

De Effectiviteit van een Schoolinterventie voor het Vergroten van een Prosociaal Klasklimaat voor Basisschoolkinderen

Amanda W.G. van Loon¹ en Tessa M.L. Kaufman¹

¹*Sociale Wetenschappen, Educatie en Pedagogiek, Universiteit Utrecht, Utrecht, Nederland*

Samenvatting Een prosociaal klasklimaat is cruciaal voor het bevorderen van het school- en psychologische welzijn van kinderen, evenals voor het optimaliseren van een effectieve leeromgeving. Schoolprogramma's kunnen een prosociaal klasklimaat bevorderen, maar er zijn weinig op evidence gebaseerde opties beschikbaar in zowel Nederland als internationaal. De huidige studie onderzocht daarom de effectiviteit van een basisschoolprogramma gericht op het bevorderen van prosocialiteit door de intrinsieke prosociale motivatie van leerlingen te verhogen door de basispsychologische behoeften te versterken (autonomie, competentie en verbondenheid). Gegevens zijn afkomstig van 2337 leerlingen (46% jongen, $M = 9.55$ jaar, $SD = 1.21$) en 225 leerkrachten (76% vrouw, $M = 38.69$ jaar, $SD = 11.09$) van 42 KiVa-scholen. Scholen werden gerandomiseerd in een experimentele conditie ($N = 20$ scholen) of een controleconditie ($N = 22$ scholen). Regressieanalyses werden uitgevoerd halverwege en aan het einde van programma-implementatie, gecorrigeerd voor baseline niveaus, waarbij condities werden vergeleken in de volledige steekproef en in een subgroep met ideale programmatrouw. Resultaten in de gehele steekproef lieten geen enkel significant effect zien voor leerlingen. Leerkrachten rapporteerden wel een positievere eigen houding ten opzichte van het stimuleren van prosociaal gedrag en autonomie. Wanneer het programma volgens plan werd geïmplementeerd (met een subgroep van $N = 202-253$ leerlingen) werden kleine effecten gevonden, voornamelijk halverwege het schooljaar, waarbij deze leerlingen gemiddeld positiever rapporteerden dan leerlingen in de controleconditie wat betreft basispsychologische behoeften, een prosociaal klasklimaat, welbevinden en veiligheid op school, en antisociaal gedrag. Er werden geen effecten gevonden voor leerkrachten. Concluderend werden er voor leerlingen geen effecten in de volledige steekproef gevonden, alleen kortdurende kleine effecten bij een kleine subgroep die ideale programmatrouw rapporteerde, wat verder onderzoek vereist. Daarnaast is inzicht in programma-implementatie essentieel, aangezien de mate van programmatrouw belangrijk lijkt te zijn voor programma-effectiviteit.

Trefwoorden schoolinterventies, kinderen, prosociaal klasklimaat, welbevinden, zelfdeterminatietheorie

Artikelgeschiedenis

Ontvangen: 14 mei 2025

Geaccepteerd: 17 december 2025

Online: 18 maart 2026

Contactpersoon

Amanda W.G. van Loon,
a.w.g.vanloon@uu.nl

Copyright

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

Financiering Onderzoek

Dit onderzoek werd financieel ondersteund door een subsidie van het Nationaal Kennisinstituut Onderwijs (NKO, voorheen NRO). De financierende instantie heeft geen rol gespeeld in het ontwerp van het onderzoek, noch in het schrijven van dit manuscript of de beslissing om het te publiceren.

Belangen

De auteurs hebben geen belangen te melden.

Themanummer:

Positieve Ontwikkeling en Veerkracht bij Kinderen en Jongeren

1 Inleiding

Om het school- en psychologisch welzijn van leerlingen te bevorderen en een effectieve leeromgeving te optimaliseren, is een prosociaal klasklimaat van cruciaal belang (Evans et al., 2009; Wang et al., 2020). Een positief, zorgzaam en ondersteunend klasklimaat wordt geassocieerd met minder pestgedrag en slachtofferschap van pesten (Saarento & Salmivalli, 2015; Thornberg et al., 2022). Prosociale klassen worden gekenmerkt door prosociaal gedrag, een prosociale norm in de klas, een hoge mate van samenwerking, weinig conflicten en hoogwaardige interacties tussen leerkrachten en klasgenoten (Khalfaoui et al., 2021; Villardón-Gallego et al., 2018; Wang et al., 2020). Het huidige onderzoek onderzoekt de effectiviteit van een schoolprogramma dat een prosociaal klasklimaat in basisscholen promoot, door het verhogen van intrinsieke prosociale motivatie en het vervullen van basispsychologische behoeften.

Meta-analyses van internationale interventies laten zien dat schoolbrede programma's voor positief gedrag en sociaal en emotioneel leren (SEL) kleine tot grote effecten hadden op het verbeteren van het schoolklimaat, prosociaal gedrag, sociale en emotionele vaardigheden, positieve attitudes, een positief zelfbeeld, psychologisch welzijn, burgerschapsvaardigheden en verminderd antisociaal gedrag bij kinderen en adolescenten (Charlton et al., 2021; Cipriano et al., 2023; Durlak et al., 2011; Gaffney et al., 2021; Sklad et al., 2012; Tanner-Smith et al., 2018; Taylor et al., 2017). Onderzoek toonde ook de effectiviteit van interventieprogramma's aan die prosociaal gedrag bij kinderen en adolescenten bevorderen (Caprara et al., 2015; Mesurado et al., 2018; Shin & Lee, 2021), met name voor programma's die ontwikkeld zijn om de sociale competentie te vergroten in plaats van probleemgedrag te voorkomen (Shin & Lee, 2021). Daarnaast toonden meta-analyses aan dat programma's gericht op het versterken van basispsychologische behoeften (dat wil zeggen autonomie, competentie en verbondenheid) geassocieerd waren met een toename van basisbehoeften, intrinsieke motivatie, en fysiek en psychologisch welzijn (Ntoumanis et al., 2021; Wang et al., 2024). Bovendien liet eerder onderzoek op basis- en middelbare scholen, waarbij programma's werden geïmplementeerd gericht op motivatie en bevrediging van basispsychologische behoeften, positieve effecten op basisbehoeften, prosociaal gedrag en een positief klasklimaat zien, evenals een afname van antisociaal gedrag (Cheon et al., 2019; Manzano-Sánchez & Valero-Valenzuela, 2019). Hoewel de meeste onderzoeken positieve effecten lieten zien, was niet elke interventie even effectief op alle uitkomsten die werden gemeten of waren er gemixte resultaten. Effecten werden bijvoorbeeld niet gevonden voor verbondenheid (Wang et al., 2024), verbale agressie (Caprara et al., 2015) of fysieke gezondheid (Taylor et al., 2017), terwijl deze effecten wel in andere studies werden gevonden. Ondanks dat meta-analyses voornamelijk positieve resultaten laten zien zijn de effecten veelal klein (bijvoorbeeld Shin & Lee, 2021; Taylor et al., 2017), zijn effecten afhankelijk van implementatie (bijvoorbeeld Cipriano et al., 2023; Durlak et al., 2011), en zijn er methodologische rigoureuzere studies nodig die effecten op een valide en betrouwbare manier evalueren (bijvoorbeeld Charlton et

al., 2021). Er is daarom behoefte aan (cluster) gerandomiseerde gecontroleerde onderzoeken om de effecten van dit soort programma's in diverse, grote steekproeven van basisschoolleerlingen uitgebreid te evalueren.

Er zijn ook Nederlandse interventies die het klasklimaat proberen te versterken, waaronder KiVa, PRIMA, Kanjertraining, De Vreedzame School, *School-Wide Positive Behavior Support* (SWPBS), de superheldenaanpak, en Kwink (Nederlands Jeugdinstituut, 2025). Voor een aantal interventies is bewijs voor positieve effecten op het verminderen van pesten en slachtofferschap van pesten en een toename van het schoolklimaat, schoolwelbevinden, prosociaal gedrag, zelfvertrouwen en burgerschapsvaardigheden (KiVa, SWPBS, De Vreedzame School, Kanjertraining, PRIMA; Huitsing et al., 2020; Nederlands Jeugdinstituut, 2025; Orobio De Castro et al., 2018; van Verseveld et al., 2021; Vermande, 2015). De andere interventies zijn nog niet onderbouwd met wetenschappelijk onderzoek, en zijn alleen erkend als goed onderbouwd (superheldenaanpak, Kwink; Nederlands Jeugdinstituut, 2025). Al deze interventies bestaan uit een bepaald aantal lessen, en niet een aanpak die doorlopend tijdens het hele schooljaar wordt uitgevoerd. Daarnaast ligt de primaire focus van deze interventies op pesten (KiVa, PRIMA, Kanjertraining), sociale vaardigheden (superheldenaanpak, Kwink), burgerschap (de Vreedzame School), of het voorkomen van ongewenst gedrag (SWPBS, Kwink). De huidige interventie richt zich voornamelijk op het verbeteren van het klasklimaat en de onderlinge relaties in de groep op een prosociale manier (SterkWerk, een Nederlandse aanpassing van *Meaningful Roles*).

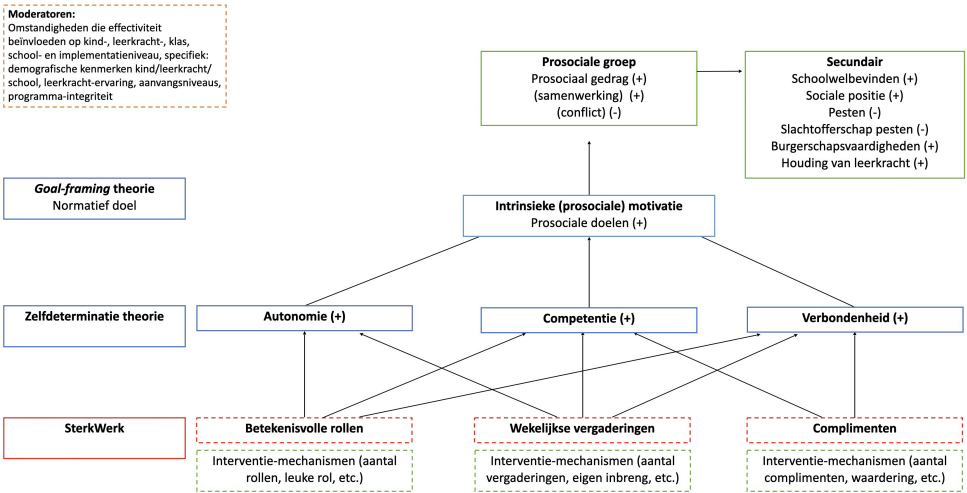
Gebaseerd op de theorie van basispsychologische behoeften (Ryan & Deci, 2017) wordt gesteld dat het creëren van een prosociaal klasklimaat strategieën vereist die de intrinsieke motivatie van leerlingen voor positief gedrag stimuleren door te voorzien in hun basispsychologische behoeften (Cheon et al., 2019; Ryan & Deci, 2000, 2017). Deze behoeften hebben betrekking op autonomie (het gevoel van controle over eigen beslissingen), competentie (het vertrouwen in eigen kunnen) en verbondenheid (het gevoel van verbondenheid en veiligheid). Vervulling van deze basispsychologische behoeften wordt geassocieerd met meer empathie voor anderen, betere vaardigheden om met conflicten om te gaan, internalisering van een prosociale klasnorm, een zorgzame klasomgeving en de motivatie voor prosociaal gedrag (Cheon et al., 2019). De *goal-framing* theorie verklaart hoe dergelijk gedrag in klassen kan blijven bestaan: de heersende groepsnorm beïnvloedt welk gedrag wordt ondersteund en versterkt en welk gedrag niet (Lindenberg & Steg, 2007). Prosociaal gedrag moet daarom een sociale norm in de klas worden om te zorgen dat dit gedrag een blijvend onderdeel wordt van het klasklimaat. Het handhaven van een dergelijke norm gebeurt door het waarnemen van prosociale attitudes en gedrag, en door de steun van anderen (Lindenberg & Steg, 2007).

SterkWerk is gebaseerd op het Amerikaanse programma *Meaningful Roles* (Ellis et al., 2016), een schoolbreed programma met rollen en complimenten. Leerlingen op een middelbare school kregen rollen toegewezen, zoals studentenadministrateur, omroeper, gangmonitor of veiligheidsingenieur. Daarnaast schreven leerlingen complimenten

voor prosociaal gedrag onder klasgenoten. Deze complimenten werden voor iedereen zichtbaar opgehangen. Het hoofddoel van het programma was om pesters prosociale alternatieven voor hun gedrag te bieden door het vergroten van autonomie en verbondenheid. De implementatie van de interventie ging gepaard met significante verminderingen in alle vier de factoren op schoolniveau: minder ruzies per maand, minder dagen afwezig per maand, een lagere frequentie van ziekte/letsel per maand en minder nablijven per maand (Ellis et al., 2016). Het programma is aangepast voor basisschoolleerlingen en focust zich niet specifiek op pesters en het bieden van alternatieve, prosociale, manieren om belangrijke sociale en materiele doelen te bereiken, maar op de gehele groep/klas en het promoten van prosociaal gedrag voor elke leerling. Sommige rollen zijn bijvoorbeeld aangepast zodat ze meer passend zijn voor de basisschoolcontext (zoals computer- of rekenexpert, feestdagkapitein, tuinier).

SterkWerk is een aanpak gedurende het hele schooljaar voor groep 5 tot en met 8 die uitgevoerd wordt door leerkrachten. Het doel van SterkWerk is om het klasklimaat en de onderlinge verhoudingen in de groep te verbeteren, op een positieve, prosociale manier. Het programma bestaat uit drie componenten: 1) rollen: leerlingen betekenisvolle rollen bieden om verantwoordelijkheid en samenwerking te stimuleren, 2) klassenvergaderingen: democratische bijeenkomsten in de klas faciliteren om wederzijds begrip en betrokkenheid te bevorderen, en 3) complimenten: leerlingen stimuleren om complimenten te geven om verbindingen en zelfvertrouwen op te bouwen. Eerder onderzoek liet positieve effecten zien van deze (losse) componenten, waaronder verantwoordelijkheid en prosociaal gedrag voor rollen (Ellis et al., 2016; Embry, 2004), verantwoordelijkheid, verbondenheid en prosociaal gedrag voor klassenvergaderingen (Battistich, 2005; Edwards & Mullis, 2003; Gfroerer et al., 2024), en sociale competentie en prosociaal gedrag voor complimenten (Embry, 2004; Embry & Biglan, 2008). Op basis van de theorie van basispsychologische behoeften en de goal-framing theorie is de verwachting dat deze drie componenten de basispsychologische behoeften van leerlingen vervullen, wat zorgt voor intrinsieke motivatie voor prosociaal gedrag, wat leidt tot meer prosociaal gedrag (primaire uitkomsten) en een positief klasklimaat (bijvoorbeeld hoger schoolwelzijn, minder antisociaal gedrag; secundaire uitkomsten) (zie Figuur 1 voor het conceptuele model).

Om de hiaten uit eerder onderzoek en (Nederlandse) interventies aan te pakken, beoogt deze studie de kennis over onderzoek naar welzijn op school op twee manieren te vergroten door: a) empirische ondersteuning te bieden aan het theoretisch kader dat het versterken van basispsychologische behoeften een prosociaal klasklimaat kan bevorderen, en b) gebruik te maken van een cluster gerandomiseerd gecontroleerd design (RCT) – een gouden standaard in effectiviteitsonderzoek (Hariton & Locascio, 2018). Specifiek onderzoekt de huidige studie – als eerste doel – de effectiviteit van een schoolprogramma dat beoogt een prosociaal klasklimaat te bevorderen (de Nederlandse versie van Meaningful Roles, zie Figuur 1 voor het conceptuele model). We verwachten voor de primaire uitkomsten 1) vervulling van basispsychologische behoeften (dat wil zeggen autonomie,



Figuur 1 Conceptueel model

competentie, verbondenheid), 2) een verhoogde intrinsieke prosociale motivatie, en 3) toename van prosociaal gedrag (dat wil zeggen prosociaal gedrag, prosociale groepsnorm, meer samenwerking, minder conflicten). Voor de secundaire uitkomsten verwachten we 4) een verhoogd schoolwelzijn en veiligheid op school, 5) een afname van pesten en slachtofferschap van pesten, 6) een toegenomen positieve perceptie van leerlingen over de leerkrachthouding ten opzichte van pesten, 7) een toegenomen positieve (eigen) houding van leerkrachten ten opzichte van het stimuleren van prosociaal gedrag en de vervulling van basispsychologische behoeften van leerlingen (gerapporteerd door leerkrachten) en 8) verbeterde burgerschapsvaardigheden van leerlingen (gerapporteerd door leerkrachten) op scholen die de interventie implementeren in vergelijking met controlescholen.

Als tweede doel hebben we ook de effecten van programmatrouw bepaald – wanneer een programma wordt geïmplementeerd zoals bedoeld (Ma et al., 2023; Sanetti et al., 2021). Gebrek aan effectiviteit kan namelijk voortkomen uit een slechte implementatie, en niet omdat het een inherent ineffectief programma is (Lane et al., 2004; Sanetti et al., 2021). Programma's die met een hogere getrouwheid worden geïmplementeerd, leveren over het algemeen positievere resultaten op, wat de rol van implementatiekwaliteit op programmasucces benadrukt (Durlak et al., 2011; Haataja et al., 2014).

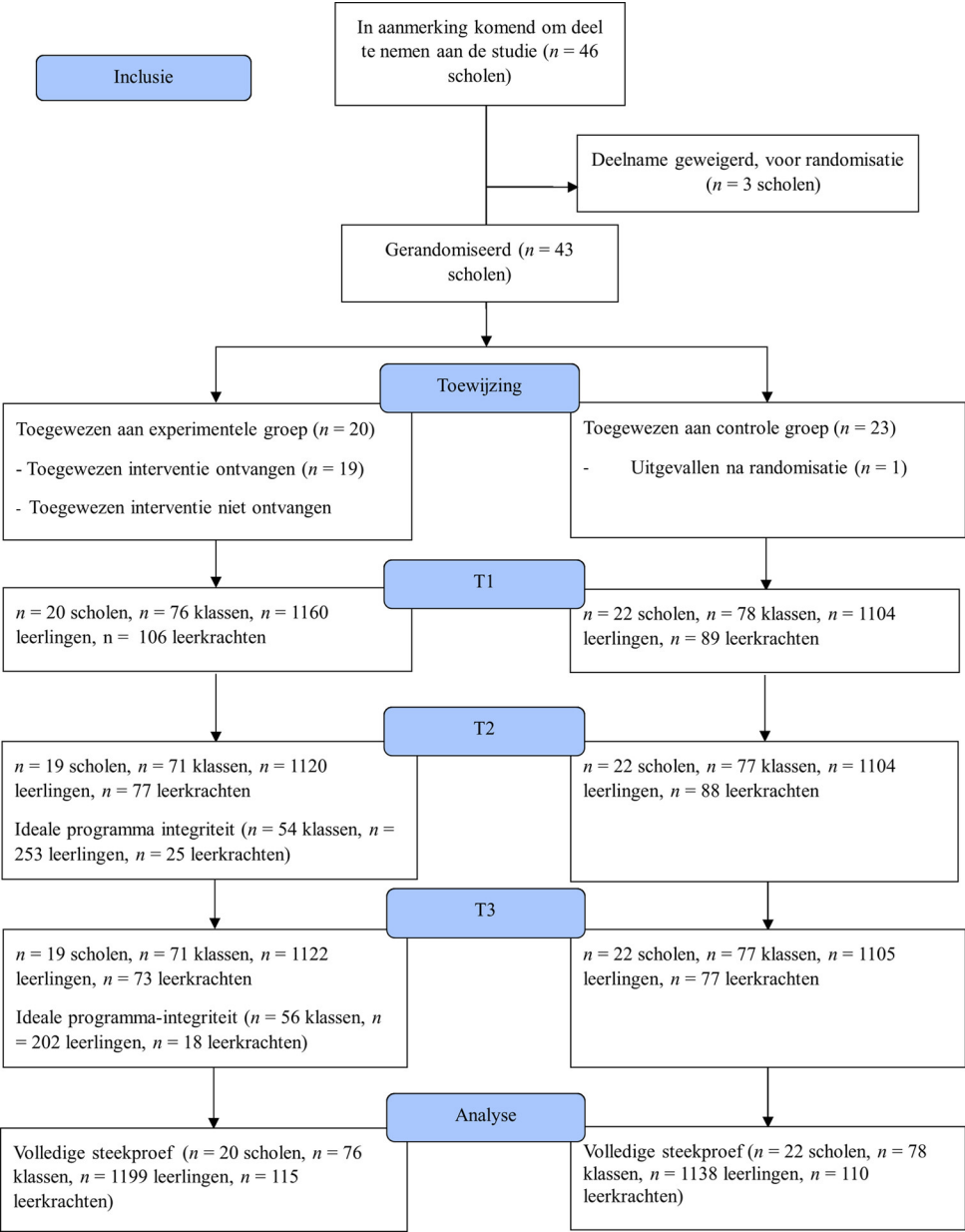
2 Methode

2.1 Procedure en Design

Vanaf januari 2023 werden meer dan vierhonderd scholen die het KiVa-programma (antipestprogramma) implementeerden (KiVa-scholen) benaderd om deel te nemen aan het huidige onderzoek, zowel persoonlijk, per email als telefonisch. We hebben ervoor gekozen om alleen KiVa-scholen te werven omdat SterkWerk is ontwikkeld door dezelfde organisatie als die het KiVa-programma heeft ontwikkeld. Ten eerste vanuit praktisch oogpunt: er zijn ruim 450 scholen die KiVa implementeerden en die daarvoor al in contact stonden met het implementatieteam, waardoor het makkelijker was om deze scholen te werven in de korte tijdsperiode die we hadden om scholen te werven. Ten tweede vanuit een theoretisch perspectief: de verwachting was dat een enigszins stabiele basis nodig was om dit programma te implementeren, en die zou er in principe moeten zijn bij deze scholen (net zoals bij andere scholen die effectieve antipestprogramma's implementeren). Omdat zowel interventie- als controlescholen KiVa implementeren, gaat het in deze studie om het effect van de huidige interventie bovenop het effect van KiVa. Scholen werden in april 2023 gerandomiseerd op basis van schoolgrootte (aantal deelnemende klassen) met behulp van een online randomisatietool (*Research Randomizer*, 2023) in een verhouding van 1:1. Er werd een cluster RCT uitgevoerd. Na randomisatie was er één controleschool uitgevallen vanwege externe omstandigheden (schoolreorganisatie). Figuur 2 geeft het stroomdiagram weer van de studie.

Vanaf mei 2023 verzamelden scholen toestemming voor deelname aan het onderzoek. Scholen informeerden ouders en leerlingen over het project via een brochure en informatiebrief, met papieren en/of digitale toestemmingsformulieren voor ouders, leerlingen (> 11 jaar) en leerkrachten/schoolleiders. Gemiddeld was de oudertoestemming 65 %, wat de verwachtingen overtrof op basis van eerdere studies die gemiddelde ouderlijke toestemmingpercentages tussen de 30 % en 60 % rapporteerden (Tigges, 2003; Wolfenden et al., 2009).

Het design bestond uit drie dataverzamelmomenten. Het eerste meetmoment vond plaats in september–oktober 2023 (T1), de tweede in februari–maart 2024 (T2) en het laatste meetmoment in mei–juni 2024 (T3). Na het eerste meetmoment haakte één interventieschool af (zowel met de implementatie als met deelname aan het onderzoek), omdat ze de tijdsinvestering in combinatie met andere ontwikkelingen binnen de school hadden onderschat. De huidige studie is geregistreerd in ClinicalTrials (NCT05891067) en de ethische commissie van de faculteit Sociale Wetenschappen van de Universiteit Utrecht heeft het design van de studie goedgekeurd (23-0082) (voor het onderzoeksprotocol, zie Van Loon & Kaufman, 2023).



Figuur 2 Stroomdiagram van het onderzoek

2.2 Participanten

In totaal waren 46 scholen geïnteresseerd in deelname aan het onderzoek, waarvan drie scholen vóór de randomisatie afhaakten. Scholen werden gerandomiseerd in een experimentele conditie ($N = 22$ scholen, $N = 76$ klassen) of een controleconditie ($N = 20$ scholen, $N = 78$ klassen). Van de totale steekproef haakte één controleschool af na de randomisatie en één interventieschool na het eerste meetmoment ($N = 4$ klassen, T₁). Er waren geen andere uitvallers op schoolniveau op T₂ of T₃.

Tabel 1 toont de demografische gegevens van de leerlingen ($N = 2337$ leerlingen van 6–13 jaar, $M_{\text{leeftijd}} = 9.55$, $SD = 1.21$; sekse jongen 48 %, gender 46 %; sekse meisje 48 %, gender 43 %; andere sekse 4 %, gender 12 %). De meeste leerlingen voelden zich verbonden met westerse landen ($N = 1932$, 88 %; bijvoorbeeld Nederland, Italië of Frankrijk), anderen met niet-westerse landen of een mix ($N = 264$, 12 %; bijvoorbeeld Turkije, Marokko of Nederland-Suriname). De deelname over leerjaren was ongeveer gelijk (met een iets groter aandeel in de hogere leerjaren).

Tabel 2 toont de demografische gegevens van leerkrachten ($N = 225$ leerkrachten van 21–65 jaar; $M_{\text{leeftijd}} = 38.7$, $SD = 11.1$; sekse vrouwelijk 77 %, gender 76 %, 1 % wil het liever niet zeggen). De meeste leerkrachten voelden zich verbonden met een westers land (voornamelijk Nederland, 97 %). De deelname over schooljaren was ongeveer gelijk, vergelijkbaar met de leerling rapportages. Deze studie volgde de *Consolidated Standards of Reporting Trials Guidelines* (CONSORT; Begg et al., 1996), zoals weergegeven in Figuur 2.

2.3 SterkWerk Programma

De Nederlandse versie van Meaningful Roles (SterkWerk) is een klassikaal programma dat wordt uitgevoerd door leerkrachten in de groepen 5–8 van het basisonderwijs (leerlingen tussen de 8–12 jaar oud). Leerkrachten van interventiescholen kregen tussen augustus en oktober 2023 twee trainingssessies (twee keer drie uur), verzorgd door getrainde en ervaren onderwijsprofessionals ($N = 4$, die ieder 3–6 scholen trainde). Tijdens de training kregen leerkrachten informatie over het programma (onder andere over het theoretisch kader) en hoe ze de componenten konden implementeren in de klas. Leerkrachten kregen bijvoorbeeld advies hoe ze verantwoordelijkheden (rollen) konden geven aan leerlingen in plaats van opdrachten (taken). Daarnaast werd de opzet van de klassenvergaderingen besproken en hoe deze het beste begeleid konden worden, evenals het stimuleren en zichtbaar maken van (het geven van) complimenten. Het eerste meetmoment vond zo vroeg mogelijk in het schooljaar plaats, vaak na de eerste leerkrachttraining (zodat de trainer de school nog kon herinneren aan de toestemmingsformulieren of andere stappen die genomen moesten worden) en vóór de tweede trainingssessie. Leerkrachten mochten pas beginnen met het implementeren van SterkWerk nadat de vragenlijsten waren afgenomen.

Tabel 1 Demografische gegevens op T1 en verschillen op T1 tussen de experimentele conditie (N = 1149) en controleconditie (N = 1084) voor leerlingen

	Experimentele conditie	Controleconditie	Verschillen op T1
	<i>M (SD)</i>	<i>M (SD)</i>	<i>t (df)</i>
Leeftijd (jaren) ^a	9.59 (1.20)	9.51 (1.22)	-1.643 (2227)
	<i>N (%)</i>	<i>N (%)</i>	χ^2 (<i>df</i>)
Sekse			2.258 (2)
Jongen	547 (47.6)	534 (49.3)	
Meisje	553 (48.1)	516 (47.6)	
Anders of weet ik niet	49 (4.3)	34 (3.1)	
Gender			1.360 (2)
Jongen	524 (45.6)	503 (46.4)	
Meisje	500 (43.5)	449 (41.5)	
Anders ^b	125 (10.9)	131 (12.1)	
Etnische identiteit			0.014 (1)
Westers	995 (87.9)	937 (88.1)	
Mix westers-niet westers – niet-westers	137 (12.1)	127 (11.9)	
Groep/leerjaar ^c			5.197 (3)
5	210 (18.3)	237 (21.9)	
6	287 (25.0)	276 (25.5)	
7	324 (28.3)	286 (26.4)	
8	325 (28.4)	284 (26.2)	

Noot. ^a N = 2229 leerlingen rapporteerden leeftijd (N = 1147 experimentele conditie, N = 1082 controleconditie).
^b Onder de categorie ‘anders’ vielen de keuzemogelijkheden: jongen en meisje, geen meisje of jongen, weet ik niet, wil ik niet zeggen. ^c N = drie leerlingen rapporteerden dat ze in groep/leerjaar 4 zitten, allen afkomstig uit de experimentele conditie. Deze leerlingen werden niet in deze analyses opgenomen.
*p < .05, **p < .01, ***p < .001

Met betrekking tot de rollen krijgen leerlingen specifieke rollen om actief te werken aan autonomie en verantwoordelijkheid, competentie en prosociaal gedrag. Deze rollen, zoals nieuwslezer of plantenverzorger, zijn prosociaal van aard, omdat leerlingen een of meer leerlingen helpen of bijdragen aan de hele klas, of rekening houden met andere leerlingen in de klas. Er wordt verwacht dat deze activiteit interactie stimuleert, en daarmee verbondenheid. Bovendien verwachten we dat de rollen leerlingen autonomie en verantwoordelijkheid geven, omdat de leerlingen vrij zijn om de rol te vervullen zoals zij dat willen. De rollen zijn erop gericht om de sterke punten en talenten van elk kind te optimaliseren, door leerlingen de ervaring te geven dat ze iets kunnen bijdragen dat hun vertrouwen in hun eigen kunnen vergroot.

Tijdens de wekelijkse klassenvergaderingen nemen leerlingen de leiding in discussies over het functioneren en de afspraken in de klas, ze bespreken hoe ze zich voelen en hoe

Tabel 2 Demografische en descriptieve gegevens op T1, en verschillen op T1 tussen de experimentele- (N = 115) en controleconditie (N = 110) voor leerkrachten

	Experimentele conditie	Controleconditie	Verschillen op T1
	<i>M (sd)</i>	<i>M (sd)</i>	<i>t (df)</i>
Leeftijd (jaren)	39.38 (10.75)	37.95 (11.45)	-0.905 (194)
	<i>N (%)</i>	<i>N (%)</i>	χ^2 (<i>df</i>)
Sekse			11.605 (2)**
Man	11 (12.5)	33 (32.7)	
Vrouw	76 (86.4)	68 (67.3)	
Wil ik liever niet zeggen	1 (1.1)	0 (0.0)	
Gender			10.972 (2)**
Man	11 (12.5)	32 (31.7)	
Vrouw	77 (87.5)	68 (67.3)	
Wil ik liever niet zeggen	0 (0.0)	1 (1.0)	
Etnische identiteit			2.626 (1)
Westers	87 (100.0)	98 (97.0)	
Mix westers-niet westers	0 (0.0)	3 (3.0)	
Groep/leerjaar			4.125 (2)
5-6	51 (44.3)	48 (43.6)	
7-8	58 (50.4)	48 (43.6)	
Anders (bijvoorbeeld 5-8, 6-8, 6-7)	6 (5.2)	14 (12.7)	

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

de sfeer in de klas is. De verwachting is dat dit zorgt voor interactie en verbinding met elkaar en het gevoel van competentie, autonomie en verantwoordelijkheid van leerlingen stimuleert in sociale klasprocessen.

Ten slotte leren leerlingen hoe ze effectieve complimenten kunnen geven en ontvangen en oefenen ze dit op een structurele basis, wat naar verwachting het vertrouwen in anderen en het vertrouwen in hun eigen kunnen versterkt. Door regelmatig en expliciet complimenten te geven met de nadruk op rollen en elkaars sterke punten, krijgen leerlingen waardering voor zichzelf en anderen en creëren ze een groepsnorm die prosociaal gedrag structureel versterkt. Het wordt aangeraden om de complimenten op een zichtbare plek te tonen in de klas, en aandacht te geven aan de complimenten tijdens de klassenvergaderingen.

2.4 Controleconditie

De controlegroep ontving geen materiaal of training met betrekking tot SterkWerk tijdens de implementatie van het programma in de experimentele conditie (*treatment-*

as-usual) en werd gevraagd om geen nieuw sociaal veiligheidsprogramma te starten of (onderdelen van) SterkWerk te implementeren. Dit betekent dat de controlescholen alleen het huidige sociaal veiligheidsprogramma gebruikten tijdens het onderzoek, wat in dit geval KiVa was (een antipestprogramma voor sociale veiligheid en positieve groepsvorming in de klas). Scholen in de controlegroep kregen de mogelijkheid om het materiaal en de training van SterkWerk te ontvangen nadat de dataverzameling was afgerond.

2.5 Instrumenten

Tabel B6 in Appendix B geeft een overzicht van de instrumenten op elk meetmoment. Informatie over prosociaal klasklimaat (prosociaal gedrag, prosociale groepsnorm, samenwerking en conflict in de klas), basispsychologische behoeften (autonomie, competentie, verbondenheid), welzijn en veiligheid op school, en antisociaal gedrag (pesten, slachtofferschap van pesten, leerkrachthouding ten opzichte van pesten) werd door leerlingen gerapporteerd. Daarnaast verstrekten leerkrachten informatie over hun houding ten opzichte van het stimuleren van de vervulling van basispsychologische behoeften en burgerschapsvaardigheden van leerlingen. Appendix A bevat de gedetailleerde informatie over de instrumenten, zoals de categorisering van demografische gegevens en validiteit van de vragenlijsten.

2.6 Statistische Analyses

Statistische analyses werden uitgevoerd met SPSS en Mplus. We hanteerden een *intention-to-treat* aanpak en namen alle deelnemers op in de analyses, inclusief deelnemers uit klassen die niet met het programma waren begonnen of het programma niet (goed) hadden geïmplementeerd. We volgden deze aanpak om motivatie als mogelijk *confounding* effect te verminderen en de effectiviteit van de toegewezen interventie te onderzoeken (Montori & Guyatt, 2001). Hoewel het programma op klassenniveau werd geïmplementeerd, varieerden de individuele ervaringen en betrokkenheid aanzienlijk (bijvoorbeeld of leerlingen een rol hadden), daarom zijn de analyses op individueel niveau uitgevoerd. Om te controleren of de randomisatie succesvol was, werden voor de analyses mogelijke baseline (T1) verschillen in demografische gegevens en uitkomstvariabelen (baseline meetmoment) tussen leerlingen in de interventie- en controleconditie gecontroleerd met onafhankelijke *t*-toetsen of ANOVA-toetsen (voor continue variabelen) en χ^2 analyses (voor categorische variabelen). Leerlingen die geen toestemming gaven of geen toestemming kregen van hun ouders om mee te doen aan het onderzoek hebben geen enkele onderzoeksvragenlijst ingevuld. We hebben dus geen gegevens over deze leerlingen, en kunnen geen analyses uitvoeren om te onderzoeken of deze leerlingen verschilden van deelnemende leerlingen.

2.6.1 Effectiviteit van het Programma in de Volledige Steekproef

Om de effectiviteit van het programma te onderzoeken (doel 1), werden regressieanalyses uitgevoerd in Mplus. De afhankelijke variabelen waren de uitkomsten halverwege het programma (T₂) of aan het einde van het schooljaar (T₃), conditie was de voorspellende variabele (1 = interventie, 0 = controle), en baseline levels (T₁) van de uitkomstmaten, potentiële demografische verschillen tussen de experimentele en controleconditie op T₁, en/of T₂ metingen (wanneer de afhankelijke variabelen de uitkomstmaten op T₃ waren) waren de covariaten. Voor primaire en secundaire uitkomsten op T₂ en T₃ werden aparte modellen gedraaid (in totaal vier modellen), zodat de onderlinge samenhang tussen effecten werd meegenomen. De covariantie van conditie en de specifieke variabele(n) werden ook in de modellen opgenomen. Het toevoegen van de covariantie zorgt voor een meer realistische modellering van de data en een betere schatting van de regressiecoëfficiënten (vanwege de afhankelijkheid van de voorspellers). In de regressieanalyses werd gecorrigeerd voor clustering binnen scholen ('TYPE = COMPLEX') en *Robust Maximum Likelihood* (MLR) werd gebruikt als *estimator*. De MLR is robuust tegen (mogelijke) afwijkingen van normaliteit. Daarnaast is er gebruik gemaakt van *Full Information Maximum Likelihood* (FIML), een methode die rekening houdt met missende waarden waarbij alle beschikbare gegevens worden benut. Effectgroottes (Cohen's *d*) werden berekend door middel van een online calculator (Lenhard & Lenhard, 2022), op basis van de gestandaardiseerde bèta's. Een positieve Cohen's *d* weerspiegelt verbeteringen ten gunste van de interventieconditie ten opzichte van de controleconditie. Kleine effectgroottes werden beschouwd als $d = 0.20$, middelmatige effectgroottes als $d = 0.50$ en grote effectgroottes als $d = 0.80$ (Cohen, 1988).

2.6.2 Effectiviteit van het Programma voor Ideale Programmatrouw

Voor het tweede doel hebben we analyses uitgevoerd om de impact van programmatrouw op de effectiviteit van het programma te bepalen. We hebben de regressieanalyses herhaald in een subgroep van leerlingen met een ideale programmatrouw. Om dit te bepalen, gebruikten we zelfrapportagegegevens van leerlingen over de drie belangrijkste programmacomponenten, namelijk betekenisvolle rollen, klassenbijeenkomsten en complimenten. Deze gegevens werden gecombineerd om één variabele te creëren voor elk meetmoment (T₂ en T₃). Voorwaarden waren dat leerlingen in de afgelopen maanden minstens één rol hadden gehad en minstens één keer per maand bezig waren met de rol(len) (T₂ $N = 751$, 68%; T₃ $N = 679$, 61%), minstens één klassenbijeenkomst per week hadden (T₂ $N = 560$, 51%; T₃ $N = 473$, 43%) en minstens één compliment per week ontvingen (T₂ $N = 657$, 59%; T₃ $N = 655$, 59%), zie Tabel B1 in Appendix B voor de beschrijvende statistieken. Deze subgroep (hoge trouw) omvatte alleen leerlingen die aan alle drie de voorwaarden voldeden (T₂ $N = 253$, 23%; T₃ $N = 202$, 18%). Analyses werden ook herhaald voor de subgroep met lage trouw. De betrouwbaarheidsintervallen van beide groepen (hoge en lage trouw) werden met elkaar vergeleken om te bepalen of de subgroepen van elkaar verschilden. Als de betrouwbaarheidsintervallen van beide

groepen met elkaar overlappen, is er geen (significant) verschil tussen de groepen ('nee'); als de betrouwbaarheidsintervallen wel met elkaar overlappen is er een (significant) verschil ('ja').

Regressieanalyses werden herhaald in een subgroep van leerkrachten met een ideale programmatrouw, met de volgende voorwaarden: leerkrachten gaven aan dat leerlingen de rol(len) (zeer) zelfstandig konden uitvoeren ($T_2 N = 56, 74\%$; $T_3 N = 53, 77\%$), minstens één klassenbijeenkomst per week hadden ($T_2 N = 33, 44\%$; $T_3 N = 25, 36\%$) en dat leerkrachten aangaven dat ze aandacht besteedden aan complimenten en ervoor zorgden dat complimenten minstens één keer per week werden gegeven ($T_2 N = 67, 92\%$; $T_3 N = 63, 93\%$), zie Tabel B2 in Appendix B voor de beschrijvende statistieken. Deze subgroep (hoge trouw) bestond alleen uit docenten die aan alle drie de voorwaarden voldeden ($T_2 N = 25, 34\%$; $T_3 N = 18, 26\%$). Analyses werden ook herhaald voor de subgroep met lage trouw, en beide groepen werden met elkaar vergeleken door te kijken naar de betrouwbaarheidsintervallen.

3 Resultaten

3.1 Randomisatie, Descriptieve Waarden en Drop-outs

Wat betreft de zelfrapportage van leerlingen waren er op T_1 geen verschillen tussen de experimentele en de controleconditie voor demografische gegevens (zie Tabel 1). Voor de studievariabelen waren er op T_1 verschillen tussen de experimentele en controleconditie wat betreft prosociaal gedrag (individueel en samenwerking in de klas), veiligheid op school, slachtofferschap van pesten, en leerkrachthouding ten opzichte van het bespreken van pesten (Tabel 3). De controleconditie scoorde hoger op prosociaal gedrag en slachtofferschap van pesten, terwijl de experimentele conditie hoger scoorde op veiligheid op school en leerkrachthouding ten opzichte van het bespreken van pesten. De waarden voor scheefheid en gepiektheid waren voor de meeste variabelen acceptabel, behalve voor pestgedrag en slachtofferschap van pesten. We verwachtten echter niet dat pesten en slachtofferschap van pesten normaal verdeeld is, aangezien dit niet vaak in de populatie voorkomt (bijvoorbeeld Demaray et al., 2016; Romera et al., 2021; Tabel 3). Bovendien verwachtten we geen problemen met de analyses, aangezien we gebruik maken van MLR. Er waren enkele verschillen op T_1 tussen deelnemers en drop-outs op T_2 en T_3 . Drop-outs hadden vaker een niet-westerse etnische identiteit, kwamen vaker uit lagere leerjaren (Tabel B3 in Appendix B) en rapporteerden een lager niveau van samenwerking in de klas en een hoger niveau van leerkrachthouding over de effectiviteit om pesten te verminderen (Tabel B4 in Appendix B). Er waren geen andere verschillen op T_1 tussen de experimentele en controleconditie en tussen deelnemers en drop-outs met betrekking tot demografische gegevens en studievariabelen (zie Tabellen B3 en B4 in Appendix B).

Tabel 3 Descriptieve statistieken, verschillen op t1 tussen de experimentele- (N = 1160) en controleconditie (N = 1105), en regressieanalyses voor leerlingen

	Experimentele conditie ^a						Verschillen T1			Regressie ^b						
	T1		T2		T3		T1	T2	T3	Uitkomst op T2			Uitkomst op T3			
	M (sd)	M (sd)	M (sd)	M (sd)	M (sd)	M (sd)				M (sd)	β	d	B (se)	β	d	B (se)
Primaire uitkomsten																
Autonomie	2.96 (0.52)	2.94 (0.53)	2.93 (0.55)	2.94 (0.54)	2.92 (0.54)	2.94 (0.55)	-1.00 (2252)	-0.22	0.07	0.01	0.01	0.12	-0.02	-0.02	(0.03)	0.14
Competentie	2.59 (0.42)	2.59 (0.43)	2.59 (0.43)	2.57 (0.42)	2.58 (0.43)	2.57 (0.45)	-1.11 (2254)	-1.18	1.11	-0.01	-0.01	-0.01	0.01	0.01	(0.02)	-0.02
Verbondenheid	4.14 (0.72)	4.05 (0.75)	4.04 (0.76)	4.13 (0.75)	4.06 (0.79)	4.08 (0.78)	-0.32 (2252)	-1.04	1.20	-0.01	-0.02	-0.03	-0.02	-0.02	(0.03)	0.13
Intrinsieke (prosociale) motivatie	2.65 (0.38)	2.65 (0.38)	2.64 (0.38)	2.66 (0.38)	2.64 (0.39)	2.63 (0.42)	0.80 (2256)	-1.23	1.35	0.03	0.02	0.15	0.01	0.01	(0.03)	-0.02
Prosociaal gedrag (individueel)	2.80 (0.51)	2.85 (0.50)	2.84 (0.52)	2.88 (0.51)	2.86 (0.52)	2.88 (0.53)	3.45 (2258)***	-0.09	-0.10	0.02	0.02	0.14	-0.02	-0.02	(0.03)	0.13
Prosociaal gedrag: injunctieve norm	2.48 (0.49)	2.41 (0.53)	2.36 (0.57)	2.50 (0.51)	2.40 (0.53)	2.37 (0.58)	0.88 (2262)	-0.97	0.44	0.02	0.02	0.14	-0.01	-0.01	(0.03)	0.12
Prosociaal gedrag: samenwerking	3.92 (0.59)	3.88 (0.59)	3.87 (0.65)	3.98 (0.62)	3.92 (0.64)	3.95 (0.63)	2.42 (2257)*	-0.42	0.77	-0.02	-0.02	-0.03	-0.04	-0.04	(0.03)	0.15
Prosociaal gedrag: conflict	2.11 (0.85)	2.17 (0.86)	2.18 (0.84)	2.17 (0.87)	2.16 (0.86)	2.13 (0.86)	1.79 (2258)	0.51	-0.10	0.02	0.03	0.14	0.03	0.06	(0.05)	-0.06
Secundaire uitkomsten																
Schoolwelzijn	3.18 (0.60)	3.12 (0.62)	3.11 (0.63)	3.15 (0.65)	3.10 (0.65)	3.11 (0.64)	-1.24 (2247)	-0.68	0.18	0.00	0.00	0.00	-0.01	-0.01	(0.03)	-0.02

Tabel 3 Descriptieve statistieken, verschillen op t1 tussen de experimentele- en controleconditie, en regressieanalyses (vervolg)

	Experimentele conditie ^a						Verschillen T1		Regressie ^b					
	T1	T2	T3	T1	T2	T3	M (sd)	t (df)	Uitkomst op T2			Uitkomst op T3		
	M (sd)	M (sd)	M (sd)	M (sd)	M (sd)	M (sd)			γ_1	γ_2	β	B (se)	β	B (se)
Veiligheid op school	4.06 (0.89)	4.08 (0.84)	4.08 (0.84)	3.98 (0.97)	4.01 (0.92)	4.05 (0.88)		-2.02 (2248)*	-1.18	1.70	0.02	0.04 (0.05)	0.15	-0.00 (0.06)
Pestgedrag	1.19 (0.61)	1.18 (0.58)	1.18 (0.59)	1.24 (0.68)	1.20 (0.65)	1.18 (0.59)		1.78 (2172)	3.83	16.41	-0.01	-0.02 (0.02)	-0.03	0.02 (0.03)
Slachtofferschap van pesten	1.46 (0.99)	1.47 (0.95)	1.39 (0.84)	1.62 (1.10)	1.56 (1.06)	1.47 (0.95)		3.62 (2195)***	2.17	3.93	-0.02	-0.04 (0.05)	-0.04	-0.02 (0.04)
Slachtofferschap van pesten (max)	1.78 (1.19)	1.77 (1.17)	1.62 (1.06)	1.96 (1.28)	1.87 (1.26)	1.78 (1.19)		3.38 (2244)***	3.05	10.68	-0.01	-0.03 (0.08)	-0.03	-0.05* (0.05)
Leerkrachthouding: bespreken pesten	2.61 (1.19)	2.73 (1.15)	2.60 (1.20)	2.44 (1.21)	2.66 (1.17)	2.45 (1.26)		-3.33 (2234)***	0.34	-0.68	0.01	0.02 (0.11)	0.12	0.04 (0.09)
Leerkrachthouding: houding pesten	4.48 (0.95)	4.61 (0.76)	4.63 (0.76)	4.46 (0.95)	4.56 (0.84)	4.56 (0.84)		-0.54 (2235)	-2.01	3.80	0.03	0.04 (0.04)	0.14	0.04 (0.04)
Leerkrachthouding: effectiviteit verminderen pesten	3.53 (1.04)	3.34 (1.01)	3.20 (1.09)	3.45 (1.09)	3.41 (1.07)	3.24 (1.14)		1.91 (2209)	-0.38	-0.42	-0.05	-0.10 (0.06)	-0.10	-0.01 (0.06)

Noot. Aantal deelnemers varieert per variabele en tijdstipmoment (experimentele conditie N = 113-116; controleconditie N = 1086-1105). γ_1 = scheefheid op T1 en, γ_2 = gepiektheid op T1 voor de gehele steekproef.

^aConditie is gecodeerd als controlegroep = 0 en experimentele groep = 1. ^bGecorrigeerd voor clustering op schoolniveau en voor onderlinge samenhang tussen effecten.

* p < .05, ** p < .01, *** p < .001.

Wat betreft de zelfrapportage van leerkrachten waren er op T1 verschillen tussen de experimentele en controleconditie voor geslacht en gender (zie Tabel 2). Er waren minder vrouwen in de controleconditie (zowel voor geslacht als gender). De waarden voor scheefheid en gepiektheid waren voor alle variabelen acceptabel (zie Tabel 4). Er waren op T1 geen verschillen tussen deelnemers en drop-outs voor de studievariabelen op T2 en T3, alleen wat betreft geslacht en conditie. Drop-outs waren vaker mannen en kwamen vaker uit de experimentele conditie (zie Tabel B5 in Appendix B).

3.2 Effectiviteit van het Programma in de Volledige Steekproef

Ten eerste werd onderzocht in hoeverre er effecten waren in de volledige steekproef, ongeacht de mate van programmatrouw. Resultaten over de covariaten zijn gepresenteerd in Appendix C. Er waren geen significante effecten op de uitkomstmaten, zowel voor primaire als secundaire uitkomsten (zie Tabel 3), behalve op slachtofferschap van pesten (maximale score) op T3. Omdat de effectgrootte zeer laag was ($d = .09$), interpreteren we dit als geen betekenisvol effect.

Tabel 4 presenteert de resultaten van de regressieanalyses voor leerkrachten. De resultaten toonden aan dat er geen significant effect was voor burgerschapsvaardigheden van leerlingen, zowel op T2 als op T3. Voor leerkrachthouding over de stimulatie van basisbehoeften op T3 werden significante effecten van het programma gevonden voor het stimuleren van pro sociaal gedrag en autonomie. Leerkrachten zijn verbeterd in het belangrijk vinden en stimuleren van pro sociaal gedrag en autonomie bij leerlingen. Dit effect werd niet gevonden op T2. De effecten waren klein tot middelmatig (Cohen, 1988).

3.3 Effectiviteit van het Programma bij Ideale Programmatrouw

Bij ideale programmatrouw werden effecten gevonden op de primaire uitkomsten basispsychologische behoeften (autonomie, competentie, en verbondenheid), intrinsieke (prosociale) motivatie, en pro sociaal gedrag (individueel gedrag, en samenwerking en conflict in de klas; Tabel B7B in Appendix B). De meeste van deze effecten waren kortdurend (op T2, niet blijvend tot T3), met uitzondering van competentie en verbondenheid, waarvoor de effecten zowel op T2 als op T3 werden gevonden (Tabel B8B in Appendix B). Er werd geen significant effect gevonden op de groepsnorm (injunctieve norm), op zowel T2 als T3 (Tabellen B7B en B8B in Appendix B). De effecten voor alle uitkomsten waren klein (Cohen, 1988). Voor de primaire uitkomsten op T2 werden verschillen gevonden tussen hoge en lage trouw voor autonomie, verbondenheid, intrinsieke (prosociale) motivatie, individueel pro sociaal gedrag en samenwerking in de klas, maar niet voor competentie en conflict in de klas (Tabel B7B in Appendix B). Op T3 werden alleen verschillen gevonden tussen hoge en lage trouw voor verbondenheid, niet voor competentie (Tabel B8B in Appendix B).

Tabel 4 Verschillen op T1 tussen de experimentele- (N = 100) en controleconditie (N = 90), en regressieanalyses op T2 en T3 voor leerkrachtge-rapporteerde uitkomsten

	Experimentele conditie ^a						Verschillen T1		Regressie ^b					
	T1		T2		T3		Y ₁	Y ₂	Uitkomst op T2		d	Uitkomst op T3		d
	M (sd)	M (sd)	M (sd)	M (sd)	M (sd)	M (sd)			β	B (se)		β	B (se)	
Leerkrachthouding: pro sociaal gedrag	7.79 (0.68)	7.86 (0.79)	8.15 (0.73)	8.00 (0.86)	8.07 (0.72)	8.07 (0.74)	-0.56 (188)	1.10	-0.13 (0.127)	-0.181 (0.127)	-0.25	0.14* (0.09)	0.20 (0.09)	0.39
Leerkrachthouding: autonomie	7.61 (0.86)	7.86 (0.85)	8.23 (0.79)	8.05 (0.87)	8.04 (0.73)	8.12 (0.71)	-0.46 (188)***	0.97	0.03 (0.114)	0.042 (0.114)	0.16	0.14* (0.09)	0.20 (0.09)	0.38
Leerkrachthouding: competentie	8.02 (0.82)	8.02 (0.81)	8.29 (0.81)	8.16 (0.91)	8.11 (0.76)	8.30 (0.69)	1.11 (185)	0.95	-0.06 (0.101)	-0.092 (0.101)	-0.12	0.06 (0.07)	0.09 (0.07)	0.23
Leerkrachthouding: verbondenheid	8.25 (0.75)	8.25 (0.81)	8.38 (0.74)	8.43 (0.81)	8.33 (0.73)	8.44 (0.71)	1.54 (185)	-0.11	-0.06 (0.110)	-0.087 (0.110)	-0.12	-0.02 (0.10)	-0.03 (0.10)	-0.04
Burgerschapsvaardigheden	3.03 (0.39)	3.06 (0.41)	3.09 (0.39)	3.02 (0.35)	3.05 (0.33)	3.03 (0.36)	-0.18 (168)	-0.05	0.00	0.001	0.10	0.09	0.07 (0.050)	0.28

Noot. Aantal deelnemers varieert per variabele en tijdstipmoment (experimentele conditie N = 73–100, controleconditie N = 77–90). Gender is geïncludeerd als covariaat in alle analyses (alleen categorieën man en vrouw). Y₁ = schreefheid op T1 en Y₂ = gepiektheid op T1 voor de gehele steekproef.

^aConditie is gecodeerd als controlegroep = 0 en experimentele groep = 1. ^bGecorrigeerd voor clustering op schoolniveau en voor onderlinge samenhang tussen effecten.

*p < .05, **p < .01, ***p < .001.

Op secundaire uitkomsten werden effecten gevonden voor schoolwelzijn en veiligheid op school, evenals slachtofferschap van pesten (gemiddelde en maximum) en houding van de leerkracht tegenover pesten (Tabel B7B in Appendix B). De meeste van deze effecten waren kortdurend (op T2, niet blijvend tot T3), met uitzondering van slachtofferschap van pesten en houding van de leerkracht tegenover pesten, waarvoor de effecten zowel op T2 als T3 werden gevonden. Er is een effect gevonden die wel op T3 maar niet op T2 aanwezig was, namelijk het effect voor de leerkrachthouding over het bespreken van pesten (Tabel B8A in Appendix B). De effecten voor alle uitkomsten waren klein (Cohen, 1988). Er werden geen significante effecten gevonden voor pestgedrag en leerkrachthouding over de effectiviteit van het verminderen van pesten, op zowel T2 als T3 (Tabellen B7B en B8B in Appendix B). Voor de secundaire uitkomsten op T2 werden alleen verschillen gevonden tussen hoge en lage trouw voor schoolwelzijn en veiligheid op school, niet voor slachtofferschap van pesten en houding van de leerkracht tegenover pesten (Tabel B7B in Appendix B). Op T3 werden alleen verschillen gevonden tussen hoge en lage trouw voor slachtofferschap van pesten, niet voor de leerkrachthouding over pesten (Tabel B8B in Appendix B).

Kortom, leerlingen die aangaven SterkWerk te hebben ontvangen zoals bedoeld voelden zich – vergeleken met kinderen in de controlegroep – op korte termijn autonomer, waren gemotiveerder om zich prosociaal te gedragen en vertoonden ook vaker dit gedrag. Daarnaast voelden ze zich op korte termijn veiliger en gelukkiger op school, en voelden ze zich zowel op korte termijn als enkele maanden later competent en meer verbonden met anderen. Op langere termijn ervoeren deze leerlingen dat ze minder gepest werden en dat de leerkracht pesten meer afkeurde.

Wat betreft de zelfrapportage van leerkrachten, werden er geen positieve effecten gevonden voor de groep leerkrachten die aangaven dat het programma werd geïmplementeerd zoals bedoeld (Tabellen B9 en B10A in Appendix B). Er waren ook geen verschillen tussen hoge en lage trouw (Tabellen B9 en B10B in Appendix B).

4 Discussie

Een positief, prosociaal klasklimaat is van cruciaal belang om een effectieve leeromgeving te optimaliseren en het school- en psychologische welzijn van kinderen te bevorderen (Evans et al., 2009; Wang et al., 2020). Schoolinterventies kunnen een ideale aanpak zijn om dit te bereiken, maar er is een gebrek aan kennis over de effectiviteit van programma's gebaseerd op de theorie van de vervulling van basispsychologische behoeften en (cluster) RCT-designs om dergelijke effecten te evalueren. Om deze hiaten aan te pakken, onderzocht de huidige studie de effectiviteit van SterkWerk – een schoolinterventie gericht op het bevorderen van een prosociaal klasklimaat in basisscholen door de intrinsieke prosociale motivatie van leerlingen te versterken door het vervullen van de drie basispsychologische behoeften (autonomie, competentie en verbondenheid).

Onze bevindingen toonden aan dat er voor de totale steekproef geen significante effecten waren voor zowel primaire als secundaire uitkomstenmaten halverwege en aan het einde van het schooljaar (T2 en T3). Alleen voor een kleine subgroep van leerlingen die aangaven dat het programma optimaal was geïmplementeerd (ongeveer een kwart van de gehele steekproef), werden gemiddeld positievere effecten gevonden ten opzichte van de controlegroep met betrekking tot de primaire uitkomsten basispsychologische behoeften en een prosociaal klasklimaat, en de secundaire uitkomsten welzijn en veiligheid op school, en antisociaal gedrag. Alle effecten waren klein en kortdurend (voornamelijk halverwege het schooljaar (T2)). Leerkrachten die SterkWerk hadden geïmplementeerd rapporteerden een positievere (eigen) houding ten opzichte van het stimuleren van prosociaal gedrag en autonomie bij leerlingen, alleen aan het einde van het schooljaar (T3). Wanneer leerkrachten aangaven dat ze de interventie hadden uitgevoerd zoals bedoeld, werden er geen significante effecten gevonden.

Er werden geen positieve effecten voor leerlingen gevonden in de gehele steekproef (doel 1), wat tegengesteld is aan bevindingen van soortgelijke interventies (bijvoorbeeld Manzano-Sánchez & Valero-Valenzuela, 2019; Mesurado et al., 2018; Sklad et al., 2012). Het kan zijn dat één of meerdere componenten van de interventie niet effectief waren, of dat de vervulling van basispsychologische behoeften niet het werkzame mechanisme van het programma is. Een andere verklaring voor het niet vinden van positieve effecten kan zijn dat er meer tijd nodig is voor de implementatie. Een schooljaar is mogelijk niet genoeg om effecten te vinden voor een aanpak dat onderdeel moet worden van de schoolcultuur (ten opzichte van interventies met een bepaald aantal sessies of onderdelen). Eerder onderzoek laat zien dat effecten van programma's groter zijn naarmate een programma langer loopt (Huitsing et al., 2020; Limber et al., 2018). Een aanbeveling is daarom ook om in toekomstig onderzoek de evaluatie over een langere periode dan een schooljaar uit te voeren. De nulresultaten kunnen ook het gevolg zijn van de lage variabiliteit van de scores voor de variabelen (onder andere individueel prosociaal gedrag) of een plafond-effect (relatief hoge gemiddelde scores), wat mogelijk het vermogen om effecten te detecteren heeft beperkt (bijvoorbeeld Kloczek et al., 2024). Het feit dat het onderzoek alleen op KiVa-scholen is uitgevoerd kan hierbij ook een rol hebben gespeeld, omdat dit scholen zijn die waarschijnlijk al relatief hoog scoren op een positief klasklimaat. Als laatste is een mogelijke reden dat de interventie niet geïmplementeerd is zoals bedoeld, wat een belangrijke voorwaarde kan zijn voor de effectiviteit van een interventie (bijvoorbeeld Durlak & DuPre, 2008), wat we ook hebben gezien in onze resultaten (dat wil zeggen alleen effecten bij leerlingen met optimale programmatrouw, dat wil zeggen leerlingen die zelf aangaven dat ze een rol hadden gehad en hier minimaal maandelijks mee bezig waren geweest, wekelijks een compliment ontvingen en wekelijks een vergadering hadden). Verschillende factoren kunnen hier van invloed op zijn, gerelateerd aan bijvoorbeeld de aanbieders (bijvoorbeeld competentie, bekwaamheid of motivatie van de leerkracht), aspecten van het ondersteuningssysteem (bijvoorbeeld organisatorisch functioneren van de school, steun van collega's, hoge werkdruk), of de ontvankelijkheid

van het programma (bijvoorbeeld behoefte of motivatie van leerlingen) (bijvoorbeeld Combs et al., 2022; Rojas-Andrade & Bahamondes, 2019; Waller et al., 2017). Vervolgonderzoek is nodig om uit te zoeken welke factoren belangrijk zijn en wat er nodig is voor optimale implementatie.

Onze bevindingen benadrukken de rol van programmatrouw bij het interpreteren van programma-effectiviteit (doel 2). Ondanks dat de effecten klein waren en er bij sommige effecten vooral een (sterkere) verslechtering was in de controlegroep ten opzichte van de interventiegroep, zijn de positieve effecten in de subgroep van leerlingen met hoge programmatrouw een belangrijke aanwijzing over het belang van optimale programma-implementatie. Dat programmatrouw belangrijk is voor programma-effectiviteit komt overeen met eerder onderzoek naar programma's die gericht zijn op het bevorderen van sociaal en emotioneel leren en het verminderen van antisociaal gedrag (Durlak et al., 2011; Durlak & DuPre, 2008; Haataja et al., 2014). Onze resultaten wijzen erop dat dit ook geldt voor programma's die zich richten op basispsychologische behoeften; om prosociale klassen te bevorderen, moeten alle componenten frequent en consistent worden uitgevoerd. In het geval van SterkWerk betekent dit dat leerlingen verantwoordelijkheid en autonomie krijgen door middel van zinvolle opdrachten (klassenrollen), dat klassenbijeenkomsten worden gefaciliteerd om de groepsdynamiek en -sfeer te bespreken, en dat leerlingen worden aangemoedigd elkaars sterke punten en positief gedrag te herkennen en waarderen door middel van gestructureerde complimenten.

Opvallend werden de meeste effecten voor programmatrouw bij leerlingen op korte termijn gevonden, halverwege het programma, en niet aan het einde van het schooljaar. De uitdoving van deze effecten kan te wijten zijn aan motivatieproblemen tijdens de implementatie, moeite om aandacht te blijven geven aan de componenten, of druk van andere taken die in de tweede helft van het schooljaar verplicht zijn, zoals toetsen en vakanties – zelfs in de subgroep met hoge programmatrouw. In de eerste helft van het jaar waren leerkrachten en leerlingen waarschijnlijk het meest gemotiveerd en besteedden ze de meeste aandacht aan de implementatie, omdat ze recent de training hadden gevolgd. Dit komt overeen met de gerapporteerde programmatrouw, waarbij minder leerlingen aangaven rollen te hebben gehad en bezig te zijn geweest met de rol(len) (T2 68 %, T3 61 %) en minder leerkrachten en leerlingen aangaven wekelijkse klassenvergaderingen te hebben gehad (leerlingen: T2 51 %, T3 45 %, leerkrachten: T2 44 %, T3 36 %) op T3 vergeleken met T2. Aan de andere kant hadden de meeste effecten op T3 betrekking op de secundaire uitkomst pesten (slachtofferschap, leerkrachthouding), wat niet de primaire focus van het programma was. Effecten op negatief, hardnekkig gedrag, zoals pesten, kunnen langer duren om te ontwikkelen, aangezien deze processen zich voornamelijk ontwikkelen via complexe groepsprocessen (Thornberg, 2011). Het zou kunnen dat langere blootstelling aan het programma voornamelijk verband houdt met effecten op secundaire uitkomsten. Ten slotte is het ook mogelijk dat er minder effecten op T3 werden gevonden vanwege een lagere power (op T3 was de steekproef kleiner dan op T2, voor zowel leerlingen als leerkrachten).

De huidige interventie is gekaderd rondom de basispsychologische behoeften (Ryan & Deci, 2017) en de goal-framing theorie (Lindenberg & Steg, 2007), waarbij het creëren van een pro sociaal klasklimaat intrinsieke motivatie vereist, wat gestimuleerd kan worden door het vervullen van de drie basispsychologische behoeften. Om te zorgen dat pro sociaal gedrag een blijvend onderdeel wordt van het klasklimaat, is het van belang dat dit de sociale norm in de klas wordt en blijft (Cheon et al., 2019; Lindenberg & Steg, 2007; Ryan & Deci, 2000, 2017). Aangezien er geen effect van het programma is gevonden voor de injunctieve (groeps)norm in de klas, zou dit een verklaring kunnen zijn voor de beperkte effecten in het algemeen en voor de uitdovende effecten bij programma-trouw (minder effecten aan het einde van het jaar ten opzichte van halverwege het jaar).

4.1 Beperkingen en Sterke Punten

Er zijn verschillende beperkingen van de huidige studie waarmee rekening gehouden moet worden. Hoewel we een hoger responspercentage hadden dan in eerder onderzoek (Tigges, 2003; Wolfenden et al., 2009), nam ongeveer een derde van de in aanmerking komende niet deel aan het onderzoek. Eerdere studies lieten *biased* steekproeven zien bij het gebruik van actieve toestemming in vergelijking met passieve toestemming, zoals meer vrouwelijke en jongere deelnemers, minder etnisch diverse deelnemers en minder probleemgedrag (Courser et al., 2009; Liu et al., 2017; Pokorny et al., 2001). Een andere mogelijke *bias* van de steekproef is dat het onderzoek alleen is uitgevoerd op KiVa-scholen: scholen die werken met KiVa als sociaal veiligheidsprogramma. SterkWerk kan worden gezien als extra of toegevoegd onderdeel van een sociaal veiligheidsprogramma. Ondanks dat SterkWerk is bedoeld voor deze populatie, is het onduidelijk of onze resultaten generaliseerbaar zijn naar niet-KiVa-scholen (die werken met andere sociale veiligheidsprogramma's dan KiVa). Omdat in Nederland alle scholen een veiligheidsplan moeten opstellen om pesten tegen te gaan (Wet veiligheid op school), en er meerdere antipestprogramma's in Nederland effectief zijn bevonden (onder andere KiVa, Kanjertraining, PRIMA; Huitsing et al., 2020; Nederlands Jeugdinstituut, 2025; Orobio De Castro et al., 2018; Van Verseveld et al., 2021; Vermande, 2015)), is de verwachting dat de meeste scholen vergelijkbaar zijn met KiVa-scholen. Toekomstig onderzoek zou moeten uitwijzen of SterkWerk ook positieve effecten heeft op scholen die andere sociale veiligheidsprogramma's implementeren. Ondanks deze mogelijke *biases* waren de deelnemende interventie- en controlescholen verspreid over heel Nederland, zowel in kleinere dorpen als steden, wat suggereert dat de steekproef representatief is voor een groot deel van de Nederlandse bevolking.

Daarnaast waren er verschillen op baseline tussen de experimentele en controleconditie en tussen deelnemers en drop-outs, zowel onder leerlingen als onder leerkrachten. Hoewel we in de analyses rekening hielden met de verschillen op baseline tussen de experimentele en controleconditie, suggereren deze verschillen dat de randomisatie niet

geheel succesvol was en dat de interne validiteit van de studie mogelijk wat verzwakt is. Voor de analyses met ideale programmatrouw rapporteerden leerlingen met hoge programmatrouw over het algemeen hogere en positievere scores dan leerlingen met lage programmatrouw of leerlingen in de controlegroep op baseline. Het zou kunnen dat de positieve effecten te verklaren zijn door andere aspecten dan de interventie, bijvoorbeeld de specifieke karakteristieken van leerlingen, zoals meer intrinsieke (prosociale) motivatie of gedrag. Dit is in lijn met eerder onderzoek dat laat zien dat hogere niveaus van problemen aan de start van een (angst)interventie geassocieerd zijn met sterkere effecten van het programma (bijvoorbeeld Stjerneklar et al., 2019). We hebben in deze studie programmatrouw per leerling geanalyseerd, niet op klassenniveau, waardoor het onbekend is of er robuuste effecten waren in klassen waar het programma trouw geïmplementeerd werd. Echter leverde dit wel interessante bevindingen op, zoals dat de aparte onderdelen redelijk waren uitgevoerd, maar dat de combinatie van deze onderdelen en het bereiken van alle leerlingen in een klas moeilijk was. Vervolgonderzoek zal moeten uitwijzen of er ook effecten zijn op klassenniveau.

Methodologische beperkingen waren onze zelfgeconstrueerde vragenlijsten, de relatief lage power voor zowel leerling- als leerkrachtanalyses voor programmatrouw, en de vergelijkbare items voor programmatrouw en de uitkomsten. Hoewel de meeste instrumenten in deze studie gevalideerd en/of gebruikt zijn in eerdere studies, was dit niet het geval voor elk instrument (bijvoorbeeld injunctieve norm, intrinsieke prosociale motivatie, veiligheid op school, leerkrachtvragenlijsten). Het zou kunnen dat het gebruik van zelfgeconstrueerde vragenlijsten heeft geresulteerd in eerdere en/of grotere positieve effecten. Echter hebben we ook positieve effecten gevonden op gevalideerde vragenlijsten, niet alleen op zelfgeconstrueerde instrumenten. Bovendien zou het ook kunnen dat het juist heeft geresulteerd in kleinere effecten, bijvoorbeeld bij vragenlijsten met een lagere interne consistentie. Bij de analyses met programmatrouw hadden we een relatief lage power, waardoor we voorzichtig moeten zijn in onze uitspraken over de effectiviteit van het programma. Dit zou ook een verklaring kunnen zijn voor de niet-significante effecten bij leerkrachten, ondanks een positieve trend voor de meeste uitkomsten bij de groep leerkrachten met hoge programmatrouw. Daarnaast overlappen sommige items van de uitkomsten (zoals sociaal gedrag) met dat van de componenten (zoals het ontvangen van een compliment), wat mogelijk associaties heeft uitvergroot. Vergelijkbare resultaten met betrekking tot sociaal gedrag (bijvoorbeeld samenwerking en motivatie) ondersteunen echter de robuustheid van deze bevinding.

4.2 Praktische Implicaties en Suggesties voor Vervolgonderzoek

Onze bevindingen laten zien dat er in de gehele steekproef geen positieve effecten zijn gevonden voor SterkWerk in vergelijking met de controlescholen. Wel lijkt optimale programmatrouw belangrijk voor programma-effectiviteit, aangezien er kleine kortetermijneffecten werden geobserveerd in een subgroep van leerlingen met een hoge

programmatrouw. Programmatrouw wordt door weinig studies gerapporteerd (Bruhn et al., 2015; Harn et al., 2013; Van Loon et al., 2020), terwijl het essentieel is om correcte conclusies te trekken over de effectiviteit van een programma. Ontwikkelaars van interventieprogramma's en onderzoekers die de effectiviteit van dergelijke programma's onderzoeken moeten zich hiervan bewust zijn en worden geadviseerd zich te richten op correcte implementatie van het protocol. Om de implementatie van schoolprogramma's te verbeteren, adviseerde eerder onderzoek om positieve relaties tussen externen (bijvoorbeeld interventie-aanbieders) en schoolpersoneel te creëren en te onderhouden, het programma aantrekkelijk te maken voor gebruikers, voldoende zichtbaar te zijn en voldoende trainingsmogelijkheden te bieden aan de uitvoerders van het programma (Gee et al., 2021; Kaufman & Huitsing, 2019; Van Loon et al., 2024; Wagner et al., 2004). Daarnaast werden de meeste programma-effecten bij ideale programmatrouw waargenomen op de korte termijn, halverwege het programma. Dit geeft aan dat het belangrijk is dat aanbieders betrokken zijn bij de implementatie gedurende het gehele implementatieproces (of langer), niet alleen aan het begin.

Opvallend was dat slechts een klein deel van de leerlingen en leerkrachten ervaarde dat het programma was geïmplementeerd zoals bedoeld, ongeveer een kwart van de deelnemers. Het is essentieel om te achterhalen waarom het programma niet doorgevoerd is zoals bedoeld. Een opvallende constatering was dat de drie hoofdcomponenten afzonderlijk redelijk goed werden geïmplementeerd (de gerapporteerde ideale programmatrouw was respectievelijk 43–68 % en 36–93 % voor leerlingen en leerkrachten), maar de combinatie van alle drie niet (18–23 % voor leerlingen, 26–34 % voor leerkrachten). Het lijkt erop dat leerkrachten de interventie redelijk hebben uitgevoerd, in ieder geval de aparte onderdelen, maar dat de combinatie van deze onderdelen en het bereiken van alle leerlingen in een klas lastig was. Het is daarom belangrijk om te onderzoeken hoe de programmacomponenten worden ingezet (afgezien van de frequentie, zoals onderzocht is in deze studie), wat de werkingsmechanismen zijn, welke ondersteuningsbronnen scholen nodig hebben en welke scholen het meeste profiteren van deze ondersteuning (Gee et al., 2021; Kaufman & Huitsing, 2019). Het is ook waardevol om te onderzoeken of componenten van SterkWerk – en het programma als geheel – ook effectief zijn op scholen die werken met andere sociale veiligheidsprogramma's anders dan KiVa.

Leerkrachten in interventiescholen rapporteerden meer positieve effecten over hun houding ten opzichte van het stimuleren van prosociaal gedrag en autonomie in vergelijking met leerkrachten in de controleconditie. Hoewel er mogelijk sprake is van *bias*, omdat leerkrachten niet blind waren voor het feit dat ze deelnamen aan de interventie, lijkt het erop dat leerkrachten het belang inzien van SterkWerk, en daarbij het werken op een autonomie-stimulerende manier. Eerder onderzoek heeft aangetoond dat het toepassen van een meer autonomie-ondersteunende motiverende stijl in het onderwijs van leerkrachten positieve effecten heeft voor zowel leerkrachten als leerlingen, waaronder meer werkplezier en meer betrokkenheid van leerlingen in de klas (Cheon et al.,

2019). Omdat de rol van de leerkracht in dit programma – en de meeste schoolinterventies – cruciaal is, is het essentieel dat leerkrachten achter het programma staan en gemotiveerd zijn om met het programma te werken. Eerder onderzoek naar een preventieve interventie voor jonge kinderen liet bijvoorbeeld zien dat leerkrachten die meer zorgen uitten over de interventie, minder activiteiten van het programma uitvoerden, terwijl leerkrachten die positiever waren over hun baan, werkomgeving en directie en een grotere betrokkenheid rapporteerden, juist meer activiteiten van het programma uitvoerden (Baker et al., 2010).

4.3 Conclusie

Het is belangrijk om een prosociaal klasklimaat te bevorderen, vanwege het positieve effect op het school- en psychologische welzijn van kinderen en een optimale leeromgeving. De huidige studie onderzocht de effectiviteit van een schoolprogramma, SterkWerk, gericht op het bevorderen van een prosociaal klasklimaat door het versterken van basispsychologische behoeften (autonomie, competentie en verbondenheid). Resultaten in de gehele steekproef lieten geen enkel effect zien bij leerlingen. Leerkrachten rapporteerden wel een positievere eigen houding ten opzichte van het stimuleren van prosociaal gedrag en autonomie. Wanneer leerlingen aangaven dat het programma was geïmplementeerd zoals bedoeld, werden bij deze leerlingen gemiddeld positievere effecten gevonden vergeleken met leerlingen in de controlescholen wat betreft basispsychologische behoeften, prosociaal klasklimaat, welbevinden en veiligheid op school, en antisociaal gedrag. De effecten waren echter klein, kortdurend, en alleen gevonden in een kleine groep leerlingen, wat verder onderzoek vereist. Het is daarnaast ook belangrijk om meer inzicht te krijgen in de programma-implementatie, aangezien de mate van programmatrouw belangrijk lijkt te zijn voor programma-effectiviteit.

Korte Verklaring over de Data en Analyses

De data zijn gearchiveerd in de DANS Easy Environment van de KNAW, zoals overeengekomen met NWO (*open access policy*). Voor meer informatie kan er contact worden opgenomen met de corresponderende auteur.

Auteursbijdragen

Beide auteurs hebben bijgedragen aan de opzet en het ontwerp van de studie. AL coördineerde en voerde de dataverzameling tijdens de studie uit, TK begeleidde de dataverzameling. AL schreef de conceptversie van het manuscript, TK gaf kritische feedback. Beide auteurs hebben het definitieve manuscript gelezen en goedgekeurd.

Literatuur

- Baker, C.N., Kupersmidt, J.B., Voegler-Lee, M.E., Arnold, D.H., & Willoughby, M.T. (2010). Predicting teacher participation in a classroom-based, integrated preventive intervention for preschoolers. *Early Childhood Research Quarterly*, 25(3), 270–283. <https://doi.org/10.1016/j.ecresq.2009.09.005>
- Battistich, V. (2005). Character education, prevention, and positive youth development. *Character Education, Prevention, and Positive Youth Development*, 2000, 1–10.
- Battistich, V., Solomon, D., Watson, M., & Schaps, E. (1997). Caring school communities. *Educational Psychologist*, 32(3), 137–151. https://doi.org/10.1207/s15326985ep3203_1
- Begg, C., Cho, M., Eastwood, S., Horton, R., Moher, D., Olkin, I., Pitkin, R., Rennie, D., Schulz, K.F., Simel, D., & Stroup, D.F. (1996). Improving the quality of reporting of randomized controlled trials: The CONSORT statement. *Journal of the American Medical Association*, 276(8), 637–639. <https://doi.org/10.1001/jama.276.8.637>
- Boor-Klip, H.J., Segers, E., Hendrickx, M.M.H.G., & Cillessen, A.H.N. (2014). Perceptions of classroom peer context: Associations with social status, academic achievement, and self-esteem | Beleving van de peer context in de klas: Samenhang met sociaal functioneren, academisch functioneren en zelfbeeld. *Pedagogische Studien*, 91(5).
- 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(2), 370–389. <https://doi.org/10.1111/sode.12137>
- Boor-Klip, H.J., Segers, E., Hendrickx, M.M.H.G., & Cillissen, A.H.N. (2017). The moderating role of classroom descriptive norms in the association of student behavior with social preference and popularity. *Journal of Early Adolescence*, 37(3), 387–413.
- Bruhn, A.L., Hirsch, S.E., & Lloyd, J.W. (2015). Treatment integrity in school-wide programs: A review of the literature (1993–2012). *Journal of Primary Prevention*, 36(5), 335–349. <https://doi.org/10.1007/s10935-015-0400-9>
- Caprara, G.V., Luengo Kanacri, B.P., Zuffianò, A., Gerbino, M., & Pastorelli, C. (2015). Why and how to promote adolescents' prosocial behaviors: Direct, mediated and moderated effects of the CEPIDEA school-based program. *Journal of Youth and Adolescence*, 44(12), 2211–2229. <https://doi.org/10.1007/s10964-015-0293-1>
- Charlton, C.T., Moulton, S., Sabey, C.V., & West, R. (2021). A systematic review of the effects of schoolwide intervention programs on student and teacher perceptions of school climate. In *Journal of Positive Behavior Interventions* (Vol. 23, Issue 3, pp. 185–200). SAGE Publications Inc. <https://doi.org/10.1177/1098300720940168>
- Cheon, S.H., Reeve, J., & Ntoumanis, N. (2019). An intervention to help teachers establish a prosocial peer climate in physical education. *Learning and Instruction*, 64 (June 2018), 101223. <https://doi.org/10.1016/j.learninstruc.2019.101223>
- Cipriano, C., Strambler, M.J., Naples, L.H., Ha, C., Kirk, M., Wood, M., Sehgal, K., Zieher, A.K., Eveleigh, A., McCarthy, M., Funaro, M., Ponnock, A., Chow, J.C., & Durlak, J. (2023). The state of evidence for social and emotional learning: A contemporary meta-analysis of universal

- school-based SEL interventions. *Child Development*, 94(5), 1181–1204. <https://doi.org/10.1111/cdev.13968>
- Cohen, J. (1988). Statistical power analysis for the social sciences. *Hillsdale, NJ: Erlbaum*.
- Combs, K.M., Buckley, P.R., Lain, M.A., Drewelow, K.M., Urano, G., & Kerns, S.E.U. (2022). Influence of classroom-level factors on implementation fidelity during scale-up of evidence-based interventions. *Prevention Science: The Official Journal of the Society for Prevention Research*, 23(6), 969–981. <https://doi.org/10.1007/s11121-022-01375-3>
- Courser, M.W., Shamblen, S.R., Lavrakas, P.J., Collins, D., & Ditterline, P. (2009). The impact of active consent procedures on nonresponse and nonresponse error in youth survey data: Evidence from a new experiment. *Evaluation Review*, 33(4), 370–395. <https://doi.org/10.1177/0193841X09337228>
- Durlak, J.A., & DuPre, E.P. (2008). Implementation matters: A review of research on the influence of implementation on program outcomes and the factors affecting implementation. *American Journal of Community Psychology*, 41(3–4), 327–350. <https://doi.org/10.1007/s10464-008-9165-0>
- 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(1), 405–432. <https://doi.org/10.1111/j.1467-8624.2010.01564.x>
- Edwards, D., & Mullis, F. (2003). Classroom meetings: encouraging a climate of cooperation. *Professional School Counseling*, 7(1), 20–28.
- Ellis, B.J., Volk, A.A., Gonzalez, J.M., & Embry, D.D. (2016). The Meaningful Roles intervention: an evolutionary approach to reducing bullying and increasing prosocial behavior. *Journal of Research on Adolescence*, 26(4), 622–637. <https://doi.org/10.1111/jora.12243>
- Embry, D.D. (2004). Community-based prevention using simple, low-cost, evidence-based kernels and behavior vaccines. *Journal of Community Psychology*, 32(5), 575–591. <https://doi.org/10.1002/jcop.20020>
- Embry, D.D., & Biglan, A. (2008). Evidence-based kernels: Fundamental units of behavioral influence. *Clinical Child and Family Psychology Review*, 11(3), 75–113. <https://doi.org/10.1007/s10567-008-0036-x>
- Evans, I.M., Harvey, S.T., Buckley, L., & Yan, E. (2009). Differentiating classroom climate concepts: Academic, management, and emotional environments. *Kotuitui*, 4(2), 131–146. <https://doi.org/10.1080/1177083X.2009.9522449>
- Franck, E., De Raedt, R., Barbez, C., & Rosseel, Y. (2008). Psychometric properties of the Dutch Rosenberg Self-Esteem Scale. *Psychologica Belgica*, 48(1), 25–35. <https://doi.org/10.5334/pb-48-1-25>
- Gaffney, H., Ttofi, M.M., & Farrington, D.P. (2021). Effectiveness of school-based programs to reduce bullying perpetration and victimization: An updated systematic review and meta-analysis. *Campbell Systematic Reviews*, 17(2). <https://doi.org/10.1002/cl2.1143>
- Gee, B., Wilson, J., Clarke, T., Farthing, S., Carroll, B., Jackson, C., King, K., Murdoch, J., Fonagy, P., & Notley, C. (2021). Review: Delivering mental health support within schools and colleges – a thematic synthesis of barriers and facilitators to implementation of indicated psychological

- interventions for adolescents. *Child and Adolescent Mental Health*, 26(1), 34–46. <https://doi.org/10.1111/camh.12381>
- Gfroerer, K., Edwards, D., & Dwight, E. (2024). Utilizing class meetings to teach essential social-emotional skills and build community. *Journal of Individual Psychology*, 80(3), 188–202. <https://doi.org/10.1353/jip.00004>
- Haataja, A., Voeten, M., Boulton, A.J., Ahtola, A., Poskiparta, E., & Salmivalli, C. (2014). The KiVa antibullying curriculum and outcome: Does fidelity matter? *Journal of School Psychology*, 52(5), 479–493. <https://doi.org/10.1016/j.jsp.2014.07.001>
- Hariton, E., & Locascio, J.J. (2018). Randomised controlled trials – the gold standard for effectiveness research. *BJOG: An International Journal of Obstetrics & Gynaecology*, 125(13), 1–4. <https://doi.org/10.1111/1471-0528.15199>
- Harn, B., Parisi, D., & Stoolmiller, M. (2013). Balancing fidelity with flexibility and fit: What do we really know about fidelity of implementation in schools? *Exceptional Children*, 79(2), 181–193. <https://doi.org/10.1177/001440291307900204>
- Huitsing, G., Lodder, G.M.A., Browne, W.J., Oldenburg, B., Van der Ploeg, R., & Veenstra, R. (2020). A large-scale replication of the effectiveness of the KiVa antibullying program: a randomized controlled trial in the Netherlands. *Prevention Science*, 21(5), 627–638. <https://doi.org/10.1007/s11121-020-01116-4>
- Kärnä, A., Voeten, M., Little, T.D., Poskiparta, E., Kaljonen, A., & Salmivalli, C. (2011). A large-scale evaluation of the KiVa antibullying program: Grades 4–6. *Child Development*, 82(1), 311–330. <https://doi.org/10.1111/j.1467-8624.2010.01557.x>
- Kaufman, T., & Huitsing, G. (2019). Implementatie van een antipestprogramma: Wat zijn de succes-ingrediënten? *Basischool Management*, 3, 10–13.
- Kaufman, T.M.L., Huitsing, G., Bloemberg, R., & Veenstra, R. (2021). The systematic application of network diagnostics to monitor and tackle bullying and victimization in schools. *International Journal of Bullying Prevention*, 3(1), 75–87. <https://doi.org/10.1007/s42380-020-00064-5>
- Khalfoui, A., García-Carrión, R., & Villardón-Gallego, L. (2021). A systematic review of the literature on aspects affecting positive classroom climate in multicultural early childhood education. *Early Childhood Education Journal*, 49(1), 71–81. <https://doi.org/10.1007/s10643-020-01054-4>
- Klocek, A., Kollerová, L., Havrdová, E., Kotrbová, M., Netík, J., & Pour, M. (2024). Effectiveness of the KiVa anti-bullying program in the Czech Republic: A cluster randomized control trial. *Evaluation and Program Planning*, 106. <https://doi.org/10.1016/j.evalprogplan.2024.102459>
- Lane, K.L., Bocian, K.M., MacMillan, D.L., & Gresham, F.M. (2004). Treatment integrity: An essential – but often forgotten – component of school-based interventions. *Preventing School Failure: Alternative Education for Children and Youth*, 48(3), 36–43. <https://doi.org/10.3200/psfl.48.3.36-43>
- Lenhard, W., & Lenhard, A. (2022). *Computation of effect sizes*. Psychometrica. <https://doi.org/10.13140/RG.2.2.17823.92329>
- Lickona, D., & Davidson, M. (2003). School as a caring community profile-II. *Unpublished Manuscript*.

- Limber, S.P., Olweus, D., Wang, W., Masiello, M., & Breivik, K. (2018). Evaluation of the Olweus Bullying Prevention Program: A large scale study of U.S. students in grades 3–11. *Journal of School Psychology*, 69, 56–72. <https://doi.org/10.1016/j.jsp.2018.04.004>
- Lindenberg, S., & Steg, L. (2007). Normative, gain and hedonic goal frames guiding environmental behavior. *Journal of Social Issues*, 63(1), 117–138.
- Liu, C., Cox, R.B., Washburn, I.J., Croff, J.M., & Crethar, H.C. (2017). The effects of requiring parental consent for research on adolescents' risk behaviors: A meta-analysis. *Journal of Adolescent Health*, 61(1), 45–52. <https://doi.org/10.1016/j.jadohealth.2017.01.015>
- van Loon, A.W.G., Creemers, H.E., Beumer, W.Y., Okorn, A., Vogelaar, S., Saab, N., Miers, A.C., Westenberg, P.M., & Asscher, J.J. (2020). Can schools reduce adolescent psychological stress? A multilevel meta-analysis of the effectiveness of school-based intervention programs. *Journal of Youth and Adolescence*. <https://doi.org/10.1007/s10964-020-01201-5>
- van Loon, A.W.G., Creemers, H.E., Vogelaar, S., Saab, N., Miers, A.C., Westenberg, P.M., & Asscher, J.J. (2024). Implementation and experimental evaluation of school-based intervention programs promoting adolescent mental health: Lessons learned. *School Psychology International*, 45(1), 70–84. <https://doi.org/10.1177/01430343231183999>
- Ma, X., Shen, J., & Reeves, P. (2023). Measuring integrity and fidelity of program implementation: Validating an instrument designed for school renewal. *Evaluation and Program Planning*, 100, 102341. <https://doi.org/10.1016/j.evalprogplan.2023.102341>
- Manzano-Sánchez, D., & Valero-Valenzuela, A. (2019). Implementation of a model-based programme to promote personal and social responsibility and its effects on motivation, prosocial behaviours, violence and classroom climate in primary and secondary education. *International Journal of Environmental Research and Public Health*, 16(21). <https://doi.org/10.3390/ijerph16214259>
- Mesurado, B., Guerra, P., Richaud, M.C., & Rodruíguez, L.M. (2018). Effectiveness of prosocial behavior interventions: A meta-analysis. In *Psychiatry and Neuroscience Update: From Translational Research to a Humanistic Approach: Volume III*. <https://doi.org/10.1007/978-3-319-95360-1>
- Montori, V.M., & Guyatt, G.H. (2001). Intention-to-treat principle. *Canadian Medical Association Journal*, 165(10), 1339–1341.
- Nederlands Jeugdinstituut. (2025). *Databank Effectieve jeugdinterventies*. <https://www.nji.nl/interventies>
- Ntoumanis, N., Ng, J.Y.Y., Prestwich, A., Quested, E., Hancox, J.E., Thøgersen-Ntoumani, C., Deci, E.L., Ryan, R.M., Lonsdale, C., & Williams, G.C. (2021). A meta-analysis of self-determination theory-informed intervention studies in the health domain: effects on motivation, health behavior, physical, and psychological health. *Health Psychology Review*, 15(2), 214–244. <https://doi.org/10.1080/17437199.2020.1718529>
- Olweus, D. (1996). *The Revised Olweus Bully/Victim Questionnaire*. Mimeo. Bergen, Norway: Research Center for Health Promotion (HEMIL Center), University of Bergen. <https://doi.org/10.1037/t09634-000>
- Orobio De Castro, B., Mulder, S., Van Der Ploeg, R., Onrust, S., Van Den Berg, Y., Stoltz, S., Buil, M., De Wit, I., Buitenhuis, L., Cillessen, T., Veenstra, R., Van Lier, P., Deković, M., & Scholte, R. (2018).

Effectiviteit van kansrijke programma's tegen pesten in de Nederlandse onderwijspraktijk Wat werkt tegen pesten?

- Pauw, L.M.J. (2013). *Onderwijs en burgerschap: wat vermag de basisschool? Onderzoek naar De Vreedzame School*.
- Pokorny, S.B., Jason, L.A., Schoeny, M.E., Townsend, S.M., & Curie, C.J. (2001). When active consent procedures replace passive consent. *Sage Publications*, 25(5), 567–580.
- Research randomizer*. (2023). <https://www.randomizer.org/>
- Rojas-Andrade, R., & Bahamondes, L.L. (2019). Is implementation fidelity important? A systematic review on school-based mental health programs. *Contemporary School Psychology*, 23(4), 339–350. <https://doi.org/10.1007/s40688-018-0175-0>
- Rosenberg, M. (1965). *Society and the adolescent self-image*. Princeton University Press, Princeton, NJ.
- Ryan, R.M., & Deci, E.L. (2000). Self-determination theory and the facilitation of intrinsic motivation, social development, and well-being. *American Psychologist*, 55(1), 68–78. <https://doi.org/10.1037/0003-066X.55.1.68>
- Ryan, R.M., & Deci, E.L. (2017). Self-determination theory: Basic psychological needs in motivation, development, and wellness. In *Elgar Encyclopedia of Organizational Psychology*. <https://doi.org/10.4337/9781803921761.00125>
- Saarento, S., & Salmivalli, C. (2015). The role of classroom peer ecology and bystanders' responses in bullying. *Child Development Perspectives*, 9(4), 201–205. <https://doi.org/10.1111/cdep.12140>
- Sanetti, L.M.H., Cook, B.G., & Cook, L. (2021). Treatment fidelity: What it is and why it matters. *Learning Disabilities Research and Practice*, 36(1), 5–11. <https://doi.org/10.1111/ldrp.12238>
- Sheldon, K.M., Ryan, R., & Reis, H.T. (1996). What makes for a good day? Competence and autonomy in the day and in the person. *Personality and Social Psychology Bulletin*, 22(12), 1270–1279.
- Shin, J., & Lee, B.Y. (2021). The effects of adolescent prosocial behavior interventions: a meta-analytic review. *Asia Pacific Education Review*, 22(3), 565–577. <https://doi.org/10.1007/s12564-021-09691-z>
- Sklad, M., Diekstra, R., de Ritter, M., Ben, J., & Gravesteyn, C. (2012). Effectiveness of school-based universal social, emotional, and behavioral programs: do they enhance students' development in the area of skill, behavior, and adjustment? *Psychology in the Schools*, 49(9), 892–909. <https://doi.org/10.1002/pits>
- Stjerneklar, S., Hougaard, E., & Thastum, M. (2019). Guided internet-based cognitive behavioral therapy for adolescent anxiety: Predictors of treatment response. *Internet Interventions*, 15 (January), 116–125. <https://doi.org/10.1016/j.invent.2019.01.003>
- Tanner-Smith, E.E., Durlak, J.A., & Marx, R.A. (2018). Empirically based mean effect size distributions for universal prevention programs targeting school-aged youth: A review of meta-analyses. *Prevention Science*, 19(8), 1091–1101. <https://doi.org/10.1007/s1121-018-0942-1>
- Taylor, R.D., Oberle, E., Durlak, J.A., & Weissberg, R.P. (2017). Promoting positive youth development through school-based social and emotional learning interventions: A meta-analysis of follow-up effects. *Child Development*, 88(4), 1156–1171. <https://doi.org/10.1111/cdev.12864>

- Thornberg, R. (2011). 'She's Weird!' – the social construction of bullying in school: A review of qualitative research. *Children and Society*, 25(4), 258–267. <https://doi.org/10.1111/j.1099-0860.2011.00374.x>
- Thornberg, R., Wegmann, B., Wänström, L., Bjereld, Y., & Hong, J.S. (2022). Associations between student–teacher relationship quality, class climate, and bullying roles: A Bayesian multilevel multinomial logit analysis. *Victims and Offenders*, 17(8), 1196–1223. <https://doi.org/10.1080/15564886.2022.2051107>
- Tigges, B.B. (2003). Parental consent and adolescent risk behavior research. *Journal of Nursing Scholarship*, 35(3), 283–289. <https://doi.org/10.1111/j.1547-5069.2003.00283.x>
- Van Widenfelt, B.M., Goedhart, A.W., Treffers, P.D.A., & Goodman, R. (2003). Dutch version of the Strengths and Difficulties Questionnaire (SDQ). *European Child and Adolescent Psychiatry*, 12(6), 281–289. <https://doi.org/10.1007/s00787-003-0341-3>
- Veenstra, R., Lindenberg, S., Huitsing, G., Sainio, M., & Salmivalli, C. (2014). The role of teachers in bullying: The relation between antibullying attitudes, efficacy, and efforts to reduce bullying. *Journal of Educational Psychology*, 106(4), 1135–1143. <https://doi.org/10.1037/a0036110>
- Vermande, M.M. (2015). Schoolbrede interventie-programma's. In *Pesten op school*. www.veiligheld.nl
- Villardón-Gallego, L., García-Carrión, R., Yáñez-Marquina, L., & Estévez, A. (2018). Impact of the interactive learning environments in children's prosocial behavior. *Sustainability (Switzerland)*, 10(7), 1–12. <https://doi.org/10.3390/su10072138>
- Wagner, E.F., Tubman, J.G., & Gil, A.G. (2004). Implementing school-based substance abuse interventions: Methodological dilemmas and recommended solutions. *Addiction*, 99 (s2), 106–119. <https://doi.org/10.1111/j.1360-0443.2004.00858.x>
- Waller, G., Finch, T., Giles, E.L., & Newbury-Birch, D. (2017). Exploring the factors affecting the implementation of tobacco and substance use interventions within a secondary school setting: A systematic review. *Implementation Science*, 12(1), 1–18. <https://doi.org/10.1186/s13012-017-0659-8>
- Wang, M., Te, L., Degol, J., Amemiya, J., Parr, A., & Guo, J. (2020). Classroom climate and children's academic and psychological wellbeing: A systematic review and meta-analysis. *Developmental Review*, 57, 100912. <https://doi.org/10.1016/j.dr.2020.100912>
- Wolfenden, L., Kypri, K., Freund, M., & Hodder, R. (2009). Obtaining active parental consent for school-based research: A guide for researchers. *Australian and New Zealand Journal of Public Health*, 33(3), 270–275. <https://doi.org/10.1111/j.1753-6405.2009.00387.x>