventie Rots en Water relevant is. Wat voor interventies zijn er op dit terrein en waarom is het nodig om juist naar deze interventie onderzoek te doen? Wat kan dit onderzoek bijdragen aan de kennis over interventies voor het verbeteren van welbevinden of psychosociaal functioneren van adolescenten?

Wij danken de reviewer voor het aangeven dat de relevantie van het onderzoeken van de effectiviteit van R&W sterker naar voren kan komen. De relevantie van het huidige onderzoek zit vooral in het onderzoeken van een interventie die in Nederland veelvuldig wordt toegepast, terwijl tot nog toe nauwelijks onderzoek gedaan is naar de effectiviteit daarvan. We hebben ons hierbij gericht op een groep leerlingen die met name gebaat zouden zijn bij een interventie ter ondersteuning van hun psychosociaal functioneren. In de inleiding geven we nu expliciet aan dat R&W vaak wordt toegepast in Nederland en dat het daarom van belang is meer inzicht te krijgen in de interventie-effecten van R&W.

“Hoewel R&W in veel landen geïmplementeerd wordt (bijv. Nederland, Australië, China), 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.” p. 6

- 2) Wat voor eerdere of andere studies zijn er gedaan naar deze interventie? Er wordt gesproken van ‘Verschillende kleinschalige studies’ en een studie naar R&W bij seksuele agressie. Deze laatste lijkt een heel andere doelgroep te betreffen. Kan hier meer informatie over gegeven worden.

We hebben de beschrijving van de kleinschalige studies verder toegelicht door aan te geven waar, door wie en in welke context deze studies zijn uitgevoerd. Het onderzoek van De Graaf en collega's (2015) is belangrijk om aan te halen, omdat net als in de huidige studie De Graaf en collega's onderzoek hebben gedaan onder vmbo-leerlingen. We hebben de tekst aangepast om deze overeenkomst in doelgroep, en daarmee de relevantie van die studie, te benadrukken.

“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 coping strategieën 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 zelf gerapporteerde dwingende strategieën en verbale manipulatie, terwijl zelfregulatie toenam.” p. 6

- 3) Aansluitend hierop: De keuze voor een populatie VMBO leerlingen en de specifieke problematiek van deze leerlingen krijgt relatief weinig aandacht in de inleiding. Een bespreking van meer empirisch onderzoek zou de argumentatie voor deze populatie kunnen versterken.

Zoals aanbevolen door de reviewer halen we aanvullende empirisch onderzoek aan bij het beschrijven van de problematiek van vmbo-leerlingen en onze keuze voor de populatie vmbo-leerlingen. We gaan in de inleiding nu meer in op zowel de problemen die vmbo-leerlingen ervaren als de uitdaging om deze groep leerlingen te betrekken bij een schoolinterventie.

“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 autonomie en 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.” p. 5

- 4) Wat betreft de metingen van de uitkomsten valt op dat de validiteit van de instrumenten niet wordt besproken (niet in het algemeen en ook niet specifiek voor een vmbo populatie).

In ons onderzoek hebben wij gebruik gemaakt van gestandaardiseerde vragenlijsten die regelmatig in onderzoek toegepast worden. De validiteit van deze vragenlijsten is niet onderzocht voor specifiek een vmbo populatie, aangezien dit een zeer specifieke, Nederlandse, doelgroep is. We kunnen dus niet de validiteit van de vragenlijsten voor een vmbo populatie rapporteren. We hebben ervoor gekozen om niet de validiteit van de vragenlijsten in het algemeen te rapporteren; voor dergelijke informatie is het voor de geïnteresseerde lezer informatiever om de psychometrische artikelen over de desbetreffende vragenlijsten te lezen.

- 5) Daarnaast wordt er enkel gebruik gemaakt van leerlingrapportage op vragenlijsten, wat een risico vormt voor single-method, single-informant bias.

We zijn het eens met de reviewer dat enkel het gebruik van leerlingrapportage op vragenlijsten een limitatie is. We hebben deze aan de beperkingen toegevoegd in de discussie sectie (zie ook reviewer II).

“Ten derde, de uitkomsten zijn enkel gemeten door middel van zelfrapportage van de leerlingen.” p. 22-23

Andere punten van aandacht:

- Zou er een power analyse toegevoegd kunnen worden? Hoeveel clusters zijn er nodig om kleine effecten te kunnen detecteren (rekening houdend met het aantal condities)? In het protocol paper van dit onderzoeksproject staat de power analyse uitgebreid beschreven. In het huidige manuscript hebben we middels een voetnoot bij de paragraaf “Statistische Analyses” naar deze power analyse verwezen en de minimale benodigde steekproefgrootte ($N = 720$) expliciet genoemd.

“Uit de power analyse in het protocol paper (zie GEMASKEERD VOOR REVIEW) blijkt een minimale steekproefgrootte van 720 participanten nodig voor deze analyses.” p. 13

- Kan er informatie worden gegeven over de steekproef leerkrachten die R&W implementeerden?

We hebben de demografische gegevens van de leerkrachten die R&W implementeerden toegevoegd bij de beschrijving en uitvoering van R&W.

“In totaal verzorgden 49 leerkrachten tussen de 23 en 64 jaar oud ($M = 36.77$, $SD = 10.02$) de R&W lessen, van wie iets meer dan de helft man was ($n = 27$; 56).” p. 10

- Uit hoeveel leerkrachten bestond een “kernteam”? Hoe werden taken verdeeld? Werden sessies door meerdere leerkrachten gegeven?

De grootte van een kernteam was afhankelijk van de grootte van een school, c.q. het aantal klassen dat R&W lessen zou ontvangen. Het aantal leerkrachten dat getraind werd en de precieze manier waarop R&W geïmplementeerd werd, werd bepaald door de scholen zelf in overleg met de ontwikkelaars van R&W en de R&W supervisors. Het doel van ons onderzoeksproject was om de effectiviteit van R&W te onderzoeken onder ‘real world’ omstandigheden, dus zoals R&W geïmplementeerd zou worden buiten de kaders van ons onderzoeksproject. Met de kanttekening dat scholen wel gebonden waren aan de uitvoering binnen de kaders van de conditie (Light, Standaard, of Plus) waar ze in ingedeeld waren.

- Is er onderzocht of er verschillen waren tussen condities in de uitkomstmaten in de voormeting (T1)?

De interventie-effecten zijn onderzocht door middel van ongespecificeerde groei modellen. Bij groei modellen wordt de voormeting gebruikt als startpunt van de verandering over tijd, dus als intercept. Wij hebben de dummy variabelen die ‘conditie’ weergaven ook als voorspeller van de intercept toegevoegd om op die manier te corrigeren voor verschillen tussen condities in de uitkomstmaten (ongeacht over deze verschillen significant waren).

- Is er onderzocht of subpopulaties van leerlingen anders reageerden op de interventie (bv. leerlingen met een verhoogd risico op psychosociale problemen). Hoewel dit geen doel was van de studie, had het wel interessante inzichten kunnen opleveren (of een piste voor vervolgonderzoek).

Wij danken de reviewer voor deze suggestie en sluiten ons hier graag bij aan; naast het onderzoeken van interventie-effecten voor de gehele steekproef, is het ook interessant om interventie-effecten te onderzoeken bij subgroepen binnen de steekproef. Zo hebben wij bijvoorbeeld onderzocht of persoonlijkheidskenmerken van leerlingen samenhangen met de mate waarin ze konden profiteren van de interventie en hebben

hierover in een internationaal tijdschrift gepubliceerd. Dit betreft echter een andere onderzoeksvraag dan waar we ons binnen het huidige manuscript op hebben gericht. We hebben er daarom voor gekozen de nadruk te leggen op het gedetailleerd analyseren van de interventie-effecten voor de gehele steekproef en tussen de verschillende condities.

- Kan er inzicht worden gegeven of er een 2-level model onderzocht is met demografische gegevens op leerlingniveau en conditie op schoolniveau? Hoewel 13 scholen binnen een onderzoeksproject een behoorlijk aantal is, zijn dit te weinig clusters om betrouwbare analyses op schoolniveau uit te kunnen voeren. Bovendien lag onze interesse met betrekking tot de effectiviteit van de interventie meer op leerling- dan op schoolniveau. We hebben er daarom voor gekozen de data te analyseren op leerling niveau en in onze analyses te corrigeren voor het clusteren op schoolniveau middels de *complex sample cluster* (zoals beschreven in ons analyse plan, p. 13).
- Zou het kunnen dat ouderbetrokkenheid samenhang met demografische gegevens en/of psychosociaal functioneren van de leerlingen? Mogelijk werpt dat extra licht op de resultaten.
De reviewer haalt hier een interessant punt aan. Helaas kunnen wij dit met onze data niet betrouwbaar onderzoeken. Wij hebben alleen informatie van ouders die de vragenlijsten hebben ingevuld en het wel of niet invullen van een vragenlijst voor onderzoek kan niet één op één vertaald worden naar de mate van ouderbetrokkenheid. Wat betreft ouderbetrokkenheid hebben wij ons binnen dit onderzoek vooral gericht op betrokkenheid bij de interventie in de *Plus* conditie om te controleren of de condities inderdaad van elkaar verschilden zoals wij van te voren hadden bedacht. Verdere uitspraken over ouderbetrokkenheid kunnen wij binnen dit onderzoeksproject niet doen.
- Kan de generalisatie van de gegevens naar andere populaties meer aandacht krijgen in de discussie?
We hebben de interventie binnen een specifieke doelgroep onderzocht; een doelgroep van wie de leermethode goed aan lijkt te sluiten bij de aanpak van de interventie. Hierdoor is generalisatie van de gegevens naar andere populaties niet zonder meer te maken. We benoemen dit nu expliciet in de discussie sectie bij de beperkingen.
“Tot slot, in deze studie is R&W onderzocht onder een specifieke doelgroep, namelijk vmbo-leerlingen. 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).” p. 23
- De lijnen in de grafieken zijn moeilijk van elkaar te onderscheiden, zou dat op een andere manier weergegeven kunnen worden?
We hebben de lijnen in de grafieken iets verdikt en op elke lijn per conditie tekens toegevoegd als extra identificatie.

Tweede versie van het manuscript – 14-3-2023

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 intrapersonaal 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:

5556 (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 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. Nederland, Australië, China), 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 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]).

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).

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

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.

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.

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 (voornamelijk gymleerkrachten) 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. In totaal verzorgden 49 leerkrachten tussen de 23 en 64 jaar oud ($M = 36.77$, $SD = 10.02$) de R&W lessen, van wie iets meer dan de helft man was ($n = 27$; 56%).

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 = *helemaal niet* tot 5 = *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 = *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 beantwoorden deze vragen op een 5-puntsschaal (1 = *nooit* tot 5 = *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. Participanten waren geclusterd in scholen waarvoor de *complex sample cluster* in Mplus (versie 8.2; Muthén & Muthén, 1998-2010) is toegepast. 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

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

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.

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

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 intrapersonaal 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>

- Modecki, K. L., Zimmer-Gembeck, M. J., & Guerra, 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., 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.

Tweede reactie reviewer 1 – 19-4-2023

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

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.

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.

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.

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 benoemd worden in de discussie als een limitatie.

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.

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.

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.

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).

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.

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.

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.

Minor feedback:

-Ik denk niet dat het klopt dat metacognitieve vaardigheden “nog in de eerste fase van ontwikkeling zijn”; ik zou dat iets genuanceerder formuleren.

Ik hoop dat deze feedback behulpzaam is. Ik kijk uit naar de volgende versie van het artikel.

Tweede reactie reviewer 2 - 15-3-2023

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.

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 intrapersoonlijk 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-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%).” (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 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:

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 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.

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 = *helemaal niet* tot 5 = *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 = *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 beantwoorden 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 daadwerkelijk hebben deelgenomen aan de interventie. We hebben 1-level-modellen geanalyseerd met een correctie voor participanten geclusterd in scholen middels de *complex*

² 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.

Interpersoonlijke	3,90	3,81	3,83	3,73	3,98	3,87	3,81	3,80	4,07	4,00	3,92	3,84	3,91	3,78	3,77	3,75
relaties in de klas	(0,64)	(0,68)	(0,67)	(0,69)	(0,63)	(0,72)	(0,73)	(0,75)	(0,62)	(0,63)	(0,69)	(0,81)	(0,59)	(0,65)	(0,67)	(0,72)
Externaliserend	0,59	0,47	0,48	0,45	0,54	0,43	0,50	0,42	0,55	0,46	0,55	0,49	0,49	0,42	0,45	0,43
gedrag	(0,38)	(0,39)	(0,39)	(0,41)	(0,39)	(0,38)	(0,44)	(0,42)	(0,35)	(0,36)	(0,38)	(0,45)	(0,34)	(0,37)	(0,39)	(0,39)
Agressie	1,79	1,71	1,69	1,66	1,85	1,84	1,94	1,83	1,73	1,72	1,80	1,74	1,78	1,76	1,81	1,85
	(0,56)	(0,59)	(0,59)	(0,66)	(0,61)	(0,73)	(0,73)	(0,81)	(0,46)	(0,60)	(0,58)	(0,81)	(0,56)	(0,65)	(0,67)	(0,78)
Pesten	1,13	1,18	1,15	1,21	1,13	1,18	1,15	1,19	1,10	1,15	1,20	1,20	1,08	1,18	1,22	1,22
	(0,48)	(0,61)	(0,55)	(0,64)	(0,41)	(0,60)	(0,58)	(0,64)	(0,41)	(0,47)	(0,68)	(0,64)	(0,36)	(0,59)	(0,71)	(0,71)
Gepest worden	1,32	1,43	1,30	1,31	1,28	1,35	1,22	1,22	1,36	1,37	1,40	1,40	1,37	1,56	1,41	1,34
	(0,77)	(0,96)	(0,77)	(0,86)	(0,73)	(0,87)	(0,77)	(0,70)	(0,88)	(0,85)	(0,96)	(0,93)	(0,94)	(1,16)	(1,04)	(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 leermethode 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. 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., 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. (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., & Guerra, N. (2017). Emotion regulation, coping, and decision making: Three linked skills for preventing externalizing problems in adolescence. *Child Development, 88*, 417–426.

Moosj, 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., 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 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 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.
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 geïnccludeerd, 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 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.
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 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 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 leerkrachterservaring. 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 = *helemaal niet* tot 5 = *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 = *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 beantwoorden deze vragen op een 5-puntsschaal (1 = *nooit* tot 5 = *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

³ 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.

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

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 intrapersoonlijk 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, Happytes) 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 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]).

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.

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
--	-----------	-----------	-----------	-----------	--------	--------	-------

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 = *helemaal niet* tot 5 = *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 = *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 beantwoorden deze vragen op een 5-puntsschaal (1 = *nooit* tot 5 = *meerdere keren per week*).

Statistische Analyses⁴

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

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

Vijfde reactie reviewer 1– 13-9-2023

Aanvaard manuscript