Skills and Struggles in the intra- and interpersonal domain

Effectiveness of universal school-based interventions

E.C.A. Mertens

Skills and Struggles in the Intra- and Interpersonal Domain

Effectiveness of Universal School-Based Interventions

Colofon

Printed by Ridderprint | www.ridderprint.nl
Cover design by Sjaak van der Vooren
Layout and design by Dagmar van Schaik | www.persoonlijkproefschrift.nl
ISBN: 978-90-393-7349-1
https://doi.org/10.33540/63

Skills and Struggles in the Intra- and Interpersonal Domain

Effectiveness of Universal School-Based Interventions

Bekwaamheid en Belemmering in het Intra- en Interpersoonlijk Domein:

Effectiviteit van Universele Schoolinterventies
(met een samenvatting in het Nederlands)

Proefschrift

ter verkrijging van de graad van doctor aan de
Universiteit Utrecht
op gezag van de
rector magnificus, prof.dr. H.R.B.M. Kummeling,
ingevolge het besluit van het college voor promoties
in het openbaar te verdedigen op

vrijdag 15 januari 2021 des middags te 2.30 uur

door

Esther Catharina Adriana Mertens

geboren op 2 januari 1990
te Dongen

Promotor: Prof. dr. M. Deković

Copromotoren: Dr. E. Reitz

Dr. W.M. van Londen - Barentsen

Dit proefschrift werd (mede) mogelijk gemaakt met financiële steun van De Nederlandse Organisatie voor Gezondheidsonderzoek en Zorginnovatie (ZonMw) subsidienummer: 531001106.

Beoordelingscommissie

Prof. Dr. J. J. Asscher	Utrecht University
Prof. Dr. S. Branje	Utrecht University
Prof. Dr. F. Gardner	University of Oxford
Prof. Dr. B. Orobio de Castro	University of Amsterdam
Prof. Dr. S. Thomaes	Utrecht University

Table of Contents

1

General Introduction

General introduction

Skills and struggles in the intrapersonal and interpersonal domain are important predictors of adolescents' psychosocial wellbeing and their development into adulthood (Barber, 2005; Shek & Leung, 2016). The *intrapersonal* domain refers to the ability to manage one's own feelings, emotions, and attitudes about the self (Barber, 2005). This domain concerns the subjective processing of behaviors, thoughts, and emotions pertained by the individual self (Dufner, Gebauer, Sedikides, & Denissen, 2019; Finkel & Vohs, 2006). Evaluating and regulating one's own inner world and experiences can facilitate positive personal functioning (e.g., psychological wellbeing and resilience), whereas difficulties in this process can create psychological problems (e.g., internalizing behavior; Dufner et al., 2019). The *interpersonal* domain refers to the ability to build and maintain positive relationships with others, to understand social situations, roles and norms, and to respond appropriately (Pellegrino & Hilton, 2012; Shek & Leung, 2016). By planning one's own behavior and predicting the behavior of others one can act in a socially appealing way, such as building positive interpersonal relations, or in a more destructive way, such as behaving aggressively or bullying (Finkel & Vohs, 2006). Both domains bidirectionally influence each other. For instance, how individuals view themselves can influence the way they approach social interactions and vice versa (Finkel & Vohs, 2006).

Although the domains are intertwined, they are regarded as distinct domains. While the intrapersonal domain reflects subjective personal functioning, the interpersonal domain reflects social functioning (Dufner et al., 2019). Not only do factor and profile analyses support this distinction (Gilman & Anderman, 2006; Park, Tsukayama, Goodwin, Patrick, & Duckworth, 2017), the two domains are also related to different outcomes. For instance, the intrapersonal domain predicts academic achievement, whereas the interpersonal domain predicts positive peer relations (Park et al., 2017). Individuals can develop competencies in both domains by mastering relevant cognitive, affective, and social skills such as the ability to identify emotions (intrapersonal domain), to take perspective, and to help others (interpersonal domain; Durlak, Weissberg, Dymnicki, & Taylor, 2011; Hughes, Kratsiotis, Niven, & Holman, 2020). When students struggle with mastery of (some of) these skills, individuals have an increased risk of developing problems in the intrapersonal domain, such as internalizing behavior (White, Jarrett, & Ollendick, 2013), as well as in the interpersonal domain, such as externalizing behavior (Modecki, Zimmer-Gembeck, & Guerrà, 2017).

Competencies within both the intra- and interpersonal domain are particularly pivotal during adolescence for two main reasons. First, adolescents are consolidating their own identity (Barber, 2005). To foster the process of identity formation, adolescents need to be able to identify their strengths, limitations and values, and need to have a positive attitude about the self (Barber, 2005; Shek & Leung, 2016). Positive and negative feelings and emotions take a central role in the process of

identity formation as these can influence how experiences and memories are processed (Haviland, Davidson, Ruetsch, Gebelt, & Lancelot, 1994). Second, adolescents are becoming more aware of others while spending more time outside the home encountering varying contexts in which they need to interact with others, such as the school, occupational, and romantic contexts (Barber, 2005). Adolescents' ability to positively engage in interactions and constructively manage social situations is stimulated by their level of awareness of social norms and cues, and understanding that others may not have the same thoughts and feelings as they do (Shek & Leung, 2016). Hence, adolescents' psychosocial development can benefit from enhancing their competencies and preventing development of problems in both the intra- and interpersonal domains.

The school environment has been recognized as a powerful context to address adolescents' intra- and interpersonal domains (Pellegrino & Hilton, 2012). Schools have the responsibility to actively cultivate their students' attitudes, values, and social support in addition to stimulating students' academic development (Langford et al., 2014; World Health Organization, 1995). Given that adolescents spend a lot of their time at school, it also creates a perfect opportunity to involve adolescents in an intervention by implementing it during school hours (Langford et al., 2014). This is especially beneficial for those interventions that aim to involve a group of adolescents that are more difficult to reach outside the school context (Liber, De Boo, Huizenga, & Prins, 2013).

In the Netherlands there are three educational tracks in secondary school, starting at 7th Grade: Preparatory vocational track (prevocational track), preparatory college track, and preparatory university track. Compared to students in the preparatory college and university tracks, prevocational students report higher levels of behavioral problems (20% versus 13% and 7%), have more problems with peers (20% versus 10% and 8%; Stevens & De Looze, 2018), and are more likely to engage in risk behaviors (e.g., binge-drinking: 48% versus 37% and 25%; smoking: 17% versus 5% and 2%; Harakeh, De Looze, Schrijvers, Van Dorsselaer, & Vollebergh, 2012). These increased risks for psychological and behavioral problems indicate the necessity of effective interventions for this group of students, as more than half (54%) of the entire Dutch student population follows this prevocational track (Central Bureau for Statistics, 2020). However, prevocational students are a challenging group to involve in an intervention. They generally show lower levels of autonomy, show less intrinsic motivation for school and (verbal) learning, and have lower cognitive capacities compared to students following the other two tracks (Timmermans, Naaijer, Keuning, & Zijsling, 2017). These characteristics suggest that prevocational students might require a specific intervention approach that fits in well with their learning processes to be able to fully benefit from an intervention.

The main aim of this dissertation was to study whether and under what circumstances universal interventions in secondary schools can successfully stimulate students' competencies and prevent problems in the intra- and interpersonal domain.

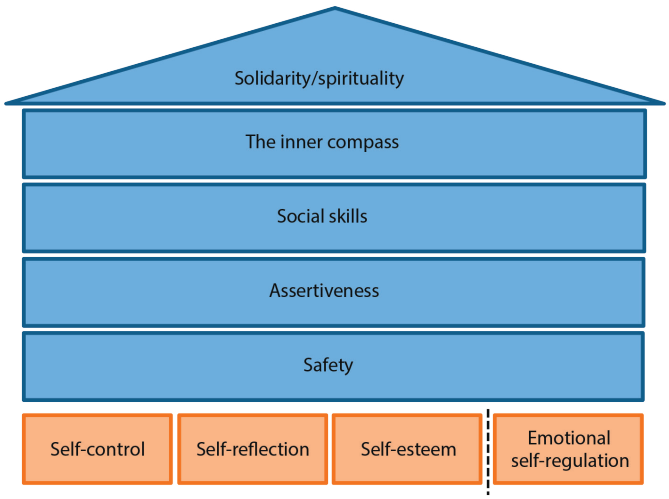
I evaluated this type of school-based interventions by means of a Randomized Controlled Trial (RCT) and a meta-analysis. I conducted the RCT to gain a detailed understanding of one specific intervention, Rock and Water (R&W; Ykema, 2002; 2018) – with a focus on prevocational students – and conducted the meta-analysis to increase insights in these interventions in general. More specifically, through the RCT and the meta-analysis I examined 1) the effectiveness of universal school-based interventions, 2) whether heterogeneity in the context and in the student population affected intervention effects, and 3) working mechanisms of universal school-based interventions. In addition, using both research approaches enabled me to discuss findings concerning one specific universal school-based intervention, as well as to relate these intervention specific findings to general findings of school-based interventions.

Rock and Water

R&W (Ykema, 2002; 2018) is a 2-year universal school-based intervention aiming to improve competencies and to prevent the development of problems in the intra- and interpersonal domain. The intervention is based on the "Rock & Water house" (see Figure 1). The R&W house consists of five levels representing the five modules of the intervention: Safety, assertiveness, social skills, the inner compass, and solidarity/spirituality. In these modules, R&W teaches students how to feel safe, and how to prevent and deal with violence (i.e., safety), how to deal with difficult situations without losing self-control (i.e., assertiveness), how to have discussions with people who have different opinions, how to set and respect boundaries, and how to understand and use non-verbal communication (i.e., social skills). Students learn to value their own intuition and preferences, and to make choices based on their intuition (i.e., the inner compass), and increase insight in themselves and others as they are all connected with each other (i.e., solidarity/spirituality). The foundation of the house consists of the three pillars of self-control, self-reflection, and self-esteem. Closely related to these pillars is emotional self-regulation, a potentially fourth pillar. The R&W house theory states that students need to improve their self-control, self-reflection, self-esteem, and emotional self-regulation to be able to develop themselves within the modules of the intervention.

The R&W intervention lessons are provided by teachers, mostly physical education teachers, during physical education classes. These teachers complete the 3-days basic R&W training course to become certified R&W trainers. Optionally, the rest of the school staff can follow an introduction training course of R&W during which the basic principles of the intervention are introduced. School staff learns how to apply the intervention principles during regular classes (e.g., math class) and how to support the R&W trainers within their school. In the manual, the role of the R&W trainer is explicitly described with a specific focus on the responsibility of the trainer

Figure 1. Rock & Water house.



to create a safe and trusting environment. It is emphasized that the trainer functions as a role model for the students and should reinforce students' positive behaviors during the lessons. The manual also describes the intervention lessons elaborately for the R&W trainers. For each lesson, the exercises and games are described and are accompanied with a brief theoretical explanation, specific instructions that can be used during the lesson, and pictures showing examples of the exercises/games.

The intervention uniquely combines a physical approach with a more common psychological approach, that is, a psychophysical approach. Through games and exercises students learn how to make (physical) contact with others, and explore, respect, and set own and other's boundaries. After the physical activities, there is a short moment of reflection to share their thoughts and feelings with each other, and to discuss how these skills can be transferred to their daily lives. During the intervention lessons, communication is based on the symbolic principles of "rock" and "water", representing opposite ends of a continuum. Rock indicates an uncompromising attitude in which one is able to resist pressure from others. Water indicates a flexible and cooperative attitude in which one is open to opinions, thoughts, and feelings of others. For both attitudes students need to be "grounded" and "centered". This means that both feet should be on the ground with the weight divided equally over both feet, feet are slightly placed apart, and the knees are slightly bend (i.e., grounded). In addition, breathing is low instead of high in the chest (i.e., centered).

The psychophysical approach and the used communication are evident in the exercises. For instance, in the exercise "Chinese boxing" two students stand oppose each other, both grounded and centered, and hold their hands in front of their body with the palms of the hands directed towards the other student (see Figure 2). The goal

of the game is to tap the hands of the opponent and try to push the other student out of balance. The most important aim of the exercises is that students remain grounded and centered while playing. In the game one can keep his/her body tense as a rock (Rock action) or move along with the movement of the opponent as water (Water action). In another exercise, students punch a bag on command of the R&W trainer who regularly interrupts the rhythm of commands to punch. This requires students to actively control their behavior, inhibiting the inclination to keep punching in the rhythm of commands. Yet another exercise involves a student who "stands strong", that is, grounded and centered, so that another student can lean against this student. After each exercise, students reflect on how grounded and centered they were and discuss that sometimes a rock attitude was more effective during the exercises and other times a water attitude was more effective. The intervention's theory suggests that with this approach students learn that a rock and a water attitude are on the same continuum and that students can shift between these attitudes. However, a good choice between a Rock action and a Water action can only be made when someone is centered and grounded. Additionally, R&W trainers stimulate students to think about how they can apply this in their daily lives. In these examples, playing the actual game represents the physical approach and the short moment of reflection the psychological approach.

Figure 2. Example of "Chinese Boxing".

Due to the psychophysical approach, R&W seems especially promising for prevocational students as this approach seems to fit in well with their learning processes. The recommended learning environment of this group of students is preferably characterized by a practical orientation in which students can practice the learned skills. When teaching prevocational students, short moments of instruction or reflection should be alternated with practical exercises. It is preferred to use activating methods and stepwise, structured tasks to help students regulate their learning processes. This teaching context is suggested to enable students to understand, use, and apply the learned knowledge more effectively (De Bruijn et al., 2005; Koopman, Den Brok, Bijaard, & Teune, 2011). Hence, the psychophysical approach of R&W seems a fitting intervention approach to target prevocational students as it uses a combination of shortly explaining or reflecting on the intervention techniques (i.e., a psychological approach) and physically practicing the intervention techniques in a relevant social context (i.e., a physical approach).

Although R&W is broadly implemented in multiple countries (e.g., the Netherlands, Australia, China, Peru), little is known about its effectiveness. In the Netherlands 18,000 professionals have attended the basic training to become a certified R&W trainer and 65 schools have the R&W certificate (i.e., all students receive at least 10 R&W lessons each year and the entire school staff is trained in R&W). Results of previous research indicate that R&W is a promising intervention establishing improvements in both the intrapersonal (e.g., resilience, identity formation, coping styles; Ykema, Hartman, & Imms, 2006) and the interpersonal (e.g., coercive strategies; De Graaf, De Haas, Zaagsma, & Wijsen, 2016) domain. However, these studies were limited in samples (e.g., small sample size, only boys) and measurements (e.g., narrowly defined outcomes, self-reports, only a pre- and post-measurement). Therefore, I examined the effectiveness of R&W with an RCT design in a relatively large sample, including both boys and girls, measured a broad range of outcomes belonging to either the intrapersonal domain or the interpersonal domain, used students' self-reports as well as parental reports and observations, and conducted measurements before, during, and after the intervention.

Aim 1: Effectiveness of Universal School-Based Interventions

In the current dissertation, the effectiveness of universal school-based interventions aiming to stimulate competencies and to prevent problems in the intra- and interpersonal domain was studied for R&W in specific and for universal school-based interventions in general. The effectiveness of R&W was examined with an RCT design, which is considered the "golden standard", providing the most rigorous assessment of intervention effects as possible (Bonell, Fletcher, Morton, Lorenc, & Moore, 2012). Furthermore, the intervention was examined under real-world conditions rather than under optimal conditions highly controlled by researchers. For instance, R&W was

implemented by trained regular school teachers instead of experienced professional facilitators. This elevates the findings from evidence for efficacy to evidence for effectiveness as stated in the "Standards of Evidence" formulated by the Society for Prevention Research (Flay et al., 2005).

To evaluate the general effectiveness of universal school-based interventions fostering students' development in the intra- and/or interpersonal domain, I conducted a meta-analysis. I first examined the overall effectiveness of these interventions in the intrapersonal domain and in the interpersonal domain in general. Second, I analyzed intervention effects on specific competencies and problems within the intrapersonal (i.e., resilience, self-esteem, self-regulation, general wellbeing, and internalizing behavior) and the interpersonal (i.e., sexual health, social competence, school climate, aggression, and bullying) domain. Studying domains in general as well as specific competencies and problems within these domains allowed me to account for the multidimensionality of the general domains (O'Mara, Marsh, Craven, & Debus, 2006). Moreover, it gives a detailed overview of which competencies and problems can currently be successfully affected by universal school-based interventions and which need more attention.

Studying intervention effects of universal school-based interventions in general provides the framework to compare the effectiveness of specifically R&W with the effectiveness of school-based interventions in general. Previously, obtained intervention effect were mostly interpreted according to Cohen's (1988) classification of effect sizes. Recently, researchers are stressing the importance of interpreting effect sizes in relevant contexts such as intervention effects found in previous studies (Durlak, 2009). In order to compare intervention effects in a relevant context, Durlak (2009) proposes three guidelines to take into consideration: 1) quality of the research, 2) comparisons across similar interventions, and 3) practical relevance. I followed these guidelines to establish a relevant context in which I could interpret the intervention effects of R&W obtained in my RCT. That is, in my meta-analysis I took the quality of research (e.g., randomization, type of control group) into account in the calculation of effect sizes when relevant, and I included interventions similar to R&W (e.g., implemented during regular school hours, universal, aim). Practical relevance, the third guideline, is determined by reflecting on the extent to which students' competencies and problems are meaningfully improved after the intervention. Thus, my meta-analysis not only provides an overview of the effectiveness of universal school-based interventions, it also establishes a relevant context for interpretation of the intervention effects of R&W.

Aim 2: Heterogeneity in Contexts and in Population

Information regarding under what circumstances (context) and for whom (population) an intervention is effective is not only essential to gain a detailed understanding of the

intervention's effectiveness, but also for theory building and implementation of the program (Bonell et al., 2012). Contexts and populations are by nature heterogeneous which could influence intervention effects. What works in one context may not work in another context. Similarly, some individuals with certain characteristics may improve, or deteriorate, while the total population might show no or weak changes in the outcome of interest (Bonell et al., 2012; Farrell, Henry, & Bettencourt, 2013; Greenberg & Abenavoli, 2017). Knowing whether characteristics of contexts and participants affect intervention's effectiveness gives an indication to what extent the intervention effects can be generalized (Rowe & Trickett, 2018). This is especially important in universal interventions as these are typically implemented in a broad context targeting a heterogeneous population (Farrell et al., 2013). In addition, insights in how characteristics of the context and participants affect intervention effects are eminent for theory development. If intervention effects are affected by certain characteristics, this suggests differences in the underlying working mechanism (Kazdin, 2007).

The importance of studying under what circumstances and for whom an intervention is effective has been long acknowledged and is even incorporated in the "Standards of Evidence" of the Society for Prevention Research (Flay et al., 2005). However, most studies neglect characteristics of the context and only examine routinely collected characteristics of participants such as sex and ethnicity. Moreover, most studies determine moderators in post hoc analyses rather than based on theory and a priori hypotheses (Farrell et al., 2013; Kazdin, 2007). For instance, Rowe and Trickett (2018) showed in their meta-analysis that only 20 of the 50 included moderation analyses were supported by literature or a priori hypotheses. In the present dissertation, I attempted to overcome these limitations by determining moderators a priori and by examining characteristics of the context as well as characteristics of participants. This enabled me to move beyond moderation analyses based on convenience (i.e., on routinely collected characteristics) and test theoretically relevant characteristics (Kazdin, 2007).

As characteristics of the context, I focused on two aspects of intervention dosage. The first dosage characteristic was the ecological width of an intervention, that is, the involvement of multiple systems in the intervention. The social ecological model of Bronfenbrenner (1979) states that behavior is determined by the interactions of multiple systems such as the individual, family, and school systems. In some interventions stronger effects were found when more systems were actively involved in the interventions (e.g., Flay, Graumlich, Segawa, Burns, & Holliday, 2004), while in other interventions no effects of involving multiple systems were found on their effectiveness (e.g., Durlak et al., 2011). In the present dissertation, I manipulated the extent to which systems were involved in R&W resulting in three levels of ecological width, i.e., 1) "Light condition": Only R&W trainers (core team of teachers) were involved in the intervention (i.e., class system), 2) "Standard condition": The entire teaching staff of the school was involved (i.e., class and school systems), 3) "Plus condition": The entire teaching staff of the school and parents were involved (i.e., class, school, and family

systems). Even though manipulation of dosage is typically not feasible in evaluation studies (Farrell et al., 2013), the design of my RCT enabled me to manipulate dosage and randomly assign schools to these different conditions.

The second dosage characteristic I examined concerned the time span of the intervention. It is reasonable to expect that the rate of change in participants during an intervention might not be linear. Therefore, I conducted measurements before, during, and after the intervention in order to model trajectories of change and examine intervention effects over time, as recommended by Greenberg and Abenavoli (2017). Overall, with this design I was in the position to evaluate the effect of the dosage of the intervention to which students were exposed (i.e., both ecological focus and time span of the intervention) on intervention effects.

As characteristic of participants, I focused on the extent to which students' Big Five personality traits affected intervention effects. According to the vulnerability theory (Tackett, 2006), certain personality traits can put individuals at increased risk of developing problems in the intra- and interpersonal domain. Especially these more vulnerable students might benefit most from the intervention. As the Risk moderation hypothesis (Spoth, Shin, Guyll, Redmon, & Azevedo, 2006) states, vulnerable students have more room to improve on competencies or to decrease on problems, indicating the potential for a compensatory effect, whereas less vulnerable students might experience a ceiling or a floor effect (Nehmy & Wade, 2014). Furthermore, certain personality traits might enhance the transfer of skills learned during the intervention to one's daily life facilitating generalization of the skills. Research showed that Big Five personality traits can indeed affect intervention effects. However, the results of these studies do not show a clear pattern indicating which personality traits are most important to take into account in interventions. For instance, Huppert and Johnson (2010) found stronger intervention effects on wellbeing for individuals with high levels of Agreeableness and Neuroticism. De Vibe and colleagues (2015) also found stronger effects on wellbeing for high levels of Neuroticism, but not for Agreeableness. Additionally, Wang and colleagues (2017) found no moderation of intervention effects on wellbeing by personality traits. The present study examined personality traits as moderators of intervention effects across a broad range of outcomes in the intra- and interpersonal domains aiming to reveal a pattern of moderation that could help clarifying the role of personality traits in interventions.

Aim 3: Working Mechanisms in Interventions

Knowledge about how an intervention establishes change is essential for understanding what is critical to an intervention and how interventions can be optimized (Kazdin, 2007). Studying working mechanisms can confirm theories, may lead to new insights for theory development (Michie, Fixsen, Grimshaw, & Eccles, 2009), or might inform directions for intervention optimization (Kazdin, 2007). In the present

dissertation, I investigated working mechanisms through examining mechanisms of change as well as intervention components. By analyzing mechanisms of change (i.e., mediators) I examine changes within the participants (Longabaugh & Magill, 2011), whereas by analyzing components of interventions I examine aspects of interventions related to intervention effects (O'Rourke & MacKinnon, 2018).

First, I studied mechanisms of change (i.e., mediators). Mechanisms of change indicate behaviors or processes within the participants that have changed as a result of an intervention and can be related to subsequent changes in the ultimate outcome (Longabaugh & Magill, 2011). Mediation analyses provide insights of practical as well as theoretical importance and can be used to inform directions for future research. For instance, the intervention might have failed in successfully addressing the mediator (i.e., action theory failure), the mediator might not be related to the outcome (i.e., conceptual theory failure), or both (O'Rourke & MacKinnon, 2018). I examined whether the effect of R&W on competencies and problems in the interpersonal domain was mediated by changes in classmates' deviant and prosocial modeling and reinforcement. As stated in the social learning theory (Bandura, 1977), peers influence each other through the mechanisms of modeling (i.e., learning new behaviors and tendencies by observing peers) and reinforcement (i.e., learning new behaviors and tendencies based on positive peer feedback). R&W explicitly addresses these two mechanisms in exercises and reflections aiming to subsequently improve the peer context in the classroom, represented by interpersonal relations in the class and victimization by bullies (aspects of the interpersonal domain). When classmates' modeling and reinforcement of deviant behaviors decrease and those of prosocial behaviors increase, the overall peer context in the classroom might become more positive (e.g., Dishion & Tipsord, 2011; Telzer, Van Hoorn, Rogers, & Do, 2018). In addition, I examined whether classmates' influences on the peer context was dependent on classmates' dyadic mutuality (i.e., the level of responsiveness, reciprocity, and shared understanding). Dyadic mutuality can represent the relationship quality in various relations (Piehler & Dishion, 2007), including relations between classmates. Taking relationship quality into account is eminent when examining peer influences as research suggests that students are more strongly influenced by peers with whom they have high quality relationships (Barry & Wentzel, 2006; Berndt, 2002; Piehler & Dishion, 2007).

Previous research studying peers' influences often focused on deviant *or* prosocial behavior within a specific type of relationship (e.g., Dishion, Spracklen, Andrews, & Patterson, 1996) and mostly relied on questionnaires (e.g., Hofmann & Müller, 2018). I examined both deviant and prosocial influences enabling a comparison between the two types of influences. Furthermore, modeling and reinforcement were assessed through video-observations of randomly selected dyads of classmates in a subsample in the "Standard" (i.e., the entire teaching staff is involved in the intervention) and Control (i.e., Care As Usual; CAU) conditions in the first year of the intervention. The randomly selected dyads represented the broad range of types of relations that exists between classmates, minimizing selection effects and enabling studying classmates'

influences on the collective classroom level (Busching & Krahé, 2020). Given that I was interested in peer influences at group level (i.e., in the classroom), I analyzed the data with multilevel mediation models in which effects *within* individuals can be separated from effects *between* clusters (i.e., classrooms; Preacher, Zyphur, & Zhang, 2010).
The second way I studied working mechanisms is through intervention components. Components of interventions are aspects associated with the content, instructional method, or structural characteristics of the intervention (Boustani et al., 2015; Lee et al., 2014) that affect participants' change. I examined all three types of components: Content, instructional, and structural components. Content components represent specific skills that are taught, such as emotion regulation (e.g., Boustani et al., 2015). Instructional components represent methods used for information delivery, such as practice during the lessons (e.g., Boustani et al., 2015). Structural components represent the structure of interventions, such as the number of sessions (e.g., Lee et al., 2014). Knowing which components are necessary to establish positive intervention effects could enhance more efficient and cost effective implementation of interventions by implementing effective and eliminating ineffective components (Michie et al., 2009). Optimization of interventions is particularly important for universal school-based interventions aiming to stimulate competencies and prevent problems in the intra- and interpersonal domain, since these interventions tend to show only small positive effects (Durlak et al., 2011). Additionally, schools have limited time and resources to invest in interventions, emphasizing the need to implement optimized interventions. Therefore, I conducted a meta-analysis to identify which components were related to stronger, or weaker, intervention effects of universal school-based interventions fostering students' development in the intra- and interpersonal domain. The results of this meta-analysis may help schools to make more informed decisions about which intervention to implement and can catalyze hypotheses generation in research regarding potentially interesting components (Wilson & Lipsey, 2007).

Outline of the Present Dissertation

To fill the gaps in the current literature, the present dissertation aims to gain insights into the effectiveness of school-based interventions stimulating competencies and preventing the development of problems in the intrapersonal and interpersonal domain. To this end, I examined 1) the effectiveness of these universal school-based interventions (i.e., what works?), 2) whether heterogeneity in the contexts and in the student population affected intervention effects (i.e., under what circumstances and for whom does it work?), and 3) working mechanisms of universal school-based interventions (i.e., how does it work?).

Regarding the first aim, the study protocol of the RCT to evaluate R&W is described in detail in Chapter 2. This chapter uses the initial concepts "socio-emotional

adjustment" and "social safety" which I later refer to as the "intrapersonal domain" and the "interpersonal domain", respectively. The results of the effectiveness of R&W, compared to CAU, are reported in Chapter 3 and the effectiveness of universal school-based intervention in general in Chapter 6. Concerning the second aim, Chapter 3 shows whether the intervention's ecological width and intervention's time span affected intervention effects. The findings to what extent students' personality affected intervention effects are presented in Chapter 4. With respect to the third aim, Chapter 5 reports whether classmates' modeling and reinforcement mediated the effect of R&W on the peer context in the classroom. Chapter 6 presents which components were related to stronger and weaker intervention effects of universal school-based interventions in general. Last, Chapter 7 discusses the main findings and put the findings of R&W in the general framework of universal school-based interventions provided by the meta-analysis.

Mertens, E. C. A.,
Deković, M.,
Van Londen, M.,
Reitz, E.
(2018).

BMC Psychology, 6, 1-12.

Author Contributions

EM, MD, MvL, and ER conceptualized the study. EM coordinated the data collection, analyzed the data and wrote the manuscript. All authors provided feedback on the study.

Acknowledgments

This study is financially supported by a grant from the Netherlands Organization for Health Research and Development (ZonMw Grant No. 531001106).

2

The Effectiveness of Rock and Water in Improving Students' Socio-Emotional Adjustment and Social Safety: Study Protocol for a Randomized Controlled Trial

Abstract

Background: Students following a low education track have an increased risk for developing problem behaviors. Rock and Water is a widespread, but still poorly evaluated, intervention that aims to improve students' socio-emotional adjustment and social safety. The aims of this study are to evaluate (1) the effectiveness of Rock and Water on socio-emotional adjustment (i.e., psychosocial wellbeing, sexual autonomy, and resilience) and social safety (i.e., perceived social security in the classroom, aggression, and bullying) and to examine (2) moderators and (3) mediators of its effects.

Methods: Schools are randomly assigned into four conditions: 'Light' (a core team of teachers is trained), 'Standard' (a core team of teachers and the whole school team is trained), 'Plus' (a core team of teachers, the whole school team is trained, and parents are involved), or 'Control condition' (Care As Usual). We aim to include 180 7th Grade students in each condition ($N = 720$) across all waves. A multi-informant (i.e., students, parents, and teachers) approach is used to assess the outcomes (socio-emotional adjustment and social safety), moderators (student, trainer, and parent characteristics) and mediators (self-control, self-reflection, self-esteem, and emotion regulation). Video-observations will be analyzed in a subsample to study the possible mediating effect of changes in deviant and prosocial communication among students on the effect on social safety.

Discussion: This project will provide information on the effectiveness of (different levels of school and parental involvement in) Rock and Water, which can be used by schools to decide upon the most efficient way to improve the care for the students. We will be able to shed more light on what works for whom and the working mechanisms of Rock and Water.

Trial registration: Dutch Trial Registration number 6554, registered on the 3rd of July 2017. The design of this study was approved by the Ethical Committee of the Faculty of Social and Behavioral Sciences of Utrecht University (FETC17-015). This study is financially supported by a grant from The Netherlands Organization for Health Research and Development, grant number 531001106.

Keywords: Rock and Water; intervention; socio-emotional adjustment; social safety; effectiveness; Randomized Controlled Trial

Background

In the Netherlands, secondary education (starting at age 12) consists of three education tracks: Preparatory vocational education track, preparatory college track, and preparatory university track. These different education tracks are attended by respectively 43%, 28%, and 29% of the total student population of secondary education (Central Bureau for Statistics, 2017). Students following the preparatory vocational education track (prevocational students) show less autonomy, less academic performance, and less school motivation and commitment than students in the other two tracks (Kuyper, Keuning, & Zijsling, 2010). They have an increased risk for psychological problems, such as substance abuse and early sexual intercourse (Harakeh, De Looze, Schrijvers, Van Dorsselaer, & Vollebergh, 2012; Schrijvers & Schuit, 2010), compared to students following the other two tracks. For instance, of the prevocational students about 17% smokes, 48% binge-drinks, and 25% has sex under the age of 17, whereas of students following the preparatory college or university track respectively 5% and 2% smokes, 37% and 25% binge-drinks, and 12% and 6% has sex under the age of 17 (Schrijvers & Schuit, 2010). Due to the high level of problems prevocational students might encounter, it is important to positively stimulate their development. Nowadays schools often implement various programs to obtain such aims, especially since the government requires schools to execute a policy to improve students' socio-emotional adjustment and social safety within the schools.

Rock and Water (R&W; Ykema, 2002; 2014) is one of such programs. It is a universal school-based intervention that aims to improve students' socio-emotional adjustment and social safety by increasing their self-control, self-reflection, self-esteem, emotion regulation and communication skills. R&W uses a psychophysical approach, that is play and exercises are used to increase the strength of youth, to teach them to make (physical) contact with others and to explore, respect and set own and other's boundaries. The name of the intervention is based on the symbolic principles of 'rock' and 'water'. Rock indicates a rigid and uncompromising attitude: Sticking to your own opinion and not bending for the opinion of others. Water on the other hand represents flexibility and cooperation: Being aware of one's own opinion, thoughts and feelings and being open to those of others at the same time, willing to cooperate with them. R&W addresses multiple themes including relaxation, self-control, physical and verbal communication, body language, assertiveness, group pressure, sexuality, and sexual violence.

Especially the psychophysical approach makes R&W eminently suitable for prevocational students. These students are less intrinsically motivated for school and (verbal) learning than students of other education tracks (Kuyper et al., 2010) and their cognitive abilities are, generally, lower (Ter Vrugte et al., 2015). Learning through play and exercise increases their motivation and is cognitively less demanding than a verbal cognitive approach. Hence, R&W fits in well with the learning style of prevocational students.

R&W is implemented in many countries (e.g., Australia, China, Singapore, France, The Netherlands). Despite this broad implementation, little information about the effectiveness of this intervention is available. Results of several small-scaled studies showed that participants feel more resilient, experience a more positive identity and use more active than passive coping styles after completing R&W (see for overview Ykema, Hartman, & Imms, 2006). Additionally, a recent study found that boys' self-reported coercive strategies and verbal manipulation decreased and their self-regulation and general efficacy increased after following the intervention (De Graaf, De Haas, Zaagsma, & Wijssen, 2015).

Notwithstanding these promising results, these studies have several limitations. First, most studies included boys only. It is unknown what the effects are for girls, for youth with different ethnic backgrounds and for prevocational students in specific. Second, only two measurement points were used, prior to the intervention and immediately after the intervention. Changes during the intervention as well as the long-term effects of R&W were not examined. Third, outcomes were narrowly defined (e.g., sexual aggression; Ykema et al., 2006) which makes it unclear what the effects are on the broader concept of socio-emotional adjustment and social safety. Fourth, often only one informant, mostly the adolescents themselves, participated and only one type of data collection was used, mostly questionnaires. Lastly, the studies did not examine potential moderators or mediators. Hence, no information is available about differential effectiveness for certain subgroups and the working mechanisms of R&W.

To increase the knowledge about the effectiveness and working mechanisms of R&W we will conduct a Randomized Controlled Trial (RCT) in which we plan to overcome the above mentioned limitations. We aim to include prevocational boys and girls from different ethnic backgrounds, conduct measurements prior, during, immediately after R&W and six months later, assess a broad range of outcomes (i.e., socio-emotional adjustment and social safety), use students, parents, and teachers as informants, utilize questionnaires, a computer task, and video-observations, and study potential moderators and mediators.

The first aim of the RCT is to examine the effectiveness of R&W in improving students' socio-emotional adjustment (i.e., psychosocial wellbeing, sexual autonomy, and resilience) and social safety (i.e., perceived social security in the classroom, aggression, and bullying). We will study R&W in three different conditions and compare it to a control group receiving Care As Usual (CAU; i.e., current school policy to enhance socio-emotional adjustment and social safety of students). We hypothesize that R&W will improve students' socio-emotional adjustment and social safety and will outperform CAU. The experimental conditions differ in the number of parties that are involved in R&W, that is a core team of teachers, a core team of teachers and the whole school team, and a core team of teachers, the whole school team and parents. The more parties are involved, the broader the ecological focus of the intervention will be. According to the social ecological model of Bronfenbrenner (1979) behavior is determined by the interaction of multiple systems (i.e., the individual, family, school). The involvement of multiple systems in the intervention could increase the

effectiveness of R&W. Although, this positive effect of involvement of multiple systems is not always found (e.g., Durlak, Weissberg, Dymnicki, Taylor, & Schellinger, 2011). We will examine if R&W is more effective the more parties are involved, as suggested by Durlak and colleagues (2011). We hypothesize that students' improvements will be more evident the more parties are involved.

The second aim is to examine potential moderators of the effect of R&W. Several student characteristics will be examined as moderators: Gender, ethnicity, and personality. In the R&W program the emphasis is on physical exercises which might lead to differential effectiveness based on gender and ethnicity. There are differences in the levels of daily physical activity between boys and girls with girls being less physically active (e.g., Minges, Chao, nam, Grey, & Whittemore, 2015; Patnode et al., 2010). Girls appear to experience higher barriers for physical activities than boys. They are more afraid of being chosen last for a team and of being embarrassed (Patnode et al., 2010). Therefore, we hypothesize that R&W is more effective for boys.

Also between ethnicities there are differences in the amount and sort of physical activities adolescents engage in. For instance, Black and Asian adolescents are less physically active and show more sedentary behavior than White adolescents (Minges et al., 2015, Brodersen, Steptoe, Boniface, & Wardle, 2007). White girls are more likely to be active in sports teams than Hispanic girls whom are more likely to be active in walking for transportation or physical activities at home such as household chores (Kelly et al., 2010). Kelly and colleagues (2010) suggest to tailor physical activity programs based on ethnicity due to differences in factors related to physical activation. For example, White girls appear to have higher levels of self-efficacy related to physical activities which makes them more physically active than Black and Hispanic girls. It might be that White students are more familiar with the sort of physical exercises used in the R&W program than students with other ethnicities. These White students might be less out of their comfort zone due to which they can focus more on the other aspects of the intervention. Thus, we hypothesize that R&W is more effective for White students than for students with other ethnicities.

Previous intervention research has shown that personality can be a moderator of intervention effectiveness. Senf and Liau (2013) have found the strongest intervention effects on happiness for individuals with high levels of extraversion and openness and on depressive symptoms for individuals with high levels of extraversion. Asscher and colleagues (2016) have shown that Multisystemic Therapy was less effective than treatment as usual in decreasing rule-breaking behavior for adolescent with low levels of agreeableness. Stoltz and colleagues (2013) have found the strongest intervention effects on reactive aggression for children with a low level of extraversion and on proactive aggression for children with less extreme levels of conscientiousness. Hence, it is expected that personality also influences the effectiveness of R&W. It is hypothesized that R&W is most effective for students with higher levels of agreeableness and openness, and average levels of conscientiousness. No specific hypothesis concerning extraversion is stated, since the results (Senf & Liau, 2013; Stoltz et al., 2013) are inconsistent.

Besides the student characteristics, we will also study trainer characteristics: Gender, ethnicity, education level, level of self-perceived competence, expertise, and degree of training and supervision. Findings concerning characteristics of professional therapists are inconsistent. Whereas some studies have found significant impact of (one of) these characteristics on treatment outcome (e.g., Anderson, Ogles, Patterson, Lambert, & Vermeersch, 2009; Bryan, Dersch, Shumway, & Arrendondo, 2004; Wheeler, & Richards, 2007), other studies have found no relation (e.g., Huppert et al., 2001; Wampold & Brown, 2005). Characteristics of (non-professional) therapists are often neglected in intervention studies. By studying trainer characteristics of non-professional therapists as moderators we will clarify the role of these characteristics on the effectiveness of R&W. We hypothesize that R&W is more effective when trainers are males, have a Western ethnic background, have a higher education level, have more self-perceived competence, have more expertise, and received more training and supervision.

Additionally, we will analyze parental characteristics as moderators: Parental sense of parenting competence and positive parenting (i.e., parental warmth and monitoring). Parents with a high sense of competence feel capable and adequate in interactions with their child (Deković et al., 2010). It might be that due to this confidence parents are more susceptible for information about the strategies and 'language' of R&W and apply these at home. Positive parenting enhances adolescents' social competence (Taylor, Conger, Robins, & Widaman, 2015) and the parent-child relationship (Pinquart, 2013). It might be that these adolescents feel more competent to incorporate R&W into their daily lives and tell their parents about R&W within that close and trusting relationship. Parents learn more about this intervention which enables them to also apply R&W. Therefore, we hypothesize that R&W is more effective for students with parents with high levels of parental sense of competence and positive parenting.

The third aim is to study the working mechanisms of R&W. Self-control, self-reflection, and self-esteem will be examined as mediators of the effect of R&W on students' socio-emotional adjustment and social safety. Self-control, self-reflection, and self-esteem are theorized as the three pillars through which R&W aims to establish the desired developments (Ykema, 2002). Additionally, emotion regulation will be examined as mediator. According to the theory behind R&W, students will become better aware of the physical representations of their emotions, for instance muscle tensions and a-rhythmic breathing. Raising students' emotional awareness is expected to facilitate them to perceive their emotions and regulate them. This improvement in emotion regulation would lead to an increase in their socio-emotional adjustment and social safety (Ykema, 2002). Analyzing these mediators enables us to test the theory of R&W. We hypothesize that R&W will increase students' self-control, self-reflection, self-esteem, and emotion regulation which, in turn, will enhance their socio-emotional adjustment and social safety.

Furthermore, deviant and prosocial communication (i.e., verbal and non-verbal) will be analyzed as mediators of the effect of R&W on social safety. Communication is a recurrent theme throughout the intervention. It is proposed in the theory that

improving the communication of students increases their feelings of social safety, as they learn to show that they care about someone's feelings, that they are open to others and can become closer to each other (Ykema, 2002). Therefore, we hypothesize that R&W decreases deviant communication and increases prosocial communication which improves students' social safety.

In sum, in this study we will examine the effectiveness of R&W in improving students' socio-emotional adjustment and social safety. Moreover, we will examine what works for whom by studying characteristics of students, teachers and parents. Additionally, the working mechanisms of the intervention will be analyzed.

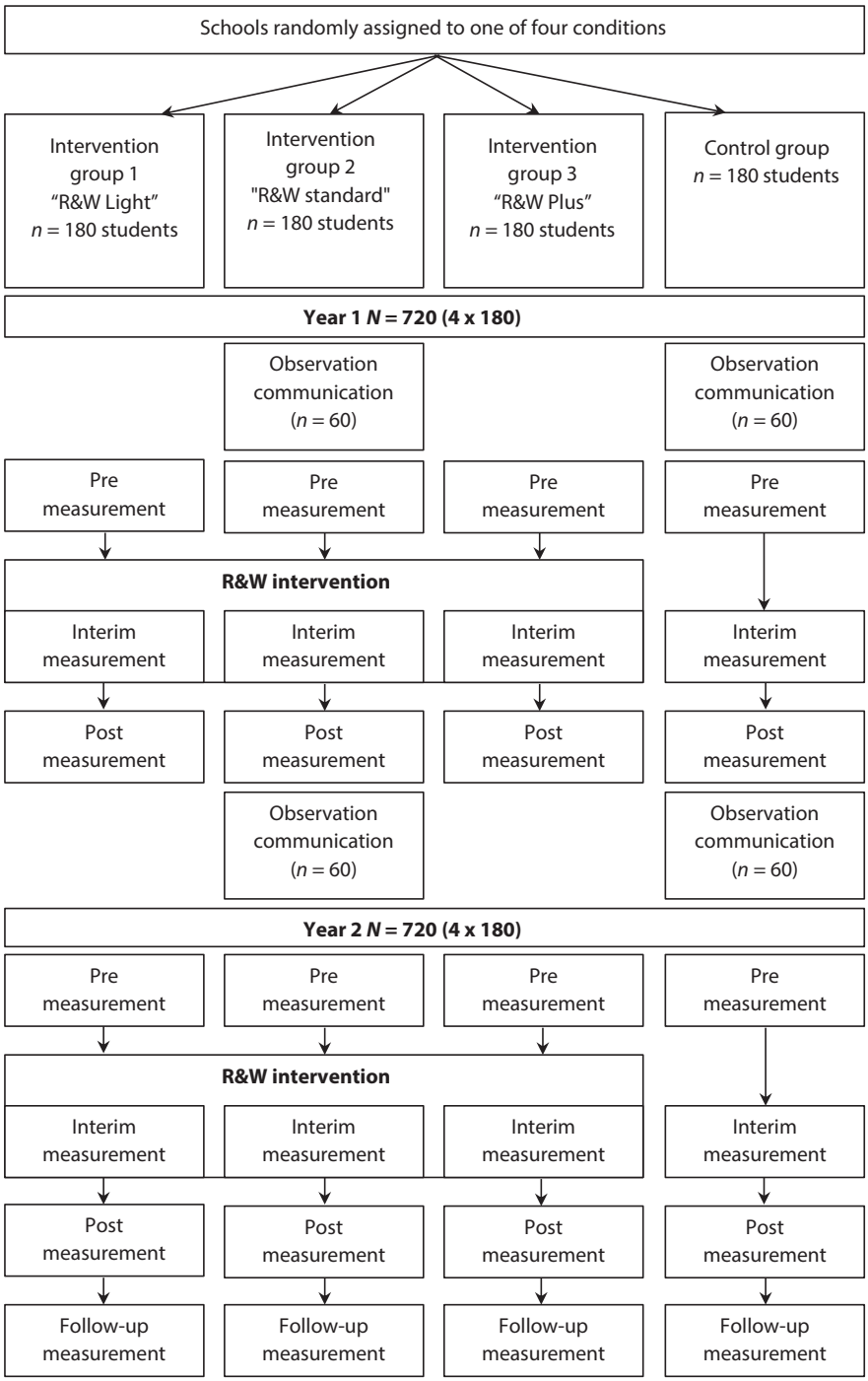
Methods

Design

To study the effects of R&W an RCT design is implemented in the 7th Grade of preparatory vocational education level. Schools from different parts of The Netherlands are randomly assigned to one of three intervention groups or to the control group (Figure 1). In the 'Light' condition a core team of teachers is trained with the three-day training course to become certified R&W trainers and implement the R&W program. In the 'Standard' condition a core team is trained to become R&W trainers and the rest of the school-team that teaches 7th Grade students follows a three-day training to learn how they can support the R&W trainers of their school and how they can apply R&W in their regular classes. The 'Plus' condition is equal to the 'Standard' condition with the addition of a parent component. The parents watch a documentary of R&W, get an invitation to join a lesson in the school, receive weekly e-mails with information about the lesson of that week and are stimulated to act on this information, for instance by communicating about R&W or using R&W language. This will not only create a supporting environment within the school, but also in students' homes. In the control condition students receive CAU; i.e., current school policy to enhance socio-emotional adjustment and social safety of students.

Students receive R&W for two years. The intervention is implemented during physical education classes by (mostly) physical education teachers who followed the three-day training. After successfully terminating this training, these teachers are certified R&W trainers. Socio-emotional adjustment (i.e., psychosocial wellbeing, sexual autonomy, and resilience), social safety (i.e., perceived social security in the classroom, aggression, and bullying), and mediators (i.e., self-control, self-reflection, self-esteem, and emotion regulation) will be assessed at multiple measurement points: Prior, during and after R&W in the first and second year and at follow-up. These concepts are assessed with online questionnaires completed by students, parents,

Figure 1. Flow chart.



and teachers. The interim measurements (during R&W) are shortened questionnaires and are completed by students after a series of three R&W lessons. Deviant and prosocial communication is assessed in the first year prior and after R&W using video-observations in a subsample. Moderators (i.e., student, teacher, and parent characteristics) are assessed at one time point.

The study is registered with the Dutch Trial Register (6554) and has been approved by the Ethical Committee of the Faculty of Social and Behavioral Sciences of Utrecht University (FETC17-015).

Study sample

In total, we aim to include 720 students in this study across all waves, 180 per condition. Participants are students in the 7th Grade, in the second year of the intervention in the 8th Grade, of preparatory vocational education level. Schools are excluded if they currently implement R&W in the whole school or have implemented the intervention in the past two years in the whole school. Additionally, schools for students with special needs are excluded from the study.

Recruitment

Schools were recruited through the network of the developers of R&W, the Gadaku Institute. An e-mail to certified R&W trainers was sent with information about this study and a message on the private R&W online forum was posted. Trainers asked within their network whether schools that did not implement R&W were interested in participating in this study. When schools were interested they contacted the Gadaku Institute which referred the school to the researchers. The researchers established whether the school was eligible for participation. If the school was eligible, the researcher provided additional information to the schools based on which the school made a decision concerning participation.

Schools were randomly assigned (1:1:1:1 ratio) by stratified block randomization, with blocks of four (i.e., the number of conditions in this study). Schools were stratified by school size (small to moderate sized schools: < 100 students in 7th Grade, large sized schools: > 100 students in 7th Grade) to establish a more equal distribution of students over the four conditions. The randomization numbers were generated by a random number generating computer program.

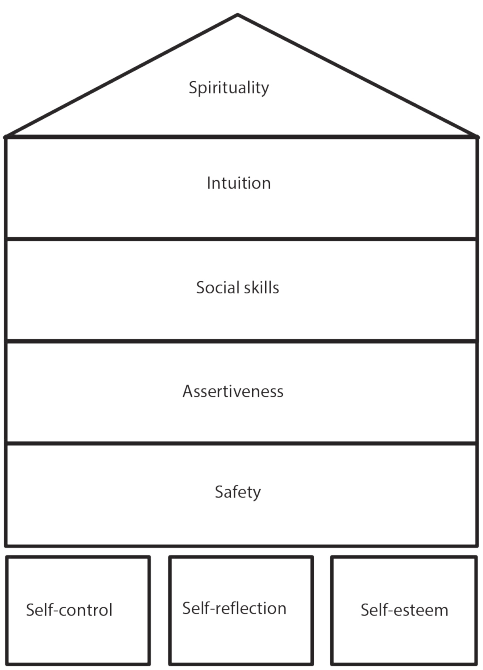
Students, parents, teachers, and R&W trainers received an information letter to inform them about R&W and the study. Parents also attended a parent-teacher evening where they received additional information about R&W. Students gave active informed consent and parents passive informed consent for participation of the student. Parents gave active informed consent for their own participation in the study. Active informed consent was also acquired for participation of teachers and R&W trainers.

Conditions

Rock and Water. The R&W program is based on the theory of 'the Rock and Water house' (Figure 2). According to this theory there are five levels of the house (themes) that are discussed during the intervention: Safety, assertiveness, social skills, intuition, and spirituality. During the theme safety, feeling safe at home, at school and in society is discussed. Safety is important for students' development and to find their own way. With the theme assertiveness students learn to deal with difficult situations without losing self-control. The third level social skills emphasizes the importance of communication in our contemporary, multicultural society. Intuition is discussed to make students aware of their preferences and choices, made by intuition, that are determinants for their lives based on their qualities, talents and possibilities. The roof of the house (fifth level) is spirituality. In this theme students learn to follow their own path and gain insight in themselves.

The foundation of the R&W house is built on three pillars: Self-control, self-reflection, and self-esteem. The R&W theory states that acquiring these three skills forms the basis for further development concerning the five themes. By learning to control and direct their energy (self-control), students are able to reflect on their behavior and the consequences (self-reflection). Knowing that they can control, reflect upon and potentially change their actions, students' self-esteem increases; they know what they are capable of and what they want.

Figure 2. The Rock and Water house.



The themes are handled and discussed using a psychophysical approach within a safe, supportive and respectful environment. R&W emphasizes the safe and supportive environment where students are allowed to make mistakes. Such an environment is expected to encourage students' exploration and learning, addressing personal needs and problems, and establishing positive relations. Respect within the group and between students is created based on the principles and ethics of martial arts on which the physical exercises are based. Students learn basic self-defense skills in which it is supposed that they train their self-control, learn what to do in violent situations and work together respectfully. Additionally, they practice their non-verbal communication by making eye contact and reading the body language of their partner during the exercises.

In the exercises, special attention is paid to students' body awareness and level of arousal. R&W aims to make students aware that physical signs such as tensions in the muscles, a-rhythmic breathing, and increased heartrate are expressions of emotions, stress and feelings. It is expected that during R&W they learn how to actively relax, find the source of the tensions and connect to their self-control and emotion regulation. This is explained using the terms 'rock' and 'water'. 'Rock' indicates a physical tense and firm position with low breathing (i.e., Being aware of one's own opinion, thoughts and feelings and being able to resist pressure from others). 'Water' implies a relaxed but alert position with low breathing (i.e., Knowing one's own opinion, thoughts and feelings and open to those of others). Students are encouraged to experience that the 'rock' and 'water' attitudes are on a continuum which they can use to regulate their behavior and interactions.

The R&W program is implemented in the schools for two years. The program starts in 7th Grade with 14 weekly lessons of one and a half hour. In 8th Grade students receive 8 weekly lessons of one and a half hour (see Table 1 for an overview of the topics per lesson). The lessons are given to the class as a whole in large spaces such as a gym hall or drama classroom. The trainers are (mostly) physical education teachers of the schools whom have participated in the three-day training course provided by the Gadaku Institute. These trained and certified teachers are monitored and supervised during the study by coaches of the Gadaku Institute. These coaches have at least once a week contact with the R&W trainers in the schools by e-mail, telephone, or face-to-face depending on the needs of the trainers. They answer questions and give advice regarding the R&W program. Additionally, the coaches observe a lesson of the trainer and provide feedback based on that observation.

During the lessons, the trainers explain, demonstrate and monitor the exercises indicating what behavior is desired and what is not by giving feedback. The skills and techniques learned in the exercises are repeated over the lessons to enable integration and internalization of these skills in students' thinking and acting. After an exercise, students reflect in the group upon what they have learned from the exercise and how they could integrate this in their daily lives. Each lesson is ended with a physical condition exercise such as push-ups or jumping rope. In the first lessons, students set individual goals for themselves concerning these physical condition exercises. At the

end of the R&W lessons, in 7th Grade and in the 8th Grade, they evaluate if their goals are met or not. After each lesson, students complete homework assignments in their workbook. In this workbook a summary of the lesson and some questions are given to the students. This aims to stimulate the transfer of the taught skills to their daily lives and to practice the skills outside of the lessons.

The physical exercises are performed with different partners which is thought to increase the coherence of the group in which the R&W program is implemented. Students learn how to work and play together as a class. During the exercises they have to control their strength, set boundaries and be respectful towards each other. It is expected that learning how to connect and listen to others while being calm and keeping their strength, they improve their social skills and increase the social support within the class.

Control condition. Students will receive CAU: Current school policy to enhance socio-emotional adjustment and social safety of students. In The Netherlands, all schools have such a policy, since it is obligatory. However, the operationalization of the policy can vary widely. Some schools have a program or training other than R&W (e.g., lifestyle lessons), other schools have a policy on management level (e.g., anti-bullying contracts with students).

Table 1 Overview of the Topics of the R&W Lessons in Year 1 and in Year 2

Year 1 (7th Grade)		**Year 2 (8th Grade)**	
Lesson	**Topic**	**Lesson**	**Topic**
1	Standing strong (Relaxation)	1	Refresh skills year 1
2	Standing strong together (Helping each other in confrontation)	2	Breathing (Relaxation)
3	Physical and mental pressure	3	Body language
4	Bullying (Ignoring)	4	Peer pressure and bullying
5	Bullying (Walking away)	5	Peer pressure (Sexual autonomy)
6	Verbal communication	6	Responsibility and making own choices
7	R&W in school	7	Sexual autonomy (Boundaries)
8	Breathing (Relaxation)	8	Positive thinking and visualization
9	Body language		
10	Personal contact		
11	Experiencing, respecting, and setting boundaries		
12	Experiencing, respecting, and setting boundaries		
13	Intuition		
14	Dealing with intimidating group		

Instruments

An overview of the concepts, instruments, measurement points and informants is presented in Additional file 1, Table 2.

Socio-emotional adjustment

To assess psychosocial wellbeing students, parents and (non-trainer) teachers complete the short version (12 items; e.g., "I am stubborn.") of respectively the Youth Self Report (YSR; Achenbach, 1991b; Verhulst, Van der Ende, & Koot, 1997b), the Child Behavioral Checklist (CBCL; Achenbach, 1991a, Verhulst, Van der Ende, & Koot, 1996) and the Teacher's Report Form (TRF; Achenbach, 1991c; Verhulst, Van de rEnde, & Koot, 1997a) based on the study of Chorpita and colleagues (2010). Additionally, students and parents fill in the subscale Psychological wellbeing (7 items; e.g., "Did you had fun?") of the KIDSCREEN-27 (Ravens-Sieberer & The European KIDSCREEN Group, 2006).

Sexual autonomy is reported by the students. For this, items from a national study in The Netherlands concerning sexual health (De Graaf, Meijer, Poelman, & Vanwesenbeeck, 2005) are used. These 5 items represent interaction competence concerning control, assertiveness and self-esteem (e.g., "I have little influence on what happens.").

To measure resilience, students, parents, and (non-trainer) teachers complete the Connor-Davidson Resilience Scale short version (CD-RISC 10; Campbell-Sills & Stein, 2007). The 10 items reflect the self-beliefs to cope with difficulties in life (e.g., "Able to adapt to change.").

Social safety

Perceived social security in the classroom is measured using the subscales Comfort (4 items; e.g., "In this class, I can be myself."), Conflict (4 items; e.g., "In this class, children argue with each other."), and Cohesion (4 items; e.g., "In this class, everyone likes each other.") of the Classroom Peer Context Questionnaire (Boor-Klip, Segers, Hendrickx, & Cillessen, 2016) completed by students and (non-trainer) teachers. It assesses the perception of school culture, for instance how positive, respectful, friendly and helpful students are towards each other and sense of belonging.

Aggression is measured with the Reactive and Proactive Aggression Questionnaire (REPRO; Dodge & Coie, 1987; Hendrickx, Crombez, Roevers, & Orobio de Castro, 2003). It assesses reactive (3 items; e.g., "If they tease or threaten me, I get angry.") as well as proactive (3 items; e.g., "If I do not like a child, I will bully him with others.") aggression. Students, parents, and (non-trainer) teachers complete this questionnaire.

To measure bullying, students complete the 2 global items of the Olweus Bully/ Victim Questionnaire (Solberg & Olweus, 2003). It measures the frequency of bullying and victimization. Additionally, students complete brief sociometric nominations assessing social acceptance, popularity and classmates' roles concerning bullying.

Moderators

Students' gender and ethnicity and trainers' gender, ethnicity, education, expertise, and degree of received training and supervision are assessed with questionnaires developed for this study.

Students' personality is reported by the student and parent using the Quick Big Five (Goldberg, 1983). It consists of 30 items (i.e., characteristics; e.g., nice, sympathetic, organized) on which the informant can indicate to what extent that characteristic suites the participant.

Parental sense of parenting competence is assessed with the subscale Competence of the Parenting Stress Index (PSI; Abidin, 1983) completed by the parent. It measures the degree to which parents feel they are capable enough and have enough skills to cope with their child. The subscale contains 8 items (e.g., "Raising my child is harder than I expected.").

Positive parenting will be measured using the subscales Warmth and Monitoring from the Co-parenting Behavior Questionnaire (CBQ; Schum & Stolberg, 2007) completed by the parent. The subscale Warmth (7 items; e.g., "I spend time doing fun things with my child."). measures the extent to which parents show parenting behavior to make their children feel comfortable, accepted and approved. The subscale Monitoring (5 items; e.g., "I know my child's after school activities.") measures parental awareness of different aspects of the children's life.

Teacher's sense of competence will be assessed with the subscale Self-efficacy for management of the Teachers' sense of self-efficacy (Wolters & Daugherty, 2007) completed by the R&W trainer. This subscale measures teachers' confidence in their skills to effectively manage their classroom. It contains 6 items (e.g., How much can you do to control disruptive behavior in the classroom?").

Mediators

Self-control is assessed with the Self-Control Scale short version (Finkenauer, Engels, & Baumeister, 2005) completed by the student. It contains 11 items (e.g., "I wish I had more self-discipline."). This questionnaire measures students' ability to change their inner responses, interrupt undesired behavioral impulses and abstain from acting on these tendencies. Additionally, halfway the questionnaire students complete a shortened version (19 items) of a delayed discounting computer task to measure self-

control including a 'catch' question (Kirby, Petry, & Bickel, 1999; Meyerson, Baumann, Green, 2014; 2017). Students can choose a smaller, immediate reward or a larger, delayed reward (e.g., "Would you prefer to receive €54 today or in 117 days €55?"). To ascertain that the students have read the questions, a catch question is added, similar in form, amount and delay: "Would you prefer to receive €59 today or in 139 days €21?".

Self-reflection will be reported by the students using the Engage in reflection subscale of the Self-Reflection and Insight Scale (SRIS; Sauter, Heyne, Blöte, Van Widenfelt, & Westenberg, 2010). It contains 6 items (e.g., "I don't often think about my thoughts.").

Self-esteem will be measured with the subscale Global self-perception of the Self-perception Profile (Harter, 1988) reported by the students. This subscale contains 5 items (e.g., "I'm often disappointed in myself.").

Emotion regulation is measured using the subscales Impulse control (6 items; e.g., "When I'm upset, I feel out of control.") and Strategies (8 items; e.g., "When I'm upset, I believe that I will remain that way for a long time.") from the Difficulties in Emotion Regulation Scale (DERS; Anderson, Reilly, Gorrell, Schaumberg, & Anderson, 2016) completed by students. These subscales measure students' ability to control their emotional impulses and the regulation strategies they apply.

Deviant and prosocial communication is assessed using video-observations of same-sex dyads of classmates in a subsample of students in the 'Standard' and control condition. This observation task is based on the Peer Interaction Task (Dishion, Andrews, & Crosby, 1995). The dyads plan an activity together, as warm-up, and subsequently discuss 3 situations concerning daily school situations. Each of these 4 segments lasts 5 minutes. The 20 minute interactions are videotaped and coded. Deviant and prosocial communication is coded based on the Conversation Topic Code (Piehler & Dishion, 2004; Van de Bongardt et al., 2017) and communication ratings (Piehler & Dishion, 2004; Dishion et al., 1989; Whalen, Henker, Collings, McAuliffe, & Vaux, 1979).

Treatment adherence

To assess treatment adherence, the trainer indicates after a series of three lessons which strategies were used, level of treatment adherence, and whether the lessons were completed. Furthermore, a subsample of lessons will be observed by an expert in R&W to assess treatment adherence, quality of delivery, participants' engagement, and adaptations. This coding schema is based on Bishop and colleagues (Bishop et al., 2014). Treatment adherence indicates the level to which the trainer has conducted the lesson as described in the manual (e.g., "Skipped the trainer exercises?). Quality of delivery is an indication of the general quality of the lesson (e.g., "Are the goals of the lesson met?"). Participants' engagement indicates the level to which the trainer

actively involves the students in the lessons and the extent to which the trainer can activate students physically (i.e., exercises) and mentally (i.e., reflection; e.g., "Do students respond to questions of the trainer?"). Adaptations are clear deviations from the manual. These can be adaptations to the exercises, structure of the exercise, instructions and adding steps to an exercise (e.g., "What percentage of the exercises of the lesson are adapted?").

Statistical analyses

The power calculations are based on the N:q rule for structural equation models (Kline, 2015). This rule states that for each free parameter (q) 10 to 20 participants (N) are needed. We took the conservative approach by taking 20 participants per free parameter for our power calculations. In our multigroup LGC model there are 9 free parameters per condition, 36 free parameters in total. Thus a total sample of 720 participants is needed for our analysis, 180 participants in each condition. Since not only students can drop out but also classes (about 20 students) and schools (about 60 to 90 students), we will include three to four schools per condition. Missing data will be handled in M*plus*.

In our data, students are nested in classes which are nested in schools. Therefore, we will examine whether there is significant intra-class correlation on one of the levels (i.e., school, and class) and we will calculate the design effect. Each level with a design effect larger than 2.0 will be modeled in the analyses which allows us to correct for the nested data, that is multilevel analyses (Muthén & Satorra, 1995).

The first aim is to examine the effectiveness of R&W in the conditions which differ in the number of parties involved in the intervention. This will be examined using an analysis of covariance (ANCOVA) for the outcomes of socio-emotional adjustment and social safety, in case the design effect is smaller than 2.0. The dependent variables will be the post-measurements after the second year (8th Grade), the independent variables the condition, and the covariates the premeasurements (7th Grade). If needed, due to large design effects, multilevel regression analyses will be used (This also holds for the other aims). Then, we will analyze the trajectories of change in socio-emotional adjustment and social safety during R&W with multigroup Latent Growth Curve (LGC) modeling in M*plus*. We will examine if these trajectories of change differ significantly between the four conditions.

The second aim, the effect of potential moderators on the effectiveness of R&W on socio-emotional adjustment and social safety, will be examined using ANCOVAs for categorical moderators and regression analyses for the continuous moderators. The interaction effects of the concerned outcome measure with the student, trainer or parent characteristics will be added as an interaction term.

The third aim, studying the working mechanisms of R&W, will be examined by analyzing multiple mediators. We will analyze whether the R&W intervention

improves students' self-control, self-reflection, self-esteem and emotion regulation by performing ANCOVAs. Furthermore, we will analyze whether the change in these concepts mediate the relation between R&W and socio-emotional adjustment and social safety through LGC modeling in M*plus*. We will model the mediators as well as the outcome measures in this mediation analysis on the assessments before, during, and after R&W (e.g., Deković, Asscher, Manders, Prins, & Van der Laan, 2012). Furthermore, we will study whether R&W decreases deviant communication and increases prosocial communication. We will analyze if changes in deviant and prosocial communication are mediators of the effect of R&W on social safety. We will model these indirect effects in M*plus* using bootstrapping.

Discussion

This study protocol presents the design of a study evaluating the effectiveness of R&W in increasing socio-emotional adjustment and social safety in prevocational students. Previous small-scaled research has shown promising results. With this study we try to overcome the limitations of previous studies by incorporating more measurement waves (i.e., prior, during, and after R&W in the first and second year, and at follow-up), and by using a multi-informant (i.e., students, teachers, parents) and multi-method (i.e., questionnaires, computer task, video-observations) approach. We will be in the unique position to not only examine the effectiveness of R&W in different levels of school and parental involvement, but also to study what works for whom, and the working mechanisms of the intervention.

A possible threat to the study that we foresee is reaching and engaging the different informants (i.e., students, teachers, parents). In general, we try to reduce this possible threat by organizing focus groups with participating schools. During these meetings we discuss the feasibility of our plans and ask for suggestions (e.g., How can we best reach parents?).

In particular, it might be that we ask students too many questions. They could lose their interest and concentration and eventually give answers without reading the questions to finish the questionnaire sooner. We try to diminish this loss of interest and concentration by digitalizing the questionnaire and by adding a computer task halfway the questionnaire. The digital questionnaire enables students to complete the questionnaire on mobile devices which is probably more interesting to the students than filling in the questionnaires on paper. The computer task makes the students feel like they are doing something different than filling in questionnaires, after which they have renewed energy to complete the second half of the questionnaire. Moreover, students can complete the questionnaire during school hours so that it does not cost them additional time.

Parents might feel disengaged from the study and do not complete the questionnaire. We try to minimize this possible threat by emphasizing the importance of it. Furthermore, both consent and the questionnaire are digital. Parents do not have to actively return a consent letter and can complete the questionnaire at home at a time convenient for them. In addition, we will send reminders for completing the questionnaire to parents through the school's parental communication system. Moreover, we will organize a raffle to motivate the parents to participate.

Teachers might not have enough time to complete the questionnaires during work hours. They have to fill in a questionnaire for each student which takes up a lot of time. We try to reduce this possible threat by asking multiple teachers per school to complete the questionnaires. Multiple teachers can decrease the burden of the questionnaires since they only have to complete the questionnaires for a subgroup of students (e.g., one class). Additionally, we have personal contact with the teachers completing the questionnaires. This enables us to directly approach a teacher from whom we have missing data.

Another possible threat is insufficient or low treatment adherence. Teachers might deviate from the manual or are not able to complete a lesson due to limited time. We try to gain insight in the quality of implementation by asking questions about treatment adherence every three lessons and by observations during a subsample of R&W lessons. This information can be taken into account in the analyses.

The current study offers the opportunity to examine whether the broadly implemented intervention R&W is effective in positively stimulating the socio-emotional adjustment and social safety of prevocational students. Furthermore, a deeper understanding of the intervention can be gained by studying moderators and mediators. When proven effective, the implementation of R&W in prevocational schools can be stimulated. Additionally, possible adjustments to increase the effectiveness of the intervention for certain subgroups (i.e., students characteristics) might be identified.

Supplementary Material

Table 2 Overview of Concepts, Instruments, Measurement Waves and Informants

Concept	Instrument	Measurement							Informant		
		Pre1	Int[1]	Post1	Pre2	Int[1]	Post2	Follow-up	St	T	P
Socio-emotional adjustment											
Psychosocial wellbeing	YSR/CBCL/TRF	X	X	X	X	X	X	X	X	X	X
	KIDSCREEN-27	X	X	X	X	X	X	X	X	X	X
Sexual autonomy	Items from "sex under 25"	X	X	X	X	X	X	X	X		
Resilience	CD-RISC 10	X	X	X	X	X	X	X	X	X	X
Social safety											
Perceived social security in classroom	Classroom Peer Context Questionnaire	X	X	X	X	X	X	X	X	X	
Aggression	REPRO	X		X	X	X	X	X	X	X	X
Bullying	Olweus Bully/Victim Questionnaire	X		X	X		X	X	X		
	Sociometric nomination	X		X	X		X	X	X		
Moderators											
Gender, ethnicity	Developed for this study	X							X	X	X
Education	Developed for this study	X								X	
Training and supervision	Developed for this study			X			X			X	
Competence	Teachers' sense of self-efficacy	X		X	X		X			X	
Expertise	Developed for this study						X			X	
Treatment adherence	Developed for this study		X			X				X	
	Observation[2]		X			X					
Personality	Quick Big Five[3]	X			X				X		X
Parental sense of competent	PSI	X									X
Positive parenting	CBQ	X									X
Mediators											
Self-control	Self-control Scale	X	X	X	X	X	X	X	X		
	Delayed discounting	X		X	X		X	X	X		
Self-reflection	SRIS	X	X	X	X	X	X	X	X		
Self-esteem	Self-perception profile	X	X	X	X	X	X	X	X		
Emotion regulation	DERS	X	X	X	X	X	X	X	X		
Deviant and prosocial communication	Adjusted Peer Interaction Task	X		X					X		

Note. Pre1 = premeasurement year 1; Int = interim measurements; Post1 = post-measurement year 1; Pre2 = premeasurement year 2; Post2 = post-measurement year 2; St = student; T = teacher (non-trainer or trainer); P = parent. [1]Interim questionnaires are shortened and only completed by the students. [2]Observation of treatment adherence will be conducted by an R&W expert. [3]Parents complete the Quick Big Five at pre1, students at pre2.

Mertens, E. C. A.,
Deković, M.,
Van Londen, M.,
Nye, E., & Reitz, E.

Manuscript submitted for publication.

Author Contributions

EM, MD, MvL, and ER conceptualized the study. EM coordinated the data collection, analyzed the data and wrote the manuscript. All authors provided feedback on the study.

Acknowledgments

This study is financially supported by a grant from the Netherlands Organization for Health Research and Development (ZonMw Grant No. 531001106).

3

Solid as a Rock, Flexible as Water? Effectiveness of a School-Based Intervention Addressing Students' Intrapersonal and Interpersonal Domains

Abstract

Background: Students following a preparatory vocational education track seem most in need of an intervention stimulating their competencies and preventing the development of problems in the intrapersonal and interpersonal domain. At the same time, these students are a challenging group to involve in interventions. The aim of the present study was to examine, first, whether a 2-year intervention using a unique combination of a physical approach with a more common psychological approach (i.e., a psychophysical approach) is effective in improving students' competencies and preventing problems in the intra- and interpersonal domain, and, second, whether the width of the ecological focus of the intervention influenced its effectiveness.

Methods: We conducted a Randomized Controlled Trial with a sample of 7th Grade students ($N = 1299$, $M_{age} = 12.38$, 54% boys). Students reported on outcomes of the intra- and interpersonal domains using digital questionnaires. The data was analyzed with Latent Growth Curve models.

Results: Results showed that the intervention was most effective when implemented with the most narrow ecological focus (i.e., only a core team of teachers involved in the intervention) and improved several outcomes in students' intrapersonal domain, interpersonal domain, and secondary outcomes (self-control and emotional self-regulation). The intervention effects were strongest, albeit small, in the first year of the intervention.

Conclusion: These results show that interventions with a psychophysical approach and a narrow ecological focus might be promising interventions for prevocational students, although effort should be put in increasing its effectiveness.

Trial registration: Dutch Trial Registration number NL6371 (NTR6554). Registered 3 July 2017.

Keywords: School-based intervention; randomized controlled trial (RCT); intrapersonal domain; interpersonal domain; implementation

Solid as a Rock, Flexible as Water? Effectiveness of a School-Based Intervention Addressing Students' Intrapersonal and Interpersonal Domains

Schools play an eminent role in fostering students' development in the intrapersonal and interpersonal domain (Pellegrino & Hilton, 2012). The intrapersonal domain refers to the ability to manage one's own feelings, emotions, and attitudes that pertain to the individual self, such as psychological wellbeing and internalizing behavior (Barber, 2005). The interpersonal domain refers to the ability to build and maintain positive relationships with others, to understand social situations, roles and norms, and to respond appropriately, such as interpersonal relations and aggression (Pellegrino & Hilton, 2012; Shek & Leung, 2016). Students can gain competencies in these domains by mastering relevant cognitive, affective, and social skills (e.g., identifying emotions, perspective taking) or can develop problems when mastery of (some of) these skills lacks or falls behind (Durlak, Weissberg, Dymnicki, & Taylor, 2011; Modecki, Zimmer-Gembeck, & Guerrà, 2017). Although the two domains influence each other, they are regarded as distinct domains. The intrapersonal domain reflects subjective personal functioning, predicting for instance academic functioning, whereas the interpersonal domain reflects social functioning, predicting for instance positive peer relations (Dufner, Gebauer, Sedikides, & Denissen, 2019; Park et al., 2017). Therefore, schools should intentionally cultivate their students' competencies and prevent development of problems in both domains (Pellegrino & Hilton, 2012).

Many schools implement universal interventions addressing students' intra- and interpersonal domains. However, these interventions show only small effects for students in general (see for a meta-analysis Mertens, Deković, Leijten, Van Londen, & Reitz, 2020) and might show even smaller effects for certain groups of students. In the Dutch secondary education system, starting at age 12 (7th Grade), there are three educational tracks: Preparatory vocational track (i.e., prevocational track), preparatory college track, and preparatory university track. More than half (54%) of the total secondary student population attends the prevocational track (Central Bureau for Statistics, 2020). These students in particular seem to be in need of school's stimulation of competencies and prevention of problems in the intra- and interpersonal domains as they report lower levels of wellbeing, more behavioral problems, more problems with peers (Stevens & DeLooze, 2018), and have an increased risk for psychological problems (Schrijvers & Schuit, 2010) compared to students in the other two educational

tracks. They can, however, be a challenging group to involve in school-based interventions. Generally, they show less autonomy, intrinsic motivation for school and (verbal) learning, and have lower cognitive capacities than students following the other two tracks (Timmermans, Naaijer, Keuning, & Zijsling, 2017). It seems pivotal for interventions that specifically target prevocational students to use an approach that fits with the learning processes of these students. Research recommends to create a learning environment in which short moments of instructions or reflections are alternated with practicing the new skills in a relevant context. Such an environment is suggested to foster students' understanding and use of the learned content (De Bruijn et al., 2005). Hence, an intervention approach that combines a psychological approach (i.e., instruction, reflection) with a physical approach (i.e., practice of skills) seems a promising approach for prevocational students.

A universal school-based intervention that uniquely combines a psychological approach with a physical approach is Rock and Water (R&W; Ykema, 2002; 2018). R&W is based on the theory of the 'R&W house' consisting of five modules. The first module addresses students' need to feel safe in order to change their behavior and develop themselves, targeting behaviors such as externalizing behavior, aggression, and bullying. The second module states that students need to learn to deal with difficult situations without losing self-control, targeting behaviors such as resilience, sexual autonomy, internalizing behavior, and victimization. The third module focusses on communication with others, targeting behaviors such as positive social interactions between classmates. The last two modules stress the importance of developing own preferences and choices to increase self-insight, targeting more general feelings of psychological wellbeing. The R&W house is built on a foundation of three pillars (i.e., self-control, self-reflection, and self-esteem). Closely related to these three pillars is emotional self-regulation. According to the theory of R&W, students' self-control, self-reflection, self-esteem, and emotional self-regulation increase during the intervention which enables students to develop themselves within each of the five modules.

Even though R&W is implemented in many countries (e.g., Australia, China, Singapore, France, the Netherlands), only little is known about the effectiveness of the intervention. Several small-scaled studies found that after completing R&W participants felt more resilient, experienced a more positive identity, and used more active than passive coping styles (Ykema, Harman, & Imms, 2006). A more recent study, examining the effectiveness of R&W on sexual aggression among prevocational boys in the 9th and 10th Grade, showed that self-reported coercive strategies and verbal manipulation decreased and self-regulation and efficacy increased after completing R&W (De Graaf, De Haas, Zaagsma, & Wijsen, 2015). Notwithstanding these promising results, a large scale study focusing on broader outcome measures and including girls as well as boys is needed to assess the effectiveness of R&W, and in specific its effectiveness for prevocational students. The first aim of the present study was therefore to examine whether R&W is effective in enhancing competencies and preventing problems in both the intrapersonal and interpersonal domain of prevocational students.

The second aim was to examine if intervention effects were influenced by the width of the intervention's ecological focus, i.e., the extent to which multiple systems are involved in the intervention. As stated in the social ecological model of Bronfenbrenner (1979), behavior is determined by the interactions of multiple systems such as the individual, family, and school systems. When more systems are actively involved in an intervention, students are more exposed to the intervention which could increase its effectiveness. For instance, Flay, Graumlich, Segawa, Burns, and Holliday (2004) examined the influence of involving multiple systems on the effectiveness of a social-emotional intervention aiming to reduce risk behaviors. They found stronger intervention effects when the intervention was not only implemented in the classroom, but also in the whole school and when parents were involved. However, this positive effect of a broad ecological focus of interventions is not always found. For instance, Durlak and colleagues (2011) found in their meta-analysis that interventions that involved the whole school and/or parents were not more effective than interventions that were only implemented in the classroom. These inconsistent findings indicate that more research is needed to clarify the influence of the width of an intervention's ecological focus on intervention effects. This knowledge is pivotal for schools in order to determine the extent of investment in intervention (e.g., number of teachers to train, whether or not to involve parents) that is necessary for intervention to work.

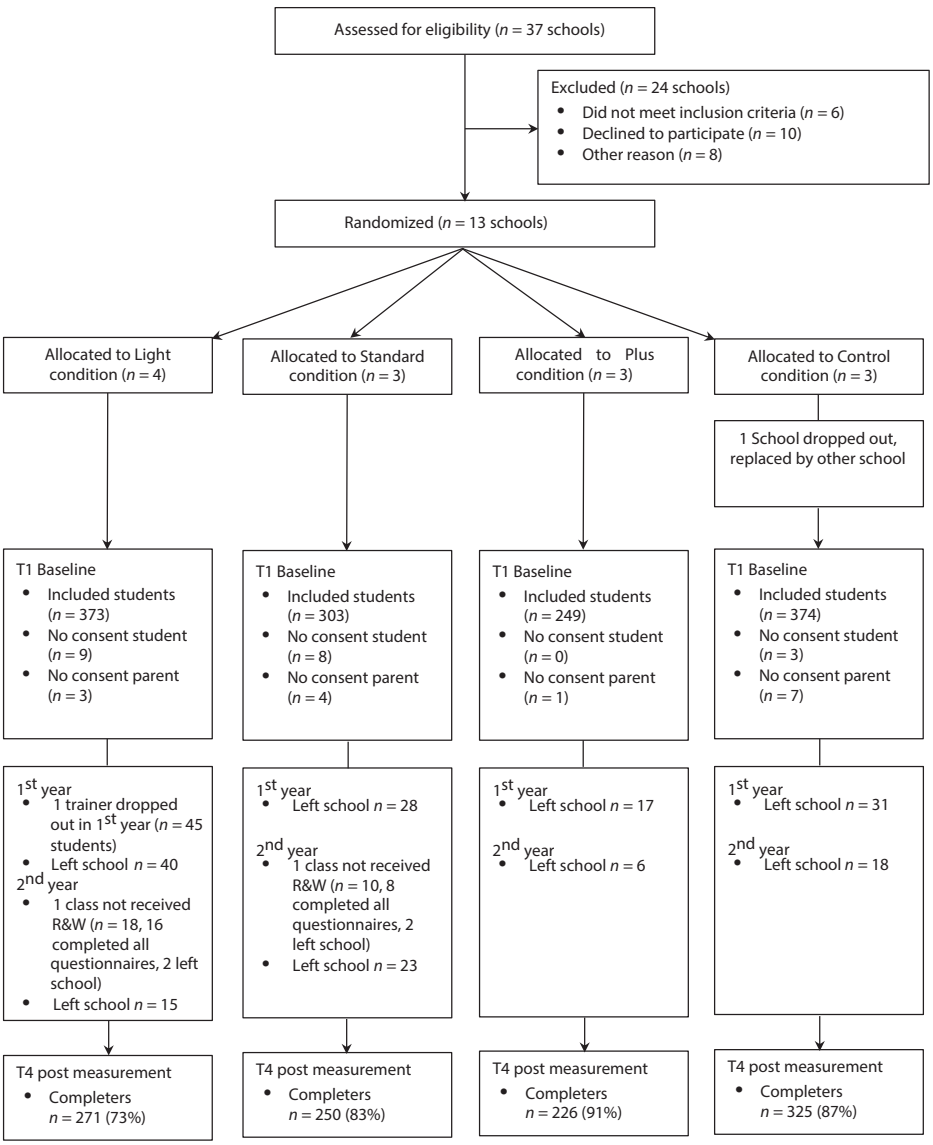
In sum, the present study had two aims. First, we examined whether a universal school-based intervention, R&W, could stimulate prevocational students' development in the intrapersonal and interpersonal domains. We hypothesized that R&W would be more effective than care as usual due to its unique approach that matches the learning processes of this group of students. Second, we examined the influence of the width of the ecological focus on intervention effects. We had no specific hypothesis concerning this aim, as research on the influence of ecological focus on the effectiveness of interventions is inconsistent.

Method

Design and Procedure

The effectiveness of R&W was examined in a Randomized Controlled Trial with four conditions that differ in width of the ecological focus. In the 'Light' condition only a core team of teachers was involved in the intervention. In the 'Standard' condition the entire teaching staff was involved. In the 'Plus' condition the entire teaching staff as well as parents were involved. In the Control condition, schools conducted their current school policy to enhance students' competencies and prevent problems in the intra- and interpersonal domain (i.e., care as usual).

Figure 1. Flow chart



Schools with a preparatory vocational education track (i.e., one of three educational tracks in the Dutch secondary school system) could participate in the present study. Schools were excluded from this study if they had implemented R&W in the last two years or if they were special education schools. Thirteen schools throughout the Netherlands, in urban and rural areas, were randomly allocated to the conditions (1:1:1:1) by stratified block randomization, with blocks of four (i.e., the number of

conditions) using an online random number generator. Schools were stratified by school size (small to moderate sized schools with < 100 students in the 7th Grade, large schools with > 100 students in the 7th Grade) to enhance an equal distribution of students over the conditions. One school, allocated to the Control condition, dropped out after randomization and before the start of data collection due to a change in school management. This school was replaced by another school (see Figure 1 for the flow chart).

Students completed digital questionnaires before the intervention started, at baseline (T1; October/November 2017), after completing the first year of R&W lessons (T2; March/April 2018), before the start of the R&W lessons in the second year of R&W (T3; October 2018), and after the intervention, at post measurement (T4; January 2019). These questionnaires were conducted by trained research assistants. Students gave active informed consent for completing the questionnaires. Parents gave passive informed consent for the participation of their child and active informed consent for their own participation. This trial was approved by the Ethical Committee of the Faculty of Social and Behavioral Sciences of Utrecht University (FETC17-05) and registered in the Dutch Trial Register, number NL6371 (NTR6554; see for protocol Mertens, Deković, Van Londen, & Reitz, 2018).

Participants

The sample consisted at baseline of 1299 7th Grade students. In the Netherlands, 7th Grade corresponds generally with ages 12 to 13 years. In our sample, the students had an average age of 12.38 years (*SD* = .62). Of the students, 661 (54%) were boys and 815 (69%) had a Western background (see Table 1 for the demographics per condition).

Table 1 Descriptives of Students' Demographics at Baseline per Condition

	Light	**Standard**	**Plus**	**Control**	**Differences at T1**		
					F/χ^2	p	$\eta^2_{partial}/\varphi$
N	373	303	249	374			
Age, *M (SD)*	12.33 (.57)	12.38 (.66)	12.34 (.60)	12.47 (.64)	3.89	.009	.009
Boys, *n* (%)	170 (48%)	161 (56%)	131 (55%)	199 (57%)	7.38	.061	.077
Western background, *n* (%)	291 (82%)	115 (43%)	211 (91%)	198 (59%)	182.01	<.001	.392

There were no differences between the conditions regarding sex distribution, but there were small differences in students' age and ethnic background (see Table 1). Students in the Control condition were slightly older than students in the Light condition. Regarding ethnic background, the Control and Standard conditions had roughly an equal distribution of students with a Western and non-Western background,

whereas the Light and Plus conditions consisted mostly of students with a Western background. Therefore we controlled for age and ethnic background in all analyses.

Attrition. Little's MCAR test indicated that the data were not missing completely at random (χ^2 (2339) = 2539.54, p = .002). Attrition analyses were conducted for the demographic and outcome variables between students who dropped out (n_{T2} = 68, n_{T3} = 83, n_{T4} = 60) and who remained in the study. Concerning the demographic variables, there were no differences on age (F(1,1230) = .24, p = .626, $\eta^2_{partial}$ = .000), sex distribution (χ^2 (1) = .1.36, p = .244, φ = -.033), and ethnic background (χ^2 (1) = 1.19, p = .276, φ = .032). Concerning the outcome variables, three MANOVAs, one per time point, showed that there were differences on the outcome variables at T1 (F(39, 3591) = 1.45, p = .035, $\eta^2_{partial}$ = .016), but not at T2 (F(39, 3363) = .72, p = .898, $\eta^2_{partial}$ = .008) and T3 (F(26, 2032) = .74, p = .824, $\eta^2_{partial}$ = .009). When further examining the differences between students who dropped out and who remained in the study after T1, the univariate test indicated that drop-outs differed from completers on externalizing behavior (F(3, 1207) = 3.00, p = .030). However, this difference was not significant anymore after Bonferroni correction (p = .124). As Little's MCAR test yield conservative results when applied to a large set of variables and because we found no differences between drop-outs and completers on the demographic or any of the outcome variables, we regarded the missing data as missing at random (Van Ness, Murphy, Araujo, Pisani, & Allore, 2007).

Conditions

R&W. R&W uses a psychophysical approach, integrating play and exercises to learn students how to make (physical) contact with others, and explore, respect, and set own and other's boundaries. The exercises are developed to enhance competencies as well as to prevent problems in students' intra- and interpersonal domains. For instance, according to the theory of R&W, students practice walking with an upright posture and experience that this has an influence on how they feel; if they walk with their head high, they feel more confident. In other exercises, students focus on their muscle tension and breathing which is theorized to raise their emotional awareness. They practice relaxing their muscles and lowering their breathing to regulate their emotions and become calm, less stressed and less aggressive. Students also engage in role-plays to practice, for instance, how to set and indicate boundaries, how to react calmly to provocations, and how to help students who get bullied. During the intervention, the symbolic principles of 'rock' and 'water' are used to indicate opposite ends of a spectrum: An uncompromising attitude in which the student is able to resist pressure from others (i.e., rock) to a flexible attitude in which the student is open to the opinions, thoughts, and feelings of others (i.e., water).

The intervention is a 2-year manualized program. In the first year, students received 14 R&W lessons and in the second year 8 R&W lessons. The lessons were implemented

weekly during 90-minute physical education lessons. Trainers were teachers of the schools, mostly physical education teachers as they have experience with teaching physical activities in class. During the lessons, students participate in physical exercises and games, reflect on the exercises, share and discuss their thoughts with each other, and address how to use the learned skills in their daily lives (see the study protocol for more information Mertens et al., 2018).

Condition check. In the 'Light' condition a core team of teachers followed the 3-day training to become certified R&W trainers and implemented the intervention. In the 'Standard' condition a core team of teachers was trained to implement R&W (i.e., R&W trainers) and the rest of the teaching staff received a 3-day training to learn how to support the R&W trainers and how to apply R&W in their regular classes. The 'Plus' condition was equal to the 'Standard' condition with the addition of a parent component; parents were invited to watch a documentary about R&W, join a R&W lesson in the school, received weekly e-mails with information about the R&W lesson of that week and were encouraged to act on this information. In all three conditions, the R&W trainers received supervision from their R&W coach during the implementation of the intervention.

To examine whether the condition manipulation was successful, trainers and parents completed questionnaires after the first and second year of the intervention. As planned, more teachers in the Standard and Plus conditions than in the Light condition were involved in the intervention; in the Light condition only the R&W trainers (i.e., trained teachers) implemented R&W, whereas in the Standard and Plus conditions it was reported that the R&W trainers as well as other teachers applied the intervention techniques. Furthermore, in the Plus condition parents were involved in the intervention; parents reported that they read the weekly information sometimes (57%) or often (39%), and some parents indicated they watched the documentary of R&W (15%) and participated in a R&W lesson at their child's school (23%). In addition, parents in the Plus condition reported to have implemented parts of the intervention more often at home than parents in the Light and Standard conditions. In conclusion, the involvement of teachers and parents was in accordance with the study design of the conditions.

Intervention fidelity. Intervention fidelity was assessed with two complementary methods: Self-reports of R&W trainers and observations of 67 R&W lessons by 3 R&W experts. R&W trainers completed a questionnaire about fidelity after each third lesson in the first year of the intervention and after each second lesson in the second year. R&W experts completed a coding schema based on Bishop and colleagues (2014) during the observation.

According to the self-reports, R&W trainers were generally able to mostly complete lessons (65%) and did not deviate or only slightly deviated from the manual (72%). R&W experts indicated that most observed lessons were completed or almost completed (86%). Trainers did not deviate much from the manual (91%). When trainers did adjust the intervention, these adjustments were generally judged as improvements (62%).

Overall, the quality of the observed R&W lessons was good (38%) to very good (54%) according to the experts. In conclusion, based on the self-reports and observations, the majority of the R&W lessons was indeed implemented and fidelity to the manual was moderate to high.

Control. In the Control condition students received care as usual, which varied between schools. For example, in one school, students had a teacher as personal coach with whom they had regular meetings, discussed their wellbeing, and could go to for advice. In another school, students could go to their mentor, a teacher, when they experienced problems, participated in a project week about 'being different', signed an anti-bullying contract, and discussed bullying in the class. In yet another school, students also had a mentor, a teacher, whom they could go to when they experienced difficulties and there was an 'anti-bullying coordinator' at the school. This coordinator facilitated actions to prevent or stop bullying which could differ per situation.

Primary Outcomes

Intrapersonal domain

Psychological wellbeing. To measure the presence of positive emotions, students completed the subscale Psychological wellbeing of the KIDSCREEN-27 (Ravens-Sieberer & The European KIDSCREEN Group, 2006). The subscale consists of 7 items (e.g., "Past week, did you have fun?") rated on a 5-point Likert-type scale (1 = *never* to 5 = *always*). Some items were recoded so that high scores indicated higher levels of psychological wellbeing (Cronbach's α = .76 - .83).

Resilience. To assess students' ability to bounce back from challenges that can arise in life, students completed the Connor-Davidson Resilience Scale – short version (Davidson & Connor, 2017) consisting of 10 items (e.g., "Able to adapt to change.") rated on a 5-point Likert-type scale (0 = *not true at all* to 4 = *true nearly all the time*; Cronbach's α = .79 - .92).

Sexual autonomy. To measure students' coping skills in sexual situations, students completed 5 items from the study Sex under 25 (e.g., "When I am with someone I like, I feel at ease."; De Graaf, Meijer, Poelman, & Vanwesenbeeck, 2005) rated on a 4-point Likert-type scale (1 = *never* to 4 = *always*). Some items were recoded so that high scores indicated higher levels of sexual autonomy. Reliability was poor at T1 (Cronbach's α = .53) and adequate at T2, T3, and T4 (Cronbach's α = .62 - .65).

Internalizing behavior. The presence of internalizing problems was measured with the internalizing subscale of the short version of the Youth Self Report (YSR; Chorpita et al., 2010). The subscale consists of 6 items (e.g., "I worry a lot.") rated on a 3-point Likert-type scale (0 = *never* to 2 = *often*; Cronbach's α = .79 - .86).

Interpersonal domain

Interpersonal relations in the class. Perceived interpersonal relations in the class were assessed using the Classroom Peer Context Questionnaire (Boor-Klip, Segers, Hendrickx, & Cillessen, 2016). The questionnaire measures negative social exchanges between classmates, the extent to which students feel comfortable around their classmates, and the unity and inclusiveness among classmates. The questionnaire consists of 12 items (e.g., "In this class students like each other.") answered on a 5-point Likert-type scale (1 = *totally not true* to 5 = *completely true*). Some items were recoded so that high scores indicated more positive interpersonal relations in the class (Cronbach's α = .80 - .85).

Externalizing behavior. To measure the presence of externalizing problems, students completed the externalizing subscale of the short version of the YSR (Chorpita et al., 2010). The subscale consists of 6 items (e.g., "I argue a lot.") rated on a 3-point Likert-type scale (0 = *never* to 2 = *often*; Cronbach's α = .65 - .79).

Aggression. Students' reactive and proactive aggression was measured with the Reactive and Proactive Aggression Questionnaire (Dodge & Coie, 1987). The questionnaire consists of 6 items (e.g., "If they tease me, I get angry.") answered on a 5-point Likert-type scale (1 = *never* to 5 = *almost always*; Cronbach's α = .65 - .83).

Bullying and victimization. To assess the frequency of bullying and experienced victimization, students completed the 2 global items of the Olweus Bully/Victim Questionnaire (Solberg, & Olweus, 2003). The items were: "How often have you taken part in bullying others?" and "How often have you been bullied?" and were preceded by a definition of bullying. Students responded on a 5-point Likert-type scale (1 = *never* to 5 = *almost always*).

Secondary Outcomes

Self-control. To assess students' ability to control their impulses and interrupt undesired behaviors, students completed the short version of the Self-Control Scale (Finkenauer, Engels, & Baumeister, 2005) consisting of 11 items (e.g., "I am good at resisting temptation.") rated on a 5-point Likert type scale (1 = *not at all* to 5 = *very much*). Some items were recoded so that high scores indicated high levels of self-control (Cronbach's α = .62 - .72).

Self-reflection. Students completed the subscale Engage in reflection of the Self-Reflection and Insight Scale (Sauter, Heyne, Blöte, Van Widenfelt, & Westenberg, 2010) to measure the extent to which students inspect and evaluate personal thoughts, feelings, and behaviors. The subscale consists of 6 items (e.g., "I often think about how I feel about something."), preceded by a definition of self-reflection, answered on a 6-point Likert scale (1 = *disagree strongly* to 6 = *agree strongly*). Some items were recoded so that higher scores represented higher levels of self-reflection. Reliability

was poor (Cronbach's α = .53 - .60). The 3 items that included a negative ("I don't think a lot about my thoughts.", "I almost never participate in 'self-reflection'.", and "I don't think about the reason why I behave the way I do.") were deleted to avoid a double negative and to improve reliability (Cronbach's α = .74 - .89).

Self-esteem. Students' level of global self-worth was measured using the subscale Global self-perception of the Self-Perception Profile (Harter, 1988) completed by the students. The subscale has 5 items (e.g., "I am satisfied with myself.") answered on a 4-point Likert-type scale (1 = *completely not true* to 4 = *completely true*). Some items were recoded so that higher scores indicated high levels of self-esteem (Cronbach's α = .73 - .75).

Emotional self-regulation. Students completed the Difficulties in Emotion Regulation Scale (Anderson, Reilly, Gorrell, Schaumberg, & Anderson, 2016) to assess students' abilities to control their emotions and their access to emotion regulation strategies (e.g., "When I'm upset, I know that I can find a way to eventually feel better."). The questionnaire consists of 14 items answered on a 5-point Likert-type scale (1 = *almost never* to 5 = *almost always*). Some items were recoded so that high scores indicate higher levels of emotional self-regulation (Cronbach's α = .88 - .91).

Statistical Analyses

Data were analyzed using an intention-to-treat approach in which students assigned to the intervention were included in the analyses regardless of whether they actually participated in the intervention or not. Participants were nested in schools in classes. We took clustering at school level into account by applying the complex sample cluster feature of M*plus* (version 8.2; Muthén & Muthén, 2010). This is a conservative clustering procedure providing unbiased estimates of the standard errors (Muthén & Muthén, 2010). Clustering at class level was not taken into account as class composition was not stable over the years (e.g., Cross et al., 2016). To include all participants in the model, we used Full Information Maximum Likelihood (FIML) procedures. Parameter estimates were obtained through Robust Maximum Likelihood estimation (MLR) which is robust to non-normality and non-independence of observations (Muthén & Muthén, 2010).

To examine the effectiveness of R&W, we tested a series of latent growth curve (LGC) models in M*plus*, as suggested by Greenberg and Abenavoli (2017). LGC models estimate for each participant an individual growth curve based on his/her initial level (i.e., intercept) and change over time (i.e., slope). The individual growth curves are indicators of latent variables describing average group growth trajectories allowing for differences in trajectories between participants (Muthén & Muthén, 2010). The slope is of main importance; when the intervention is effective compared to the Control condition, it significantly alters the slope in the desired direction. To allow for nonlinear growth, we did not specify the rate of growth at T2 and T3 (Duncan & Duncan, 2004). Growth rates at T1 and T4 were specified at respectively 0 and 3.

To assess the effects of the intervention, we constructed three dummy variables (i.e., Light, Standard, and Plus condition) with the Control condition as a reference group and regressed the intercept and slope on these three dummy variables. Students' age and ethnicity were added as covariates as the conditions differed significantly on these variables. If two or more intervention conditions appeared to be effective compared to the control group, we examined the effectiveness of those conditions compared to each other in a multigroup model by constraining the slopes of those conditions to be equal and by releasing this constraint. The model fits of the two nested models were compared using the Satorra-Bentler Scale Chi-Square test. This test applies a scaling correction to better approximate the chi-square distribution under non-normality (Satorra & Bentler, 2010). A significant Satorra-Bentler Scale Chi-Square test indicates that the unconstrained model fits better and, thus, that one intervention condition is more effective than the other.

We calculated effect sizes by multiplying the rate of change by time span divided by the standard deviation of the concerned outcome (d = (slope * duration) / SD; Feingold, 2013). We calculated effect sizes for the change between measurement points (i.e., change from T1 to T2, from T2 to T3, and from T3 to T4) and the overall change (i.e., change from T1 to T4). As there is no specific formula to calculate effect sizes for unspecified non-linear growth, the overall effect sizes were calculated using the formula for linear growth[1].

Results

Descriptive Statistics

Table 2 presents the means and standard deviations of the outcomes for the conditions on each of the measurement points. The LGC models showed acceptable fit (see Table 3; RMSEA < .08, CFI > .90, SRMR < .10; Kline, 2005). The models for bullying and victimization showed a poor fit based on the CFI, but a good fit based on the RMSEA and SRMR. The standardized factor loadings of the time points on the slope, reflecting the average change in the observed variables, indicated that students showed generally the largest change on the outcomes from T1 to T2 (see Table 3). After T2 the average change leveled off. For instance, students changed on sexual autonomy from T1 to T2 with a rate of .69, from T2 to T3 with a rate of .15, and from T3 to T4 with a rate of .05.

1 To examine the robustness of the overall effect sizes of the unspecified growth models, we modeled linear LGC models and calculated the overall effect sizes. The overall effect sizes of the linear models were in general larger than the effect sizes based on the unspecified growth models indicating that the effect sizes of the unspecified growth models were more conservative.

Table 2 Means and Standard Deviations of Outcomes per Condition per Measurement Point

	R&W Light				R&W Standard				R&W Plus				Control			
	T1	T2	T3	T4	T1	T2	T3	T4	T1	T2	T3	T4	T1	T2	T3	T4
	M (SD)	*M (SD)*	*M (SD)*	*M (SD)*	*M (SD)*	*M (SD)*	*M (SD)*	*M (SD)*	*M (SD)*	*M (SD)*	*M (SD)*	*M (SD)*	*M (SD)*	*M (SD)*	*M (SD)*	*M (SD)*
Intrapersonal domain																
Psychological wellbeing	3.94 (.62)	3.98 (.69)	3.91 (.71)	3.87 (.78)	4.02 (.61)	4.10 (.71)	4.02 (.73)	3.93 (.77)	3.97 (.62)	4.06 (.63)	3.87 (.72)	3.85 (.80)	4.06 (.59)	4.02 (.68)	3.90 (.75)	3.92 (.78)
Resilience	2.36 (.65)	2.49 (.75)	2.15 (.87)	2.47 (.82)	2.43 (.67)	2.70 (.71)	2.56 (.70)	2.53 (.90)	2.43 (.64)	2.57 (.68)	2.47 (.71)	2.51 (.84)	2.51 (.65)	2.68 (.69)	2.57 (.74)	2.61 (.85)
Sexual autonomy	3.23 (.44)	3.39 (.48)	3.40 (.44)	3.40 (.46)	3.30 (.44)	3.44 (.52)	3.38 (.54)	3.39 (.51)	3.29 (.40)	3.38 (.46)	3.34 (.45)	3.35 (.57)	3.32 (.44)	3.36 (.49)	3.32 (.50)	3.36 (.52)
Internalizing behavior	.53 (.42)	.41 (.44)	.44 (.44)	.42 (.46)	.44 (.42)	.33 (.39)	.40 (.44)	.37 (.43)	.49 (.40)	.45 (.42)	.51 (.44)	.48 (.51)	.43 (.39)	.39 (.41)	.45 (.46)	.42 (.44)
Interpersonal domain																
Interpersonal relations in the class	3.90 (.64)	3.81 (.68)	3.83 (.67)	3.73 (.69)	3.98 (.63)	3.87 (.72)	3.81 (.73)	3.80 (.75)	4.07 (.62)	4.00 (.63)	3.92 (.69)	3.84 (.81)	3.91 (.59)	3.78 (.65)	3.77 (.67)	3.75 (.72)
Externalizing behavior	.59 (.38)	.47 (.39)	.48 (.39)	.45 (.41)	.54 (.39)	.43 (.38)	.50 (.44)	.42 (.42)	.55 (.35)	.46 (.36)	.55 (.38)	.49 (.45)	.49 (.34)	.42 (.37)	.45 (.39)	.43 (.39)
Aggression	1.79 (.56)	1.71 (.59)	1.69 (.59)	1.66 (.66)	1.85 (.61)	1.84 (.73)	1.94 (.73)	1.83 (.81)	1.73 (.46)	1.72 (.60)	1.80 (.58)	1.74 (.81)	1.78 (.56)	1.76 (.65)	1.81 (.67)	1.85 (.78)
Bullying	1.13 (.48)	1.18 (.61)	1.15 (.55)	1.21 (.64)	1.13 (.41)	1.18 (.60)	1.15 (.58)	1.19 (.64)	1.10 (.41)	1.15 (.47)	1.20 (.68)	1.20 (.64)	1.08 (.36)	1.18 (.59)	1.22 (.71)	1.22 (.71)
Victimization	1.32 (.77)	1.43 (.96)	1.30 (.77)	1.31 (.86)	1.28 (.73)	1.35 (.87)	1.22 (.77)	1.22 (.70)	1.36 (.88)	1.37 (.85)	1.40 (.96)	1.40 (.93)	1.37 (.94)	1.56 (1.16)	1.41 (1.04)	1.34 (.95)
Secondary outcomes																
Self-control	3.33 (.58)	3.53 (.60)	3.52 (.61)	3.54 (.60)	3.49 (.57)	3.66 (.66)	3.56 (.65)	3.61 (.69)	3.41 (.60)	3.53 (.59)	3.43 (.61)	3.53 (.65)	3.49 (.56)	3.59 (.62)	3.57 (.65)	3.58 (.64)
Self-reflection	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)
Self-esteem	3.23 (.62)	3.28 (.66)	3.24 (.65)	3.28 (.62)	3.28 (.57)	3.33 (.60)	3.29 (.66)	3.29 (.62)	3.25 (.59)	3.33 (.57)	3.18 (.68)	3.12 (.70)	3.31 (.59)	3.29 (.64)	3.22 (.65)	3.22 (.65)
Emotional self-regulation	3.83 (.77)	3.94 (.85)	3.98 (.76)	3.97 (.75)	3.85 (.69)	3.99 (.80)	3.91 (.82)	3.91 (.84)	3.87 (.71)	3.88 (.74)	3.81 (.80)	3.83 (.85)	3.93 (.71)	3.94 (.75)	3.90 (.76)	3.92 (.76)

Table 3 Model Fit Indices and Growth Over Time of Outcomes

	Model fit statistics			Factor loadings on slope			
	RMSEA	CFI	SRMR	T1	T2	T3	T4
Intrapersonal domain							
Psychological wellbeing	.062	.907	.067	0	.19	.61	.54
Resilience	.060	.796	.084	0	.19	.04	.11
Sexual autonomy[1]	.053	.862	.073	0	.69	.84	.79
Internalizing behavior[1]	.061	.934	.068	0	.62	.73	.79
Interpersonal domain							
Interpersonal relations in the class	.053	.918	.069	0	.36	.59	.69
Externalizing behavior	.059	.919	.070	0	.28	.17	.22
Aggression	.055	.888	.069	0	.38	.59	.55
Bullying	.050	.635	.068	0	.28	.47	.60
Victimization	.058	.696	.069	0	.36	.63	.60
Secondary outcomes							
Self-control[1]	.057	.936	.077	0	.66	.76	.85
Self-reflection	.060	.912	.067	0	.54	.64	.59
Self-esteem	.068	.925	.073	0	.09	.43	.49
Emotional self-regulation	.057	.931	.075	0	.41	.64	.69

Note. [1]Variance of the baseline measurement of the concerned outcome variable was fixed to zero due to a negative residual variance of the observed variable at T1.

Effects of R&W

The standardized regression coefficients of the slope on the intervention conditions (compared to the Control condition) are reported per outcome in Table 4. The effect sizes between measurement points and overall effect sizes are reported per outcome for the intervention condition (compared to the Control condition) in Table 5.

Intrapersonal domain. Students in the Light condition showed a more beneficial trajectory of change for psychological wellbeing, sexual autonomy, and internalizing behavior compared to students in the Control condition (see Figure 2). Effect sizes were small (Cohen's d = .26 - .38; Cohen, 1988). In the Light condition students remained stable on psychological wellbeing (slope = .08), whereas students in the Control condition decreased (slope = -.25). The intervention effect is strongest from T2 to T3, i.e., in between the first and second year of the intervention. On sexual autonomy, students in the Light condition showed a steeper increase (slope = .73) than students in the Control condition (slope = .32). Students improved most in the first year of the intervention. For internalizing behavior, students in the Light and Standard conditions showed a steeper decline ($slope_{Light}$ = -.60; $slope_{Standard}$ = -.53) than students in the

Control condition (slope = -.41). In both the Light and Standard condition the strongest improvements were again in the first year. No intervention effects were found for resilience and for students in the Plus condition (see the supplementary material for the slopes of all conditions and outcomes in the intrapersonal domain).

When comparing students in the Light condition with students in the Standard condition on change over time in internalizing behavior, the unconstrained model fitted the data significantly better, $\Delta\chi^2_{SB}$ (1) = 6.60, p = .010, than the constrained model (i.e., slopes constrained to be equal across the conditions). Students in the Light condition showed a stronger decrease in internalizing behavior than students in the Standard condition ($slope_{Light}$ = -.60; $slope_{Standard}$ = -.53). Overall, students in the Light condition seemed to have a slightly more beneficial change over time in the intrapersonal domain with the intervention's primary impact early in the intervention.

Interpersonal domain. Students in the Light condition showed a more beneficial trajectory of change for aggression compared to students in the Control condition (see Figure 2). The effect size was small (Cohen, 1988). Students in the Light condition remained relatively stable over time (slope = .09), whereas students in the Control condition showed an increase in aggression (slope = .56). The strongest intervention effect was found in the first year of the intervention. Furthermore, two trends suggested that students in the Light condition also had a slightly more beneficial trajectory of change for interpersonal relations in the class and bullying compared to students in the Control condition (see Figure 2). Students in the Light condition showed a less steep decrease in positive interpersonal relations in the class (slope = -1.64) and a less steep increase in bullying (slope = 2.34) than students in the Control condition ($slope_{Interpersonal\ relations}$ = -2.03; $slope_{Bullying}$ = 2.80). Both effect sizes were small (Cohen, 1988) with the strongest effects found in the first year. No intervention effects were found for externalizing behavior and victimization. Additionally, no intervention effects were found for students in the Standard and Plus conditions (see the supplementary material for the slopes of all conditions and outcomes in the interpersonal domain). Overall, the Light condition also seemed to have a slightly more beneficial change over time in the interpersonal domain with, again, the strongest intervention effect in the first year of the intervention.

Table 4 Intervention Effects Over Time of R&W Conditions Compared to Control Condition

	Light			Standard			Plus		
	β	*SE*	*p*	β	*SE*	*p*	β	*SE*	*p*
Intrapersonal domain									
Psychological wellbeing	.15	.06	.009	.08	.08	.287	.04	.04	.228
Resilience	.09	.73	.902	.18	.27	.507	-.10	.31	.736
Sexual autonomy[1]	.18	.04	< .001	.08	.06	.144	.06	.03	.074
Internalizing behavior[1]	-.13	.04	< .001	-.07	.04	.048	.01	.03	.812
Interpersonal domain									
Interpersonal relations in the class	.05	.03	.056	-.05	.07	.499	-.04	.04	.278
Externalizing behavior	-.20	.12	.107	-.13	.11	.223	-.02	.07	.756
Aggression	-.19	.09	.030	.01	.10	.933	-.01	.05	.822
Bullying	-.12	.06	.056	-.12	.09	.149	-.05	.09	.624
Victimization	-.01	.04	.817	-.04	.06	.527	.03	.04	.530
Secondary outcomes									
Self-control[1]	.12	.03	< .001	.00	.04	.951	.02	.04	.532
Self-reflection	.01	.04	.803	-.02	.06	.761	.04	.03	.238
Self-esteem	.18	.10	.060	.09	.11	.394	-.05	.06	.327
Emotional self-regulation	.17	.05	< .001	.07	.05	.163	-.01	.02	.676

Note. [1]Variance of the baseline measurement of the concerned outcome variable was fixed to zero due to a negative residual variance of the observed variable at T1

Table 5 Effect Sizes of Change over Time in R&W Conditions Compared to Control Condition

	Light				**Standard**				**Plus**			
	T1-T2	T2-T3	T3-T4	Overall	T1-T2	T2-T3	T3-T4	Overall	T1-T2	T2-T3	T3-T4	Overall
Intrapersonal domain												
Psychological wellbeing	0.07	0.18	-0.01	0.24	0.04	0.10	-0.01	0.14	0.02	0.06	0.00	0.08
Resilience	0.04	-0.03	0.02	0.03	0.09	-0.07	0.04	0.06	-0.05	0.04	-0.02	-0.04
Sexual autonomy[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
Internalizing behavior[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
Interpersonal domain												
Interpersonal relations in the class	0.05	0.03	0.02	0.10	-0.05	-0.03	-0.02	-0.10	-0.04	-0.02	-0.02	-0.08
Externalizing behavior	0.13	-0.05	0.03	0.11	0.09	-0.03	0.02	0.08	0.02	-0.01	0.00	0.02
Aggression	0.18	0.10	0.03	0.31	-0.01	-0.01	0.00	-0.02	0.01	0.01	0.00	0.02
Bullying	0.10	0.08	0.06	0.24	0.11	0.09	0.07	0.28	0.05	0.04	0.03	0.11
Victimization	0.01	0.01	0.00	0.01	0.04	0.02	0.00	0.06	-0.03	-0.02	0.00	-0.04
Secondary outcomes												
Self-control[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
Self-reflection	0.01	0.00	0.00	0.02	-0.02	0.00	0.00	-0.03	0.06	0.01	0.00	0.07
Self-esteem	0.04	0.15	0.04	0.22	0.02	0.08	0.02	0.12	-0.01	-0.05	-0.01	-0.08
Emotional self-regulation	0.17	0.09	0.04	0.29	0.08	0.04	0.02	0.13	-0.01	0.00	0.00	-0.01

Note. [1]Variance of the baseline measurement of the concerned outcome variable was fixed to zero due to a negative residual variance of the observed variable; Effect sizes (*d*) were calculated so that positive effect sizes indicate change in the desired direction for R&W.

Figure 2. Estimated growth trajectories of the conditions concerning outcomes in the intra- and interpersonal domains.

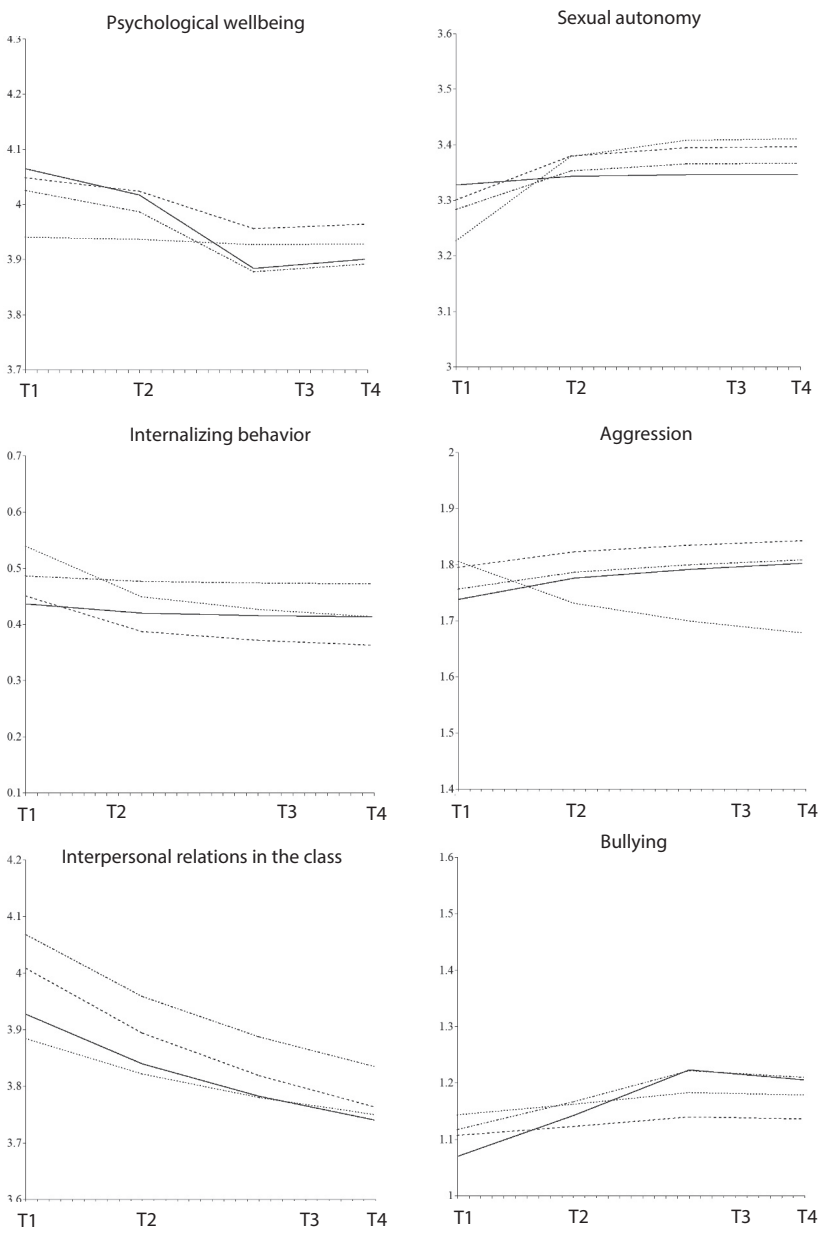
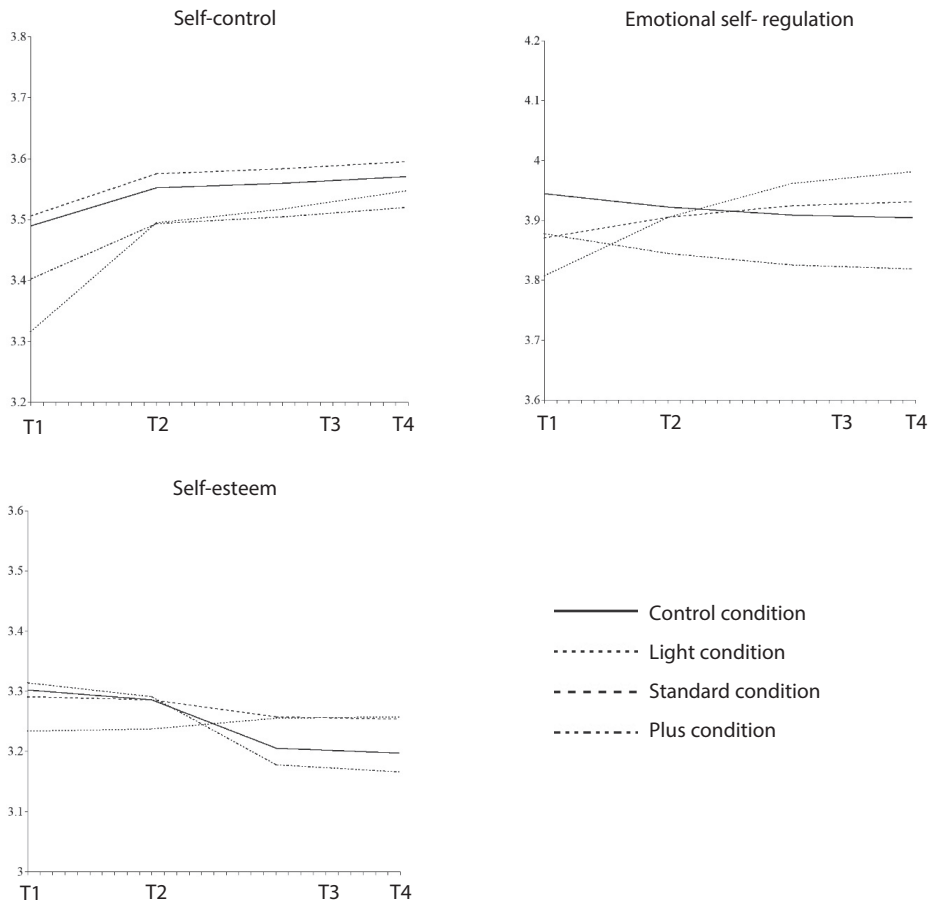


Figure 3. Estimated growth trajectories of the intervention and control conditions concerning secondary outcomes.



Secondary outcomes. Students in the Light conditions showed more beneficial trajectories of change for self-control and emotional self-regulation compared to students in the Control condition (see Figure 3). The effect sizes were small for both outcomes (Cohen, 1988). Students in the Light condition showed a steeper increase in self-control (slope = 1.75) than students in the Control condition (slope = 1.44). Regarding emotional self-regulation, students in the Light condition improved over time (slope = .24), whereas students in the Control condition slightly declined (slope = -.10). For both outcomes, students improved most in the first year of the intervention. Furthermore, there was a trend suggesting that students in the Light condition had a slightly more beneficial trajectory for self-esteem than in the Control condition. Students in the Light condition remained relatively stable (slope = -1.68), whereas students in the Control condition showed a small decrease in self-esteem (slope = -2.58). Intervention effects were small (Cohen, 1988) with the strongest effect from T2 to T3, so in between the first and second year of the intervention. No intervention effects were found for self-reflection and for students in the Standard and Plus conditions (see the supplementary material for the slopes of all conditions and outcomes in the interpersonal domain). Again, the Light condition seemed to have the most beneficial growth trajectories for the secondary outcomes with most improvement early in the intervention.

Discussion

The purpose of the present study was, first, to examine whether a psychophysical intervention could positively affect prevocational students, a challenging group to involve in interventions, and, second, to determine the extent to which the width of an intervention's ecological focus influences intervention effects. The psychophysical intervention R&W showed to be moderately effective in fostering some aspects of students' intra- and interpersonal domains, but only when the ecological focus of the intervention was narrow. R&W Light was specifically effective in stimulating competencies and preventing problems in the intrapersonal domain (i.e., psychological wellbeing, sexual autonomy, and internalizing behavior). In the interpersonal domain, R&W Light showed a potential tendency to function as a buffer against declining positive interpersonal relations in the class and against increasing aggression and bullying. In addition, R&W showed small positive intervention effects on the secondary outcomes (i.e., self-control, emotional self-regulation, and a trend for self-esteem). Although intervention effects were small (when significant effects were found they ranged from .24 to .38), the observed effect sizes are consistent with effect sizes reported for other universal school-based interventions addressing the intra- and/or interpersonal domain (e.g., Cohen's d = .22 – .27, Durlak et al., 2011; Cohen's d = .10 – 25, Mertens et al., 2020).

It appears that using a psychophysical approach is a method that fits well with interventions targeting prevocational students; the alternation between psychological instruction or reflection and physical exercises or games possibly increase students' engagement in the intervention (Ter Vrugte et al., 2015), enabling them to optimally benefit from the intervention. Most intervention effects were found in students' intrapersonal domain. Only one intervention effect (aggression, and two trends for interpersonal relations in the class and bullying) was found in the interpersonal domain. A possible explanation for the larger impact of R&W in the intrapersonal domain could be that a more physical approach is used when addressing that domain, while a more verbal approach is used when focusing on the interpersonal domain. Based on the description in the intervention's manual (Ykema 2002; 2018), the main emphasis in the exercises and games is on students' own feelings, emotions, and attitudes (i.e., the intrapersonal domain). "Was your breathing low? Were you balanced and calm during the game?" Students' attitudes and behaviors in relation to others (i.e., the interpersonal domain) is mainly addressed during (verbal) role-play and discussions when sharing thoughts together. They discuss, for instance, what bullying is and what students can do about it. "How did the bully feel when they said 'stop' together?" This more verbal approach is cognitively more demanding and could possibly hinder prevocational students to benefit from the intervention in the interpersonal domain.

Regarding the secondary outcomes, positive intervention effects were found for three of the four underlying competencies that are important according to the theory of R&W (i.e., self-control, emotional self-regulation, and a trend regarding self-esteem) indicating the intervention's potential. However, no intervention effect was found in students' self-reflection which is one of the pillars of the R&W house. Also previous interventions appeared to be ineffective in improving self-reflection of prevocational students. For instance, Ter Vrugte and colleagues (2015) examined the effectiveness of a game to improve math skills in prevocational students. This game included a reflection stimulating component, but they found no effect on students' reflection. Perhaps, self-reflection is cognitively too demanding and abstract for prevocational students since it requires thinking on the metacognition level (Sauter et al., 2010). Nonetheless, as self-reflection has been indicated as a potential effective component in universal school-based interventions (Mertens et al., 2020), future research should examine whether self-reflection is an effective intervention component for prevocational students and, if so, how self-reflection can be stimulated in these students.

The intervention effects were strongest during the first part of the intervention and leveled off to insignificant effects in the second year, indicating that it might be sufficient to only implement the first year of the intervention. A decline in intervention effects in relatively long interventions has been found previously in meta-analyses examining different types of interventions (Bakermans-Kranenburg, Van IJzendoorn, & Juffer, 2003; Cuijpers, 2000; De Mooij, Fekkes, Scholte, & Overbeek, 2019). These meta-analyses suggest that short-term interventions with a modest number of sessions are

preferred. Research has shown that participants who benefit from an intervention often show improvement early in the intervention regardless of its time span (e.g., Lutz et al., 2014; Tadić et al., 2010). Hence, the finding that the strongest effects were shown in the first part of the intervention might represent a typical trajectory of intervention effects in general. The decline in intervention effects is possibly related to a decrease in students' motivation, as motivation is found to be a moderator of intervention effects (Philips & Wennberg, 2014). It could be that the second year of R&W contains too much repetition and not enough deepening of previously learned skills or new topics. Thus, implementing only the first year of the intervention appears to minimize the burden on the students without jeopardizing the effectiveness of the intervention.

The second aim was to determine the influence of the width of the intervention's ecological focus on its effectiveness. The results showed that the condition with the most narrow ecological focus (i.e., only a core team of teachers was involved in the intervention) appeared to be most effective, suggesting that for some interventions "less is more". This finding is in line with the results of the meta-analysis of Durlak and colleagues (2011) who showed that the positive effect of involving more people and systems in interventions is not always found. Additionally, Taylor, Mumford, Liu, and Stein (2017) examined a school-based intervention with different levels of involved people and found that the intervention effects did not increase as more students and teachers were involved.

An intervention with a narrow ecological focus might benefit from trainers' feelings of responsibility for implementation. In an intervention with a narrow ecological focus only a few people are involved and thus solely responsible for properly and effectively implementing the intervention. In an intervention with a broad ecological focus many people are involved and can share the responsibility for implementation. This diffused responsibility might trigger a 'bystander effect'. A bystander effect refers to one's diminished feelings of responsibility to act in a situation when more people are present (Fischer et al., 2011). Perhaps the R&W trainers and other teachers in the Standard and Plus condition had a more passive attitude in the implementation, waiting for the other to act, whereas R&W trainers in the Light condition had a more active attitude in the implementation as they were the only ones that could act regarding the intervention. Future research should focus on the possible influence of bystander effects when responsibility for implementation is shared among a relatively large number of people.

Furthermore, interventions with a broad ecological focus may have an increased risk of sending mixed messages to the students due to the large number of people that are involved. These mixed messages could decrease intervention effects. How the intervention lessons should be implemented is explicitly described in the manual (Ykema, 2002; 2018). How the intervention techniques can be applied during regular lessons or at home is not described nor structured. This lack of structure can be especially challenging with techniques based on a physical approach. Hence, outside the intervention lessons the techniques and skills of the intervention can be applied

in different ways by other teachers or parents. Receiving various and, possibly, mixed messages could confuse students and push the intended intervention message to the background reducing the chance of students to benefit from the intervention.

The results of the present study should be interpreted in light of its strengths and limitations. Strengths of the study were the longitudinal data, the three conditions differing in the width of ecological focus, and the large sample size. This enabled us to examine change in students over two years with different levels of ecological focus. A limitation is that R&W trainers did not report *how* teachers applied the intervention techniques during regular lessons, but only how often. Information about how intervention techniques were applied during regular lessons would have provided insight in possible mixed intervention messages. Hence, future research should also measure how intervention techniques are used outside the intervention lessons to examine whether differential implementation affects intervention effects. This knowledge can shed light if and to what extent implementation outside intervention lessons should be structured. Second, the R&W trainers were providing the intervention for the first time, after three days of training. More experienced trainers might be able to establish more change. On the other hand, using first time trainers ensured that trainers in our study were comparable concerning their experience with the intervention. Last, we examined intervention effects immediately following the intervention. Future research should analyze follow-up data to examine long term effectiveness of the intervention.

Conclusion

Prevocational students seem to be most in need of an effective intervention to positively stimulate their competencies and prevent the development of problems in both the intra- and interpersonal domain, but might be at the same time a challenging group of students to target. Our study showed that an intervention using a psychophysical approach can positively affect prevocational students. The intervention was especially effective in targeting students' intrapersonal domain and showed the strongest, albeit small, effects in the first year. In the interpersonal domain, the intervention potentially functions as a buffer. R&W showed a tendency to lessen the decline in positive interpersonal relations in the class and the increase in levels of aggression and bullying. Furthermore, the present study showed that intervention effects do not per definition increase with a longer intervention or a broader ecological focus; strongest intervention effects were shown in the first year and when the intervention had a narrow ecological focus. The finding that "less is more" has important implications for the practice as it indicates that it might *not* always be worthwhile for schools to invest in implementing a long-term intervention with a broad ecological focus.

Supplementary material

Table S1. Standardized Slopes of Trajectories of the Conditions

	Light			Standard			Plus			Control		
	slope	*SE*	*p*	slope	*SE*	*p*	slope	*SE*	*p*	slope	*SE*	*p*
Intrapersonal domain												
Psychological wellbeing	.08**	1.13	.947	-.12	1.43	.935	-.26	1.46	.860	-.25	1.07	.812
Resilience	.55	.75	.462	.61	.63	.328	.52	.69	.449	.63	.78	.416
Sexual autonomy[1]	.73**	.62	.240	.50	.61	.408	.50	.65	.446	.32	.63	.613
Internalizing behavior[1]	-.60**[a]	.57	.289	-.53*	.61	.382	-.32	.57	.571	-.41	77	.594
Interpersonal domain												
Interpersonal relations in the class	-1.64†	1.00	.102	-1.71	.96	.074	-1.93	1.05	.067	-2.03	1.22	.096
Externalizing behavior	-.05	.30	.860	-.04	.61	.947	.03	.24	.894	.03	.25	.908
Aggression	.09*	1.14	.940	.31	.55	.568	.54	.90	.548	.56	.90	.530
Bullying	2.34†	1.83	.203	1.65	1.36	.225	-[3]	-	-	2.80	1.67	.093
Victimization	.29	.74	.695	.37	1.17	.750	1.02	3.83	.790	.80	7.08	.910
Secondary outcomes												
Self-control[1]	1.75**	.73	.017	1.51	.78	.050	1.44	.74	.051	1.44	.76	.056
Self-reflection[2]	.46	.55	.407	.44	.66	.503	.54	.59	.362	.43	.56	.444
Self-esteem	-1.68†	1.27	.187	-2.13	1.38	.122	-1.98	1.04	.056	-2.58	1.48	.081
Emotional self-regulation	.24**	.82	.769	.10	1.15	.934	-.18	1.30	.893	-.10	1.01	.919

Note. [a] slopes differ significantly; †$p < .10$ *$p < .05$ **$p < .01$

Mertens, E. C. A.,
Deković, M.,
Van Londen, M.,
Reitz, E.

Manuscript submitted for publication.

Author Contributions

EM, MD, MvL, and ER conceptualized the study. EM coordinated the data collection, analyzed the data and wrote the manuscript. All authors provided feedback on the study.

Acknowledgments

This study is financially supported by a grant from the Netherlands Organization for Health Research and Development (ZonMw Grant No. 531001106).

4

Personality as a Moderator of Intervention Effects of a School-Based Intervention

Abstract

Individuals with higher or lower levels of certain personality traits might benefit more from an intervention than individuals with opposite levels of these traits, as some personality traits could make individuals more vulnerable to develop problems, providing more potential to improve, whereas other personality traits could facilitate transfer of the learned skills to daily life. The aim of the present study was to examine whether Big Five personality traits affected the effectiveness of a universal school-based intervention aiming to improve competencies and prevent problems in adolescents' intra- (e.g., psychological wellbeing) and interpersonal (e.g., aggression) domains. In a two-year randomized controlled trial, adolescents ($N = 1299$, $M_{age} = 12.38$) reported on the outcomes at four different time points. Parents reported at baseline on adolescents' personality. Although most intervention effects were not moderated by personality traits, three patterns of moderation emerged. First, there was a tendency that more vulnerable adolescents – based on their levels of certain personality traits such as high levels of Extraversion, high levels of Agreeableness, high levels of Neuroticism, or low levels of Conscientiousness – benefitted most from the intervention. Second, high levels of Extraversion appeared to enable adolescents to benefit more from a universal intervention when it requires sociability from participants. Third, personality traits seemed to affect the intrapersonal domain more than the interpersonal domain, both as predictors and as moderators of intervention effects. The present study increases general insights in how personality traits might affect intervention effects for certain types of adolescents, interventions, and outcomes.

Keywords: *Personality; RCT; School-based intervention; Moderation; Intrapersonal domain; Interpersonal domain*

Personality as a Moderator of Intervention Effects of a School-Based Intervention

Personality traits are one of the most influential predictors of psychosocial development, over and above demographic variables and life events (DeNeve & Cooper, 1998; Pellegrino & Hilton, 2012). The important role personality plays in one's psychosocial development suggests that the effectiveness of interventions aiming to stimulate one's competencies and prevent the development of problems might be dependent on personality traits. The empirical evidence regarding the role of personality in affecting the effectiveness of interventions is still scarce. By unraveling whether and how personality traits affect intervention effects on both a broad level (i.e., a general domain) and a narrow level (i.e., a specific competence or problem within a general domain), interventions can be tailored to (subgroups of) individuals (Tackett, 2006). In the present study, we examined personality traits as moderators of intervention effects in a universal intervention in which we focused on the intra- and interpersonal domains in general and on specific competencies and problems within these domains. The intrapersonal domain refers to feelings, emotions, and attitudes about the self (Barber, 2005), such as psychological wellbeing and internalizing behavior. The interpersonal domain refers to the ability to build and maintain positive relationships with others and to understand social situations, roles and norms, and respond appropriately (Pellegrino & Hilton, 2012; Shek & Leung, 2016), such as interpersonal relations and aggression.

Personality traits influence the extent to which individuals are likely to develop problems, which in turn may affect intervention effects. The vulnerability theory (Tackett, 2006) states that certain personality traits can increase one's risk of developing and maintaining problems in the intra- and interpersonal domains. Particularly in the intrapersonal domain, personality traits seem important predictors of one's competencies and problems (Van Leeuwen, Mervielde, Braet, & Bosmans, 2004). Individuals who have low levels of Extraversion, Agreeableness, or Conscientiousness, or high levels of Neuroticism, or a combination of these, are suggested to be more vulnerable to develop problems (Tackett, 2006). These individuals tend to be oriented towards their inner world of subjective experiences (low levels of Extraversion), ruthless and manipulative (low levels of Agreeableness), show difficulties in delaying gratification and modulating impulsivity (low levels of Conscientiousness), and/or perceive the world as distressing or threatening (high levels of Neuroticism; Pellegrino & Hilton, 2012; Shiner & Caspi, 2003). Furthermore, these levels of personality traits

have been related to less use of effective coping strategies, such as problem solving and cognitive restructuring, and more use of problematic coping strategies, such as disengagement, a focus on negative emotions, and coping through substance use (Connor-Smith & Flachsbart, 2007). According to the Risk moderation hypothesis (Spoth, Shin, Guyll, Redmond, & Azevedo, 2006), especially individuals who experiences difficulties in certain competencies or experience problems might benefit most from interventions as they have most potential to improve. This compensatory effect has been found particularly in school-based interventions (e.g., Verdurmen, Koning, Vollebergh, Van den Eijnden, & Engels, 2014).

Personality traits could also affect the extent to which individuals can effectively transfer skills learned during an intervention to their daily lives. Actually applying the new skills in daily life facilitates generalization of the skills which possibly increases intervention effects. In particular, Conscientiousness and Openness to experiences could be relevant for transferring skills. Individuals with high levels of Conscientiousness are generally able to inhibit initial responses, are persevering, and are planful (Shiner & Caspi, 2003). Participants with these characteristics might be better able to inhibit initial responses of old behaviors and deliberately implement skills learned during the intervention than participants with lower levels of Conscientiousness. Individuals with high levels of Openness to experiences are characterized by an open mind and an interest in novel experiences (Shiner & Caspi, 2003). They tend to be flexible and have the ability to acquire and consider new perspectives and skills (Connor-Smith & Flachsbart, 2007). Individuals with high levels of Openness might have a more positive and open attitude towards the intervention and are more interested in its novelties. Their high interest together with a flexible attitude could make it easier for them to implement the newly learned skills than for participants with lower levels of Openness. Thus, individuals with high levels of Conscientiousness or Openness to experiences might benefit more from an intervention, as they might apply, and generalize, the learned skills more easily in their daily lives than individuals with low levels of these traits.

Previous intervention research showed that personality traits can indeed moderate intervention effects. However, the results do not show a clear pattern of traits that affect intervention effects in the intra- and interpersonal domains in general or on specific competencies and problems. In the intrapersonal domain, for instance, Senf and Liau (2013) found stronger intervention effects for individuals with high levels of Extraversion on both happiness and depressive symptoms. Additionally, stronger intervention effects on happiness were found for individuals with high levels of Openness to experiences. Huppert and Johnson (2010) found no moderating effects of Extraversion and Openness on wellbeing and resilience, but did find stronger intervention effects on wellbeing for individuals with high levels of Agreeableness and Neuroticism. De Vibe and colleagues (2015) also found stronger intervention effects on wellbeing for individuals with high levels of Neuroticism, but not for Agreeableness. Contrarily, Wang and colleagues (2017) found no personality traits that moderated intervention effects on wellbeing. In the interpersonal domain, Stoltz and

colleagues (2013) found stronger intervention effects on both proactive aggression and delinquent behavior for individuals with average levels of Conscientiousness and on reactive aggression for individuals with lower levels of Extraversion. Asscher and colleagues (2016) found stronger intervention effects on Conduct disorder and rule breaking behavior for individuals with higher levels of Agreeableness and Conscientiousness, but no effect of personality traits on aggression in general. In conclusion, a clear pattern of which personality traits are important to consider in interventions is lacking.

Perhaps the lack of a pattern in results concerning personality traits moderating intervention effects might be explained by the type of intervention or the design of the study. Some examined interventions were school-based (e.g., Huppert & Johnson, 2010; Stoltz et al., 2013), whereas others were online (e.g., Wang et al., 2017) or in the clinical field (e.g., Asscher et al., 2016). Some examined interventions were universal (e.g., De Vibe et al., 2015; Senf & Liau, 2013), while others were selective, i.e. targeting youth who already show problem behavior (e.g., Asscher et al., 2016; Stoltz et al., 2013). Some studies focused on the intrapersonal domain (e.g., Huppert & Johnson, 2010; Senf & Liau, 2013), and others on the interpersonal domain (e.g., Asscher et al., 2016; Stoltz et al., 2013). As a first step, it is important to examine universal interventions to gain general insights in which personality traits are important moderators of intervention effects and for which, general and specific, outcomes (Spoth et al., 2006). Universal interventions are eminently suited for this aim as they often aim to address a broad range of outcomes and include a general population (Farrell, Henry, & Bettencourt, 2013).

Our study adds to the literature in two ways. First, we examined the moderating role of personality traits across a broad range of competencies and problems in both the intrapersonal and interpersonal domains within a universal intervention. This enabled us to examine whether certain personality traits showed a pattern of moderating intervention effects across multiple specific competencies and problems and across both domains in general. Such a pattern would indicate that certain traits might be important to consider in interventions. Second, we analyzed the effect of personality traits on *trajectories of change* during the intervention. Previous research has examined personality as a moderator on a specific time point. Although this analytical procedure can show moderation of intervention effects by personality traits at that specific time point (e.g., the post or follow-up measurement), it does not show differences in change during the intervention that might be dependent on a personality trait. By examining differences in trajectories between subgroups, defined by different levels of a personality trait, we combined a person-centered approach with a variable-centered approach allowing us to analyze change within and between individuals (Greenberg & Abenavoli, 2017).

In the present study, we studied personality traits as moderators of trajectories of change during a universal and widely implemented intervention, Rock & Water (R&W; Ykema, 2002, 2014). R&W is a school-based intervention aiming to

stimulate adolescents' competencies and prevent the development of problems in the intrapersonal (i.e., psychological wellbeing, resilience, sexual autonomy, and internalizing behavior) and interpersonal domain (i.e., interpersonal relations in the class, externalizing behavior, aggression, bullying, and victimization). We hypothesized that more vulnerable adolescents – based on their levels of personality traits, i.e., low levels of Extraversion, Agreeableness, or Conscientiousness, or high levels of Neuroticism – benefit more from the intervention than less vulnerable adolescents. Furthermore, we hypothesized that adolescents who can effectively transfer skills learned during an intervention to their daily lives benefit more from the intervention (e.g., high levels of Conscientiousness or Openness to experiences).

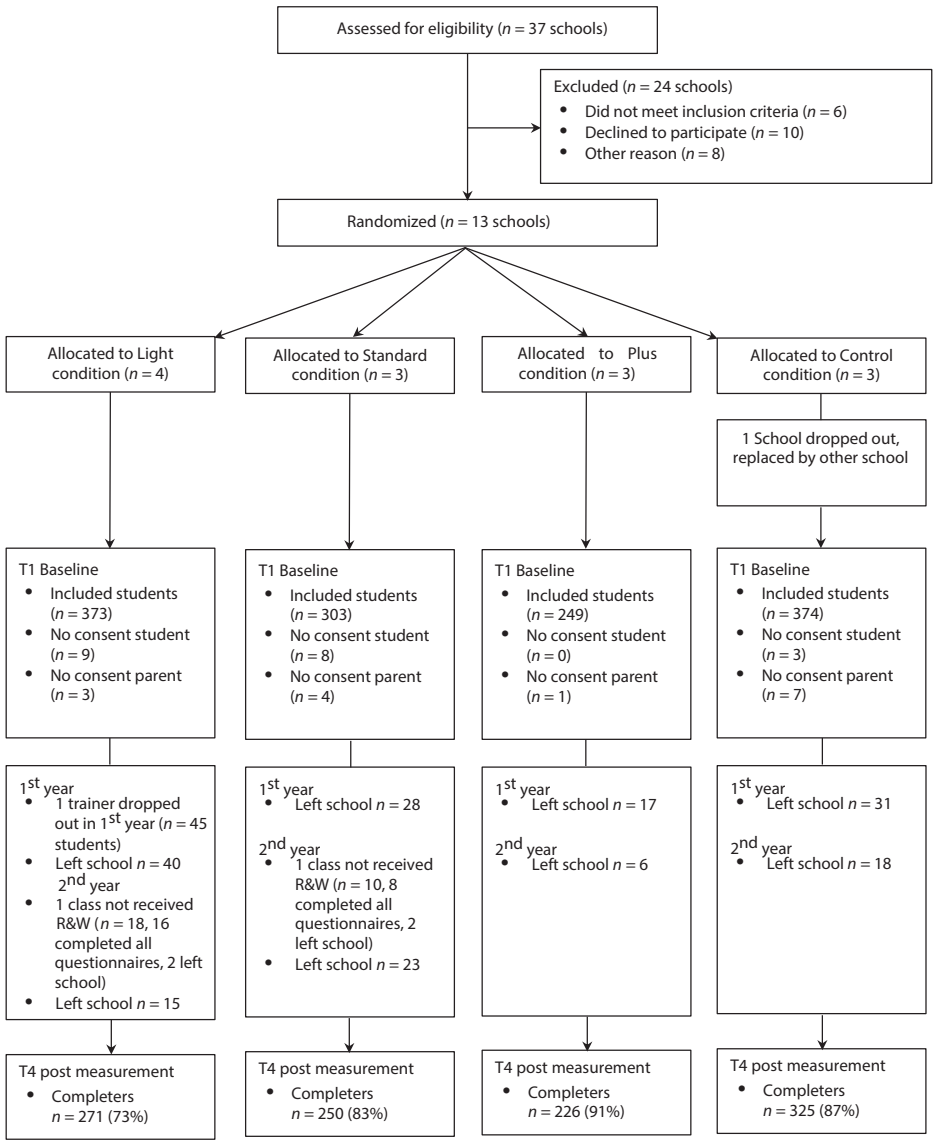
Method

Procedure and Design

We examined the effectiveness of R&W by means of a randomized controlled trial with three intervention conditions and one control condition. The intervention was examined in three different conditions as part of a larger project. One of the aims of this project was to examine whether the width of the intervention's ecological focus (i.e., the extent to which multiple systems are involved in the intervention) affected intervention effects. These different intervention conditions were not presumed to influence the moderating effect of personality. In the 'Light' condition, only a core team of teachers was involved in the intervention. In the 'Standard' condition, the entire teaching staff was involved. In the 'Plus' condition, the entire teaching staff and parents were involved. In the Control condition, current school policy to improve students' competencies and prevent problems in the intra- and interpersonal domain was implemented (i.e., care as usual).

Interested schools were screened for eligibility by the researchers. Special education schools and schools that had implemented R&W in the last two years were excluded. Thirteen schools in urban and rural areas in the Netherlands were included. An online number generator was used to randomize the schools to a condition stratified by school size (< 100 students in 7th Grade, > 100 students in 7th Grade) to enhance an equal distribution of adolescents. After randomization and before data collection started, one control school dropped out due to changed school management and was replaced by another school (see Figure 1 for the flow chart).

Figure 1. Flow chart.



Adolescents completed digital questionnaires at four time points: At baseline (T1; October/November 2017), after the first year of the intervention (T2; March/April 2018), before the start of the second year of the intervention (T3; October 2018), and at post measurement (T4; January 2019). Questionnaires were conducted by trained research assistants. Parents completed a digital questionnaire at baseline.

Adolescents and parents gave active informed consent for completing the questionnaires. In addition, parents gave passive informed consent for the participation of their child. This trial was approved by the Ethical Committee of the Faculty of Social and Behavioral Sciences of Utrecht University (FETC17-05) and registered in the Dutch Trial Register, number NL6371 (NTR6554; see for protocol Mertens, Deković, Van Londen, & Reitz, 2018).

Participants

At baseline, 1299 adolescents in the 7th Grade participated in this study with an average age of 12.38 years (*SD* = .62). In total, 661 (54%) adolescents were boys, and 815 (69%) adolescents had a Western background. Regarding the parents, 461 parents completed the questionnaires. Overall, parents' ages ranged between 28 and 77 years old (*M* = 43.16, *SD* = 5.69) and they were mostly mothers (*n* = 382, 83%). The demographics are presented per condition in Table 1.

Conditions did not differ on adolescents' and parents' sex distribution, and parents' age. There were small differences between conditions on adolescents' age and ethnic background (see Table 1). In the Light condition, adolescents were slightly younger than in the Control condition. Regarding adolescents' ethnic background, in the Light and Plus conditions, the majority were adolescents with a Western background, whereas in the Standard and Control conditions ethnic background was roughly equally distributed.

Missing data. Data were not missing completely at random according to the Little's MCAR test (χ^2 (1507) = 1606.32, p = .037). However, attrition analyses showed no differences between drop-outs and completers on the demographic variables (Age: $F(1, 1230) = .24$, $p = .626$, $\eta^2_{partial} = .000$; Sex: $\chi^2(1) = 1.36$, $p = .244$, $\varphi = -.033$; Ethnic background: $\chi^2(1) = 1.19$, $p = .276$, $\varphi = .032$) nor on the outcome variables, analyzed with MANOVAs per time point (T1: $F(27, 3609) = 1.32$, $p = .126$, $\eta^2_{partial} = .010$; T2: $F(27, 3378) = .65$, $p = .916$, $\eta^2_{partial} = .005$; T3: $F(18, 2044) = .82$, $p = .674$, $\eta^2_{partial} = .007$). As we found no differences between drop-outs and completers and because the Little's MCAR test is conservative when used on a large set of variables, we regarded the missing data as missing at random (Van Ness, Murphy, Araujo, Pisani, & Allore, 2007).

Table 1 Descriptives of Adolescents' and Parents' Demographics per Condition

	Light	Standard	Plus	Control	Differences at T1		
					F/χ^2	p	$\eta^2_{partial}/\varphi$
Demographics adolescents *(n)*	373	303	249	374			
Age, *M* (*SD*)	12.33 (.57)	12.38 (.66)	12.34 (.60)	12.47 (.64)	3.89	.009	.009
Boys, *n* (%)	170 (48%)	161 (56%)	131 (55%)	199 (57%)	7.38	.061	.077
Western background, *n* (%)	291 (82%)	115 (43%)	211 (91%)	198 (59%)	182.01	<.001	.392
Demographics parents *(n)*	178	66	107	110			
Age, *M* (*SD*)	43.11 (5.66)	41.82 (5.45)	43.85 (5.14)	43.39 (6.30)	1.82	.143	.012
Mothers, *n* (%)	155 (87%)	55 (83%)	84 (79%)	88 (80%)	4.30	.230	.097

Conditions

R&W. The theory of R&W is based on the 'R&W house'. This house consists of five levels that together represent the intra- and interpersonal domains. R&W theorizes that adolescents can develop themselves within the intra- and interpersonal domains by improving the foundation of the house (i.e., self-control, self-reflection, and self-esteem) and, closely related to that foundation, emotional self-regulation.

R&W uses the symbolic principles of 'rock' and 'water' to indicate a continuum ranging from an uncompromising attitude to resist peer pressure (i.e., rock) to a flexible attitude open to the opinions, thoughts, and feelings of others (i.e., water). The intervention uses a unique combination of a physical approach and the more common psychological approach, i.e., a psychophysical approach. Through play and exercises adolescents increase their strength, learn how to make contact with others, and explore, respect, and set own and other's boundaries. During the lessons, trainers create a safe and trusting environment in which there is respect for each other, there is room to make mistakes, and adolescents work together. Adolescents participate in physical exercises and games, reflect on the exercises, share their thoughts with each other, and address how to use the learned skills in their daily lives.

The R&W intervention is a 2-year manualized program of which the first year consists of 14 lessons and the second year of 8 lessons. Each R&W lesson lasts 90 minutes and is weekly implemented during physical education classes. R&W trainers were teachers that successfully completed the 3-day training course to become certified R&W trainers. Most teachers were physical education teachers due to the physical nature of the exercises during the lessons (see the study protocol for more information about the intervention Mertens et al., 2018).

Intervention Fidelity. Fidelity to the intervention manual was measured with questionnaires completed by R&W trainers and with observations conducted by R&W experts. R&W trainers reported that in general they could complete the lessons (65%),

and did not deviate much from the manual (72%). Three R&W experts observed 67 R&W lessons and reported that most observed lessons were (almost) completed (86%), trainers did not deviate much from the manual (91%), and adjustments were seen as improvements (62%). In conclusion, the intervention was implemented with moderate to high fidelity to the manual.

Control. Care as usual differed between the schools. In one school, it entailed teachers as personal coaches. Adolescents had regular meetings, discussed their wellbeing, and got advice from their coach. In a second school it entailed that adolescents participated in a project week 'being different', had class discussions about bullying, signed an anti-bullying contract, and had a mentor (teacher) they could contact if they experienced difficulties. Another school had appointed an 'anti-bullying coordinator' who organized activities to prevent or stop bullying and adolescents had a mentor (teacher) they could contact if they experienced difficulties.

Outcomes

Intrapersonal domain

Psychological wellbeing. The presence of positive emotions was measured with the subscale Psychological wellbeing of the KIDSCREEN-27 (Ravens-Sieberer & The European KIDSCREEN Group, 2006). The subscale contains 7 items (e.g., "Past week, have you been in a good mood?") answered on a 5-point Likert-type scale (1 = *never* to 5 = *always*). Some items were recoded with higher scores indicating more psychological wellbeing (Cronbach's α = .76 - .83).

Resilience. The ability to bounce back from challenges that can arise in life was measured with the Connor-Davidson Resilience Scale – short version (Davidson & Connor, 2017). The questionnaire contains 10 items (e.g., "Can deal with whatever comes.") answered on a 5-point Likert-type scale (0 = *not true at all* to 4 = *true nearly all the time*; Cronbach's α = .79 - .92).

Sexual autonomy. Coping skills in sexual situations was measured with 5 items from the study Sex under 25 (e.g., "When I am with someone I like, I feel at ease."; De Graaf, Meijer, Poelman, & Vanwesenbeeck, 2005) answered on a 4-point Likert-type scale (1 = *never* to 4 = *always*). Some items were recoded with higher scores indicating more sexual autonomy. Reliability was poor at T1 (Cronbach's α = .53) and adequate at T2, T3, and T4 (Cronbach's α = .62 - .65).

Internalizing behavior. The presence of internalizing problems was assessed using the internalizing subscale of the short version of the Youth Self Report (YSR; Chorpita et al., 2010). The subscale contains 6 items (e.g., "I feel worthless.") answered on a 3-point Likert-type scale (0 = *never* to 2 = *often*; Cronbach's α = .79 - .86).

Interpersonal domain

Interpersonal relations in the class. Negative social exchanges between classmates, the extent to which adolescents feel comfortable around their classmates, and the unity and inclusiveness among classmates was measured using the Classroom Peer Context Questionnaire (Boor-Klip, Segers, Hendrickx, & Cillessen, 2016). The questionnaire contains 12 items (e.g., "In this class students like each other.") rated on a 5-point Likert-type scale (1 = *totally not true* to 5 = *completely true*). Some items were recoded with high scores indicating more positive interpersonal relations in the class (Cronbach's α = .80 - .85).

Externalizing behavior. The presence of externalizing problems was assessed with the externalizing subscale of the short version of the YSR (Chorpita et al., 2010). The subscale contains 6 items (e.g., "I destroy things.") rated on a 3-point Likert-type scale (0 = *never* to 2 = *often*; Cronbach's α = .65 - .79).

Aggression. Aggression was measured with the Reactive and Proactive Aggression Questionnaire (Dodge & Coie, 1987). The questionnaire contains 6 items (e.g., "If they tease me, I get angry.") answered on a 5-point Likert-type scale (1 = *never* to 5 = *almost always*; Cronbach's α = .65 - .83).

Bullying and victimization. The frequency of bullying and experienced victimization was measured with the 2 global items of the Olweus Bully/Victim Questionnaire (Solberg, & Olweus, 2003): "How often have you taken part in bullying others?" and "How often have you been bullied?" answered on a 5-point Likert-type scale (1 = *never* to 5 = *almost always*). The items were preceded by a definition of bullying.

Moderator

Personality. To assess adolescents' personality, parents completed the Quick Big Five (Goldberg, 1992). This questionnaire contains 30 items (e.g., nice, sympathetic, organized) representing the 5 personality traits: Extraversion, Agreeableness, Conscientiousness, Neuroticism, and Openness to experiences, with each trait assessed by 6 items. Parents indicated to which extent that characteristic suites the adolescent on a 7-point Likert-type scale (1 = *not at all* to 7 = *very well*; Cronbach's α = .80 - .91).

Statistical Analyses

Analyses were conducted in M*plus* version 8.2 (Muthén & Muthén, 2010) with an intention-to-treat approach, i.e., all adolescents who participated at baseline were analyzed regardless whether they actually received the intervention or not. The nested structure of the data was taken into account by using the complex sample cluster

feature of M*plus* to correct for clustering at school level. This is a conservative clustering procedure to gain unbiased estimates of the standard errors (Muthén & Muthén, 2010). Clustering at class level was not taken into account as class composition was not stable over the years (e.g., Cross et al., 2016). Full Information Maximum Likelihood (FIML) procedures were used to include all participants in the model with Robust Maximum Likelihood estimation (MLR) for parameter estimates as this estimator is robust to non-normality and non-independence of the data (Muthén & Muthén, 2010).

For our analyses we used Latent Growth Curve (LGC) models. LGC models estimate individual growth curves and use these curves as indicators of latent variables (i.e., intercept and slope) to estimate average group growth trajectories (Muthén & Muthén, 2010). This approach is recommended by Greenberg and Abenavoli (2017) for analyzing universal interventions as these models have the potential to demonstrate (small) changes in the population curve and examine differences in trajectories between potential subgroups.

To examine adolescents' personality traits as moderators of intervention effects we analyzed LGC models with unspecified growth and interaction effects. Not specifying the growth allowed us to examine nonlinear growth. We fixed the factor loading of T1 at 0 and of T4 at 3. The factor loadings of T2 and T3 were determined by the data (Duncan & Duncan, 2004). For the interaction effects, we created three dummy variables, each representing a R&W condition compared to the Control condition, and grand mean centered the scales representing the personality traits. Subsequently, we created interaction effects of the centered trait and each dummy variable (e.g., Extraversion X Light, Extraversion X Standard, Extraversion X Plus). In the LGC model, we regressed the intercept and slope on the three condition dummy variables and the trait. Additionally, we regressed the slope on the interaction effects. A significant interaction effect indicated moderation by that trait. Adolescents' age and ethnic background were added as covariates. We estimated an LGC model per trait for each outcome separately.

In the case of a significant interaction, we conducted multigroup LGC models in which we split the sample in three groups representing adolescents scoring low, average, and high ($M \pm 1$ *SD*) on the concerned personality trait. Growth in these multigroup LGC models was specified based on the growth estimated in unspecified growth LGC models including the three dummy variables and the two covariates. All parameters were constrained to be equal across the groups, except for the slopes regressed on the dummy variables. Using these estimates we calculated effect sizes for the low, average, and high scoring groups by multiplying the rate of change by time span of the intervention as modeled in the LGC models (i.e., factor loading of T4) divided by the standard deviation of the concerned outcome (d = (slope * duration) / SD; Feingold, 2013).

Table 2 Descriptives of Outcomes and Moderator per Condition per Time Point

	R&W Light				R&W Standard				R&W Plus				Control			
	T1	T2	T3	T4	T1	T2	T3	T4	T1	T2	T3	T4	T1	T2	T3	T4
	M (*SD*)	*M* (*SD*)	*M* (*SD*)	*M* (*SD*)	*M* (*SD*)	*M* (*SD*)	*M* (*SD*)	*M* (*SD*)	*M* (*SD*)	*M* (*SD*)	*M* (*SD*)	*M* (*SD*)	*M* (*SD*)	*M* (*SD*)	*M* (*SD*)	*M* (*SD*)
Intrapersonal domain																
Psychological wellbeing	3.94 (.62)	3.98 (.69)	3.91 (.71)	3.87 (.78)	4.02 (.61)	4.10 (.71)	4.02 (.73)	3.93 (.77)	3.97 (.62)	4.06 (.63)	3.87 (.72)	3.85 (.80)	4.06 (.59)	4.02 (.68)	3.90 (.75)	3.92 (.78)
Resilience	2.36 (.65)	2.49 (.75)	2.15 (.87)	2.47 (.82)	2.43 (.67)	2.70 (.71)	2.56 (.70)	2.53 (.90)	2.43 (.64)	2.57 (.68)	2.47 (.71)	2.51 (.84)	2.51 (.65)	2.68 (.69)	2.57 (.74)	2.61 (.85)
Sexual autonomy	3.23 (.44)	3.39 (.48)	3.40 (.44)	3.40 (.46)	3.30 (.44)	3.44 (.52)	3.38 (.54)	3.39 (.51)	3.29 (.40)	3.38 (.46)	3.34 (.45)	3.35 (.57)	3.32 (.44)	3.36 (.49)	3.32 (.50)	3.36 (.52)
Internalizing behavior	.53 (.42)	.41 (.44)	.44 (.44)	.42 (.46)	.44 (.42)	.33 (.39)	.40 (.44)	.37 (.43)	.49 (.40)	.45 (.42)	.51 (.44)	.48 (.51)	.43 (.39)	.39 (.41)	.45 (.46)	.42 (.44)
Interpersonal domain																
Interpersonal relations in the class	3.90 (.64)	3.81 (.68)	3.83 (.67)	3.73 (.69)	3.98 (.63)	3.87 (.72)	3.81 (.73)	3.80 (.75)	4.07 (.62)	4.00 (.63)	3.92 (.69)	3.84 (.81)	3.91 (.59)	3.78 (.65)	3.77 (.67)	3.75 (.72)
Externalizing behavior	.59 (.38)	.47 (.39)	.48 (.39)	.45 (.41)	.54 (.39)	.43 (.38)	.50 (.44)	.42 (.42)	.55 (.35)	.46 (.36)	.55 (.38)	.49 (.45)	.49 (.34)	.42 (.37)	.45 (.39)	.43 (.39)
Aggression	1.79 (.56)	1.71 (.59)	1.69 (.59)	1.66 (.66)	1.85 (.61)	1.84 (.73)	1.94 (.73)	1.83 (.81)	1.73 (.46)	1.72 (.60)	1.80 (.58)	1.74 (.81)	1.78 (.56)	1.76 (.65)	1.81 (.67)	1.85 (.78)
Bullying	1.13 (.48)	1.18 (.61)	1.15 (.55)	1.21 (.64)	1.13 (.41)	1.18 (.60)	1.15 (.58)	1.19 (.64)	1.10 (.41)	1.15 (.47)	1.20 (.68)	1.20 (.64)	1.08 (.36)	1.18 (.59)	1.22 (.71)	1.22 (.71)
Victimization	1.32 (.77)	1.43 (.96)	1.30 (.77)	1.31 (.86)	1.28 (.73)	1.35 (.87)	1.22 (.77)	1.22 (.70)	1.36 (.88)	1.37 (.85)	1.40 (.96)	1.40 (.93)	1.37 (.94)	1.56 (1.16)	1.41 (1.04)	1.34 (.95)
Personality																
Extraversion	4.97 (1.26)				5.05 (1.20)				4.98 (1.24)				4.83 (1.19)			
Agreeableness	6.15 (.66)				6.15 (.66)				6.20 (.56)				6.15 (.66)			
Conscientiousness	3.84 (1.31)				4.22 (1.37)				4.20 (1.27)				3.86 (1.22)			
Neuroticism	3.58 (1.20)				3.33 (1.18)				3.55 (1.18)				3.42 (1.03)			
Openness	4.99 (1.08)				4.93 (.98)				5.06 (1.03)				4.94 (1.09)			

Table 3 Correlations at Baseline between Outcomes and Moderators

	1.	2.	3.	4.	5.	6.	7.	8.	9.	10.	11.	12.	13.
Intrapersonal domain													
1. Psychological wellbeing	-												
2. Resilience	.40**	-											
3. Sexual autonomy	.51**	.30**	-										
4. Internalizing behavior	-.60**	-.34**	-.49**	-									
Interpersonal domain													
5. Interpersonal relations in the class	.40**	.23**	.46**	-.31**	-								
6. Externalizing behavior	-.26**	-.13**	-.29**	.29**	-.28**	-							
7. Aggression	-.13**	-.04	-.23**	.02	-.29**	.47**	-						
8. Bullying	-.06**	-.03	-.15**	.03	-.18**	.17**	.20**	-					
9. Victimization	-.26**	-.16**	-.21**	.25**	-.34**	.11**	.12**	.23**	-				
Moderator													
10. Extraversion	.18**	.19**	.16**	-.21**	.07	.04	-.06	-.01	-.03	-			
11. Agreeableness	.12*	.07	.10*	-.01	.12*	-.12*	-.11*	.02	.01	.27**	-		
12. Conscientiousness	.05	-.02	.08	.02	.01	-.14**	-.08	-.05	.03	-.07	.23**	-	
13. Neuroticism	-.31**	-.25**	-.24**	.28**	-.14**	.12*	.01	.05	-.01	-.48**	-.17**	-.06	-
14. Openness	.07	-.04	.05	.06	-.01	-.03	-.03	.05	-.05	.21**	.44**	.28**	-.08

Note. * $p < .05$; ** $p < .01$

Results

Descriptives of the outcomes per time point and moderator at baseline are reported in Table 2 per condition. In Table 3 the correlations between outcomes and adolescents' personality traits are presented. Personality traits were more strongly correlated with outcomes in the intrapersonal than in the interpersonal domain.

Intervention Effects

The main effects of the intervention conditions, the main effects of the personality traits, and the corresponding interaction effects are reported in Table 4.

Compared to the Control condition, adolescents' trajectories in the Light condition changed in the desired direction on several outcomes in the intrapersonal (i.e., psychological wellbeing, sexual autonomy, and internalizing behavior) and interpersonal (i.e., externalizing behavior and aggression) domains. Additionally, adolescents in the Standard condition had a slightly more beneficial trajectory for

internalizing behavior than adolescents in the Control condition. No intervention effects were found for the other outcomes (i.e., resilience, interpersonal relations in the class, bullying, and victimization) or for adolescents in the Plus condition.

Personality Traits Effects

Extraversion. Extraversion was negatively related to change rates in resilience and sexual autonomy and positively related to change rate in aggression, regardless of condition. This means that higher levels of Extraversion predicted a less steep increase in resilience and sexual autonomy, and a less steep decrease in aggression over time. Concerning moderation of intervention effects, Extraversion significantly moderated intervention effects on sexual autonomy and aggression in the Standard and Plus conditions. Adolescents with high levels of Extraversion improved more on sexual autonomy (Standard: d_{low} = .32, $d_{average}$ = .46, d_{high} = .74; Plus: d_{low} = -.10, $d_{average}$ = .16, d_{high} = .59). For adolescents with lower levels of Extraversion, the intervention appeared less successful than CAU in decreasing aggression (Standard: d_{low} = -.41, $d_{average}$ = -.16, d_{high} = .38; Plus: d_{low} = -.28, $d_{average}$ = -.21, d_{high} = .01). No moderation effects were found for the other outcomes or in the Light condition.

Agreeableness. Agreeableness was positively related to aggression, regardless of condition, indicating that more Agreeableness was related to a less steep decrease in aggression over time. Furthermore, Agreeableness moderated intervention effects on internalizing behavior and aggression in the Standard condition. Adolescents with lower levels of Agreeableness decreased more in internalizing behavior (d_{low} = .98, $d_{average}$ = -.26, d_{high} = -.08). In contrast, the intervention was less successful than CAU in decreasing aggression for adolescents with low levels of Agreeableness (d_{low} = -.71, $d_{average}$ = -.07, d_{high} = .28). No moderation effects were found for the other outcomes or in the other conditions.

Conscientiousness. Conscientiousness was negatively related to internalizing behavior and bullying, and positively related to positive perceptions of interpersonal relations in the class, regardless of condition. This indicated that higher levels of Conscientiousness were related to a less steep increase in internalizing behavior and bullying, and a less steep decrease in positive perceptions of interpersonal relations in the class over time. Conscientiousness significantly moderated intervention effects on internalizing behavior in the Standard condition and bullying in the Light condition. Regarding internalizing behavior, adolescents with low to average levels of Conscientiousness improved more (Standard: d_{low} = .20, $d_{average}$ = .47, d_{high} = .09). In addition, adolescents with low levels of Conscientiousness decreased more on bullying (Light: d_{low} = .37, $d_{average}$ = -.08, d_{high} = .02). No moderation effects were found for the other outcomes or in the Plus condition.

Table 4 Standardized Main and Interaction Effects of Condition and Personality on Slope in the Intrapersonal and Interpersonal Domains

	Intrapersonal domain								Interpersonal domain							
	Psychological wellbeing		Resilience[1]		Sexual autonomy[2]		Internalizing behavior[2]		Interpersonal relations		Externalizing behavior[2]		Aggression		Bullying		Victimization	
	β	*p*	β	*p*	β	*p*	β	*p*	β	*p*	β	*p*	β	*p*	β	*p*	β	*p*
Light	.15	.009	-.03	.519	.18	<.001	-.13	<.001	.05	.057	-.10	<.001	-.18	.022	-.12	.063	-.01	.863
Standard	.08	.267	.05	.459	.08	.138	-.07	.050	-.05	.514	-.04	.215	.01	.926	-.12	.160	-.04	.568
Plus	.04	.172	-.02	.459	.06	.065	.01	.772	-.04	.281	.01	.648	-.02	.752	-.05	.622	.03	.340
Extraversion	-.01	.898	-.14	<.001	-.12	<.001	-.02	.643	-.01	.710	.03	.663	.28	.005	.14	.504	-.12	.197
Extraversion x Light	-.09	.077	.08	.102	-.01	.677	.05	.084	-.06	.124	.01	.688	-.06	.184	-.01	.919	.05	.300
Extraversion x Standard	-.05	.394	.08	.065	.03	.035	.02	.312	-.03	.052	.01	.688	-.12	.006	-.02	.796	-.01	.660
Extraversion x Plus	-.10	.097	.05	.114	.03	.005	-.01	.728	.05	.226	.04	.198	-.07	.046	-.10	.280	-.04	.263
Agreeableness	.03	.450	.02	.703	-.01	.908	-.10	.220	.04	.590	.03	.691	.16	.001	-.02	.787	-.08	.341
Agreeableness x Light	-.09	.159	.01	.754	-.02	.795	.02	.623	-.09	.085	.01	.836	-.07	.098	-.07	.326	-.03	.735
Agreeableness x Standard	-.04	.342	.00	.991	.02	.714	.07	.043	.00	.999	.01	.714	-.09	.009	.01	.810	-.04	.471
Agreeableness x Plus	.00	.946	.01	.725	-.04	.374	.12	.103	-.08	.053	.04	.332	-.01	.504	-.01	.811	.03	.445
Conscientiousness	.03	.537	.00	.942	.01	.921	-.10	.005	.12	.049	-.01	.891	.03	.748	-.28	.004	-.07	.196
Conscientiousness x Light	-.02	.682	.05	.359	.01	.644	.05	.061	-.07	.106	-.01	.773	-.04	.503	.10	.011	.10	.083
Conscientiousness x Standard	-.01	.914	.06	.120	-.01	.712	.06	.034	.05	.271	.01	.526	-.05	.504	.04	.182	.02	.553
Conscientiousness x Plus	-.08	.133	.02	.525	-.03	.258	.09	.101	-.00	.901	-.00	.979	-.01	.769	.09	.063	.02	.644

Table 4 Continued.

	Intrapersonal domain															
	Psychological wellbeing		Resilience[1]		Sexual autonomy[2]		Internalizing behavior[2]		Interpersonal relations		Externalizing behavior[2]		Aggression		Bullying		Victimization	
	β	*p*	β	*p*	β	*p*	β	*p*	β	*p*	β	*p*	β	*p*	β	*p*	β	*p*
Neuroticism	-.18	.022	.12	.111	.02	.812	.07	.329	.03	.695	.01	.838	.05	.691	.11	.269	.10	.346
Neuroticism x Light	.22	.020	-.13	.021	.06	.260	-.09	.095	.08	.229	-.06	.293	-.08	.414	-.09	.206	-.04	.550
Neuroticism x Standard	.04	.314	-.10	.040	.03	.389	-.02	.453	.01	.729	.05	.262	.00	.995	-.02	.580	-.00	.966
Neuroticism x Plus	.09	.053	-.00	.951	.00	.931	-.05	.222	-.07	.138	-.06	.170	-.06	.297	-.07	.243	-.04	.509
Openness	-.11	.324	.13	<.001	-.08	.007	-.08	.134	.03	.674	.04	.177	.14	.195	-.10	.107	-.05	.683
Openness x Light	-.04	.507	-.05	.117	.00	.717	.07	.042	.00	.969	-.01	.854	-.02	.770	.07	.124	.06	.092
Openness x Standard	.01	.916	-.01	.743	-.01	.628	.02	.340	.00	.969	.00	.954	-.07	.151	.03	.414	-.01	.778
Openness x Plus	-.07	.251	-.02	.252	-.03	.275	.12	.110	-.07	.064	.00	.965	-.01	.737	.00	.969	-.02	.583

Note. Main effects of the intervention conditions were the same in all models, therefore these main effects are only reported for the models concerning the personality trait 'Extraversion'. [1]For convergence, growth specified based on factor loadings of unspecified LGC model with dummy variables and covariates. [2]Variance of the baseline measurement of the concerned outcome variable was fixed to zero due to a negative residual variance of this observed variable (e.g., Hukkelberg, & Ogden, 2013).

Neuroticism. Neuroticism was negatively related to psychological wellbeing, regardless of condition, indicating that higher levels of Neuroticism predicted a less steep increase in psychological wellbeing over time. In addition, Neuroticism moderated intervention effects on psychological wellbeing in the Light condition and on resilience in the Light and Standard conditions. Adolescents with high levels of Neuroticism improved more on psychological wellbeing (d_{low} = -.33, $d_{average}$ = -.07, d_{high} = .74). Contrarily, the intervention was less successful than CAU in improving resilience for adolescents with high levels of Neuroticism (Light: d_{low} = .06, $d_{average}$ = -.02, d_{high} = -.89; Standard: d_{low} = .52, $d_{average}$ = .03, d_{high} = -.60). No moderation effects were found for the other outcomes or in the Plus condition.

Openness to experiences. Openness to experiences was positively related to resilience and negatively related to sexual autonomy, regardless of condition. This means that more Openness to experiences was related to a less steep decrease in resilience and a less steep increase in sexual autonomy over time. Furthermore, Openness moderated intervention effects on internalizing behavior in the Light condition. Adolescents with average levels of Openness decreased more on internalizing behavior (d_{low} = -.12, $d_{average}$ = .41, d_{high} = .06). No moderation effects were found for any of the other outcomes or in the other conditions.

Discussion

The present study expanded previous research by examining the effects of personality traits on trajectories of change during a universal intervention across a broad range of competencies and problems in the intra- and interpersonal domains. Overall, the intervention appeared moderately effective in improving several aspects of both adolescents' intra- (i.e., psychological wellbeing, sexual autonomy, and internalizing behavior) and interpersonal (i.e., externalizing behavior and aggression) domains. Moderate intervention effects were expected based on meta-analyses of school-based interventions (e.g., Cohen's *d* = .22 - .27, Durlak, Weissberg, Dymnicki, Taylor, & Schellinger; Cohen's *d* = .10 - .25, Mertens, Deković, Leijten, Van Londen, & Reitz, 2020). Personality traits affected only a few intervention effects which confirms the universality of the intervention. In general, three patterns of moderation emerged. First, there were indications that adolescents vulnerable to develop problems in the intra- and interpersonal domains – based on their levels of certain personality traits – benefitted more from the intervention than less vulnerable adolescents. Second, Extraversion appeared to be an important moderator of intervention effects in interventions requiring sociability from participants. Third, personality traits seemed to affect competencies and problems in the intrapersonal domain somewhat more than in the interpersonal domain, as predictors as well as moderators of intervention effects.

Adolescents who seemed more vulnerable to develop problems in the intra- and interpersonal domains based on their levels of Extraversion, Agreeableness, Conscientiousness, or Neuroticism appeared to benefit most from the intervention. In the present study, regardless of condition, adolescents with high levels of Extraversion increased less in sexual autonomy and decreased less in aggression, with high levels of Agreeableness also decreased less in aggression, with high levels of Neuroticism increased less in psychological wellbeing, and with low levels of Conscientiousness increased more in internalizing behavior and bullying. Hence, these levels of personality traits predicted less beneficial trajectories of change making the adolescents more vulnerable to develop problems in these outcomes. However, particularly the adolescents who were vulnerable to develop problems on a certain outcome benefitted more from R&W on this outcome than less vulnerable adolescents. This finding confirms the Risk moderation hypothesis that individuals who experience more problems can benefit more from interventions (Spoth et al., 2006). For (clinical) practice this conclusion has important implications as it suggests that the most vulnerable individuals benefit most from interventions. However, in the present study this pattern of moderation has to be interpreted with caution as the main and moderation effects were scattered across personality traits, outcomes, and conditions, and has to be replicated before stronger conclusions can be drawn.

Furthermore, in contrast to our findings, previous research suggested that low rather than high levels of Extraversion and Agreeableness were related to more problems. This difference might be explained by differences in context between studies. For instance, "talkative", a facet of Extraversion, could be considered a positive characteristic at a party, but a negative characteristic at school (Hughes et al., 2020).

Extraversion showed a consistent pattern of moderating intervention effects on sexual autonomy and aggression. Individuals with high levels of Extraversion are generally sociable, interested in other people, and focused on the outer world rather than inhibited, withdrawn, and focused on one's inner world (Pellegrino & Hilton, 2012). Their orientation towards the outer world seems to fit well with the intervention's characteristics which require social interactions, for instance when working together in physical games and exercises, when sharing thoughts in a group, and when addressing together how the learned skills can be transferred to their daily lives. These intervention characteristics might be less suited for more inhibited and withdrawn individuals hindering them to learn from the intervention. Our findings are in line with the study of Senf and Liau (2013) who also found stronger effects for extraverted individuals in their universal intervention, but contrary to the findings of Stoltz and colleagues (2013) who found stronger effects for less extraverted individuals in their selective intervention. Perhaps high levels of Extraversion are beneficial in universal interventions, whereas low levels of Extraversion are beneficial in selective interventions. More research is needed to clarify whether the moderating role of Extraversion is dependent on the characteristics of interventions, such as the level of

sociability required from participants and the intervention's selective character (e.g., universal, selective, indicated).

It is noteworthy that personality seems indeed somewhat more influential in the intrapersonal domain than in the interpersonal domain, both as predictor and as moderator of intervention effects. Especially intervention effects on internalizing behavior, a potentially problematic behavior in the intrapersonal domain, were affected by personality traits. Personality traits influence how one wants to feel and why (Huges, Kratsiotis, Niven, & Holman, 2020) and might therefore be more related to individuals' own feelings, emotions, and attitudes (i.e., the intrapersonal domain), than their interactions and social perceptions (i.e., the interpersonal domain). Our findings are in line with previous research. For instance, Van Leeuwen and colleagues (2004) found that personality traits were stronger predictors of internalizing behavior than of externalizing behavior. In addition, Hughes and colleagues (2020) showed that personality traits were more strongly linked to intrapersonal coping styles than to interpersonal coping styles. Thus, one should especially be aware of potentially different intervention effects for individuals differing in levels of certain personality traits when stimulating competencies and preventing problems in the intrapersonal domain.

Although we did not expect any differences in moderation effects between conditions beforehand, the different intervention conditions influenced intervention effects as well as moderation effects. We found only intervention effects in the condition in which few teachers were involved (i.e., Light condition). Even though it is often believed that intervention effects increase when more people are involved, this does not always seem to be the case. These results are supported by the meta-analysis of Durlak, Weissberg, Dymnicki, Taylor, and Schellinger (2011). Perhaps involving more people in an intervention lowers the responsibility people feel for effectively implementing the intervention, resembling a 'bystander effect' (Fischer et al., 2011), or results in mixed messages concerning intervention techniques (see Mertens, Deković, Van Londen, Nye, & Reitz, 2020), thereby counteracting possible intervention effects. For Extraversion, the moderation of intervention effects also differed between the conditions making these results more complex. Extraversion moderated intervention effects only in the conditions in which many people were involved in the intervention (i.e., Standard and Plus conditions). It appears that the R&W lessons alone (i.e., Light condition) were effective regardless of adolescents' level of Extraversion. In contrast, the involvement of more people in the intervention seems to have enabled extraverted adolescents to benefit from the intervention, possibly due to their sociability, whereas it counteracted the effects of the intervention lessons for less extraverted adolescents. Again, these findings underscore the importance of considering characteristics of an intervention when evaluating its effectiveness and for whom it is effective.

The present study has some strengths and limitations. Strengths of the study are the broad range of outcomes in both the intrapersonal and interpersonal domains, analyzing the effects of personality traits on change during an intervention, and

parental report of personality instead of self-report. A limitation is the relatively low number of parent reports on adolescents' personality. Missing data on parent reports were no issue when analyzing moderation with the interaction effects since missing data could be estimated in these models. However, these missing data could not be estimated in the multigroup models facilitating interpretation of the moderation effects. Furthermore, we ran 45 models with in total 135 interaction effects to examine moderation which could result in false positive results. We took this limitation into account by specifically focusing on moderation of intervention effects that showed a pattern across outcomes and conditions. It remains important that future research focuses more on the effects of personality traits on intervention effects and replicates our findings. In addition, it would be interesting to examine how combinations of personality traits influence intervention effects. For instance, future research could create a 'risk index' (e.g., Spoth et al., 2006) based on the number of present personality traits that make individuals more vulnerable to develop problems and study to what extent this risk index influences intervention effects.

In conclusion, our study showed that personality traits are potentially relevant for intervention effects, especially in the intrapersonal domain, but more research is needed. In general, most intervention effects were not affected by personality traits. As a universal intervention aims to target all subgroups in the total population equally well, few moderation effects are to be expected (Nehmy & Wade, 2014) and subscribe the universal character of R&W. Nevertheless, three patterns of moderation by personality traits emerged. First, the intervention might be especially effective for adolescents vulnerable to develop problems – based on their levels of personality traits. Second, Extraversion might be an important moderator of intervention effects when the intervention requires sociability from participants. Third, personality traits seem more important predictors and moderators in the intrapersonal domain than in the interpersonal domain. The present study highlights the importance of awareness that individuals can respond differently to interventions based on their levels of personality traits, especially in the intrapersonal domain. However, more research is needed concerning the role of personality traits in interventions' effectiveness and how this moderation interacts with intervention characteristics before interventions can be optimally tailored.

Mertens, E. C. A.,
Deković, M.,
Van Londen, M.,
Reitz, E.

Manuscript submitted for publication.

Author Contributions

EM, MD, MvL, and ER conceptualized the study. EM coordinated the data collection, analyzed the data and wrote the manuscript. All authors provided feedback on the study.

Acknowledgments

This study is financially supported by a grant from the Netherlands Organization for Health Research and Development (ZonMw Grant No. 531001106).

5

The Role of Classmates' Modeling and Reinforcement in a Universal School-Based Intervention Addressing Positive Peer Relations

Abstract

Experiences with classmates can affect adolescents' academic, emotional, and social development. We examined whether classmates' modeling and reinforcement were mechanisms of change in an intervention addressing the classroom peer context. Additionally, we examined whether relations between classmates' modeling and reinforcement and the perceived classroom peer context were moderated by dyadic mutuality. Questionnaires and observations were used in a sample of 7th Grade students ($N = 160$; $M_{age} = 12.37$; 53.8% boys). Classmates' modeling and reinforcement were not mechanisms of change. An increase in prosocial modeling, though, was related to a decrease in victimization, especially for dyads with high levels of mutuality. The results suggest that interventions should specifically focus on improving prosocial modeling and dyadic mutuality between classmates.

This trial is preregistered in the Dutch Trial Register, number NL6371 (NTR6554), registered on the 3rd of July 2017.

Keywords: peer influence; school-based intervention; perceived classroom peer context; dyadic mutuality; modeling; reinforcement

The Role of Classmates' Modeling and Reinforcement in a Universal School-Based Intervention Addressing Positive Peer Relations

The classroom is an important developmental context for students in which experiences with peers affect their academic, emotional, and social development (Rubin, Bukowski, & Parker, 2006). The classroom peer context is positive when students feel comfortable around their classmates, feel included in the group, experience few conflicts in the classroom, and experience no victimization by bullies (Boor-Klip, Segers, Hendrickx, & Cillessen, 2016). Not all students experience the classroom peer context as positive. For instance, in the Netherlands there are three educational tracks in the Dutch secondary school system of which more than half (54%; Central Bureau for Statistics, 2020) of the total student population follows the preparatory vocational education track. In this track, 20% of the students experience problems with their peers (Stevens & De Looze, 2018). Therefore, the aim of the present study was to examine the processes through which adolescents' experienced peer context in the classroom is influenced by their classmates.

Modeling and Reinforcement

According to the social learning theory (Bandura, 1977), two major ways through which peer influences occur are modeling and reinforcement. Peers' modeling and reinforcement can negatively and positively influence how students perceive the classroom peer context. Peers can model behavior that violates community or societal rules, i.e., deviant modeling, or positively evaluate such deviant behavior, i.e., deviant reinforcement (Piehler & Dishion, 2007). Both deviant modeling and reinforcement have consistently been linked to increases in adolescents' deviant behavior such as aggression, antisocial behavior (e.g., Dishion & Tipsord, 2011), bullying (e.g., Doehne, Von Grundherr, & Schäfer, 2018), and victimization (e.g., Ando, Asakura, & Simons-Morton, 2005). In contrast, peers can model behavior according to prosocial values, principles, and actions with the intention to benefit others, i.e., prosocial modeling, or respond positively to such prosocial behavior, i.e., prosocial reinforcement (Piehler & Dishion, 2007; Memmott-Elison, Holmgren, Padilla-Walker, & Hawkins, 2020). Prosocial modeling and reinforcement have been linked to decreases in problem behaviors,

for instance, antisocial behaviors (e.g., Hofmann & Müller, 2018), aggression, and depression (Memmott-Elison et al., 2020), but also to increases in adolescents' prosocial behavior (Busching & Krahé, 2020), positive interpersonal interactions in the class (e.g., Telzer, Van Hoorn, Rogers, & Do, 2018), and prosocial goal pursuit (e.g., Barry & Wentzel, 2006). Thus, increases in classmates' deviant modeling and reinforcement could negatively affect how students perceive the classroom peer context, whereas increases in classmates' prosocial modeling and reinforcement could positively affect students' perceptions of the classroom peer context.

The first aim of the current study was to examine whether the effect of an intervention on the perceived classroom peer context was mediated by changes in classmates' modeling and reinforcement. Based on the social learning theory and the empirical findings, it seems possible to positively stimulate students' perception of the classroom peer context by addressing classmates' modeling and reinforcement. An intervention that aims to improve the classroom peer context through changing classmates' modeling and reinforcement is Rock and Water (R&W; Ykema, 2002, 2018). R&W is a universal school-based intervention during which modeling and reinforcement are explicitly discussed and practiced. For instance, students model in role-play how to set boundaries in a non-aggressive way and discuss how reinforcing behaviors of peers can influence future behaviors. During the lessons, the trainer functions as a prosocial role model, reinforces students' prosocial behavior as well as students' reinforcement of prosocial behavior, and discourages and rejects students' deviant behavior. Communication is framed in the symbolic principles of 'rock', a firm position in which students are able to resist pressure from others, and 'water', a relaxed but alert position in which students are open to opinions, thoughts, and feelings of others. Students are encouraged to establish positive relations and have respect for each other. By setting negative consequences for deviant behaviors and by modeling and reinforcing alternative prosocial behaviors, R&W aims to enhance positive peer relations in the class. R&W has been found to be effective in improving resilience, positive identity formation, coping styles, self-regulation, and self-efficacy and in decreasing coercive strategies and verbal manipulation (De Graaf, De Haas, Zaagsma, & Wijssen, 2015; Ykema, Hartman, & Imms, 2006). We hypothesized that R&W would increase students' positive perceptions of the peer context in the classroom by decreasing students' deviant and increasing students' prosocial modeling and reinforcement.

Dyadic Mutuality

Our second aim was to examine whether the relation between changes in classmates' modeling or reinforcement and the perceived classroom peer context was moderated by dyadic mutuality. Dyadic mutuality indicates the degree of responsiveness, reciprocity and understanding shared between individuals. In dyads with high levels

of mutuality, peers listen and respond appropriately to each other, are genuinely interested in one another, and express affection towards each other (Piehler & Dishion, 2007). Although dyadic mutuality is related to positive (e.g., satisfaction, intimacy) and negative aspects (e.g., conflict, dissatisfaction) of friendship, it is not the same since dyadic mutuality can be low within some friendships. It is therefore important to not only focus on a specific type of relationship, but also on the quality of this relation (Berndt, 2002). Dyadic mutuality can be used to describe a variety of relations, such as friendships or parent-child relationships, and represents the variation in quality among these relations (Piehler & Dishion, 2007). In the current study, we used dyadic mutuality to describe the variation in quality of relations between classmates. Given that there are various relations between classmates (e.g., (un)reciprocal friendships, popularity, conflictual relations; Juvonen & Ho, 2008), dyadic mutuality is eminently suited to examine the extent to which adolescents are affect by peer influence in the class. Berndt (2002) theorizes that adolescents are more strongly influenced by peers with whom they have high quality relations, which is supported by empirical research (e.g., Barry & Wentzel, 2006; Piehler & Dishion, 2007). Therefore, we hypothesized that students' perception of the classroom peer context is more strongly influenced by classmates' modeling and reinforcement in dyads with higher levels of mutuality.

Current Study

In order to fully capture the complexity of the classroom peer context, the current study focused on characteristics of all four levels of the peer context as described by Hinde (1987) and Rubin and colleagues (2006). The *individual* level refers to individual characteristics students bring into the interaction (e.g., comfort). The *interaction* level refers to daily dyadic interactions in which students are interdependent; a student's behavior is both a response to and a stimulus for another student's behavior (e.g., conflict). The *relationship* level refers to a succession of interactions embedded in long-term knowledge of each other and is therefore influenced by meanings, expectations, and emotions of students towards each other (e.g., dyadic mutuality). The *group* level refers to the patterns and characteristics of interactions and relationships in a group of individuals (e.g., a class) who reciprocally influence each other. This process forms norms and shared cultural conventions that indicate which type of relations and interactions are acceptable (e.g., cohesion, victimization). These four levels are intertwined and interact together. Students' experiences with peers at one level have an influence on their experiences at the other levels.

The present study adds to the literature in three ways. First, we assessed classmates' influences and dyadic mutuality through observations within a randomized controlled trial of a universal intervention. Not only did the observations enable us to obtain more objective measures of modeling, reinforcement, and dyadic mutuality than through students' self-report, it also enabled us to examine change in students modeling and

reinforcement and its influence on the perceived classroom peer context. This is a unique approach as most studies examine these peer processes with questionnaires (e.g., Hofmann & Müller, 2018) or do not explicitly examine change in modeling and reinforcement within an intervention context (e.g., Dishion, Spracklen, Andrews, & Patterson, 1996). Second, we examined classmates' influences in not self-selected dyads to broaden the ecological validity of previous findings regarding peer influences. Focusing on involuntary, not self-selected peer groups such as classmates enabled us to emphasize relationship quality and examine peer influence beyond selection effects (Juvonen & Ho, 2008). In addition, by randomly assigning students to dyads we included various types of relations between classmates capturing the broad range of possible relations within a class. This is important as adolescents are exposed to behaviors of all their classmates and not only a selective group of classmates (e.g., friends, popular students; Busching & Krahé, 2020). Third, we examined classmates' deviant as well as prosocial modeling and reinforcement to unravel classmates' influences on students' perceptions of the classroom peer context, as recommended by Busching and Krahé (2020). By examining both deviant and prosocial modeling and reinforcement we could compare classmates' negative influences with classmates' positive influences as these might not be identical. Research examining classmates' influences has often focused on either deviant or prosocial influences (e.g., Hofmann & Müller, 2018; Juvonen & Ho, 2008) eliminating the possibility to directly compare deviant with prosocial influences.

Method

Data for the present study were collected as part of a larger study examining the effectiveness of R&W which is approved by the Ethical Committee of the Faculty of Social and Behavioral Sciences of Utrecht University (FETC17-015; see for protocol Mertens, Deković, Van Londen, & Reitz, 2018). The trial is registered in the Dutch Trial Register, number NL6371 (NTR6554).

Procedure

Schools that offered a preparatory vocational education track and had not implemented R&W in the last two years were able to participate. The Dutch secondary school system consists of three educational tracks: Preparatory vocational education track, preparatory college education track, and preparatory university education track. These tracks are attended by respectively, 54%, 23%, and 23% of the total student population (Central Bureau for Statistics, 2020).

Observation assessments were only conducted in a subsample of the larger project. In 6 schools 14 7th grade classes were randomly selected to participate in the observation task using an online random number generator (R&W condition: 7 classes; control condition: 7 classes). Using the same online generator, students were randomly matched in same-sex dyads within their class. Dyads were composed, as recommended by Huenecke and Waas (2010), of same-sex classmates as young adolescents affiliate mostly with classmates of the same sex.

Observations took place before the start and immediately after the intervention (about 4 months later) for the R&W condition and, parallel, at the same time points for the control condition. Additionally, students completed questionnaires at baseline and, about 4 months later, post intervention. Students gave active informed consent for their participation. Parents gave passive informed consent for participation of their child.

Participants

In total, 152 students (76 dyads) participated in the observation task at baseline. At post measurement, 130 students (65 dyads) participated again in the same dyad as at baseline. Eleven dyads were missing due to absence of one student (absent on the day of measurement $n = 6$, changed school $n = 4$, refused to participate $n = 1$). Missing data of the final sample ($N = 130$) was missing completely at random (Little's MCAR test: $\chi^2(18) = 18.94$, $p = .396$).

Students were between 11 and 14 years old ($M = 12.37$, $SD = .56$). Of these students, 70 (53.8%) were boys, and 85 (66.9%) had a Western background. The R&W condition consisted of 62 students of whom 34 (54.8%) were boys with an average age of 12.34 ($SD = .63$). Of these students, 30 (50.8%) had a Western background. The control condition consisted of 68 students of whom 36 (52.9%) were boys with an average age of 12.40 ($SD = .49$). Of these students, 55 (80.9%) had a Western background. An ANOVA and two Chi-squared tests showed no differences between the R&W and control condition concerning age and gender. The conditions differed slightly concerning students' background ($\chi^2(1) = 12.88$, $p < .001$, $\varphi = .318$); in the control condition more students had a Western background than in the R&W condition. There were no differences at baseline between the R&W and control condition concerning deviant and prosocial modeling and reinforcement, and outcome variables.

Conditions

Intervention. The theory of the intervention is based on the "R&W house". This house consists of 5 levels representing modules in which R&W aims to increase students' experienced safety, to learn students to deal appropriately with difficult situations, to teach about (non)verbal communication, help students to develop their

own preferences and choices, and to increase self-insight. This theoretical house is built on the three pillars of self-control, self-reflection, and self-esteem. According to the theory, strengthening students' skills concerning these pillars enables students to develop themselves within the broader domains of the R&W house (see for more information about the intervention Mertens, Deković, Van Londen, & Reitz, 2018).

R&W lessons were provided by teachers, mostly physical education teachers, who have followed the 3-day training course to become a certified R&W trainer. The rest of the teaching staff at the school received a 3-day training course to learn how they can support the R&W trainers and how to implement R&W during their regular classes.

Students received 14 weekly lessons of 90 minutes during a four months period. The lessons were provided during physical education due to the physical nature of R&W as it combines a physical approach with a psychological approach, i.e., a psychophysical approach. That is, students learn through play and exercises how to make (physical) contact with others, and explore, respect, and set own and other's boundaries. Each lesson is described in the manual and includes physical exercises, reflection, a moment of sharing thoughts with each other, and an exercise to strengthen the transfer of the learned skills to students' daily life.

Control. Students from the three schools in the control condition received care as usual. In one school this entailed a mentor students can go to with problems, a project week about 'being different' and an anti-bullying protocol. Another school had an anti-bullying coordinator and assigned a personal coach to each student with whom the student had regular meetings for advice and discussing the student's wellbeing. The third school also had an anti-bullying coordinator and a mentor students could to go to.

Measurements

Perceived classroom peer context. Levels of comfort, cohesion, and conflict in the class were measured with three subscales of the Classroom Peer Context questionnaire (Boor-Klip et al., 2016). The subscale Comfort assesses the level to which students feel at ease around their classmates (e.g., "In this class I can be myself."), Cohesion assesses unity and inclusiveness among classmates (e.g., "In this class children like each other."), and Conflict assesses students' negative social exchanges in the classroom (e.g., "In this class children fight with each other."). Each subscale contained 4 items answered on a 5-point Likert-type scale (1 = *totally not true*, 5 = *completely true*). Cronbach's α was for Comfort .71 and .84 for T1 and T2, respectively, for Conflict .83 and .88, and for Cohesion .44 and .62.

Experienced victimization was assessed with 1 item of the global measures of the Olweus Bully/Victim Questionnaire (Solberg & Olweus, 2003): "How often have you been bullied at school in the past two months?". This item was preceded by a definition of bullying. Frequency was indicated on a 5-point scale (1 = *never*, 5 = *multiple times a week*).

Deviant and prosocial modeling and reinforcement. Observations of peer interactions took place at the students' school during school hours and were videotaped by trained research assistants. The research assistant explained the procedure of the observation task and kept track of time outside the observation room. The research assistant was not present in the observation room during the discussions to enable the dyad to talk freely.

The observation task was based on the Peer Interaction Task (e.g., Dishion et al., 1996). The interaction consisted of four vignettes which students each discussed for five minutes. The first vignette was planning an activity together as a warm-up. The other three vignettes, systematically counterbalanced, concerned daily school situations involving: Student at work in the class, student with new clothes, and sitting together with classmates. For example, "classmate A is in the classroom working on an assignment in his book. Classmate B is doing nothing. Classmate B is annoying and throws pieces of paper towards classmate A." Two different versions of all vignettes were used for the baseline and post measurements. Participants were instructed to read the vignettes in turn aloud and discuss the situation together for five minutes. After five minutes, the research assistant re-entered the observation room to end the discussion and provided the next vignette. Students were also given three questions they could use in order to help them discuss the situation for the full five minutes: 1) What do you think of the situation? Could this happen at your school? 2) Imagine you are classmate A. What would you do? 3) How could this end?

The Conversation topic code (Piehler & Dishion, 2004b) was used to assess frequencies of deviant and prosocial verbal and nonverbal modeling. Verbal modeling was coded based on verbatim transcription of the discussion. Deviant modeling was all utterances that violated community or societal rules or were not appropriate to the setting or task (e.g., "I would hit him in his face."). Prosocial modeling was all utterances referring to positive or prosocial values, principles, or actions (e.g., "I wouldn't bully him."). Neutral modeling was all utterances that did not fit in the deviant or prosocial categories (e.g., "This situation happens all the time.").

Nonverbal modeling was coded, while watching the videotaped observation, when participants used gestures to support their utterance or only used gestures. Depending on the content of the gesture it was coded as deviant (e.g., making a punch movement, making weird faces) or prosocial (e.g., waving their hand as a greeting).

Proportions of (verbal and nonverbal) deviant and prosocial modeling were calculated over all verbal (i.e., prosocial, deviant, and neutral utterances) and nonverbal behaviors during the interaction, representing the proportion of prosocial and deviant modeling relative to all coded modeling. Interrater reliability of the three independent coders was good concerning deviant and prosocial modeling ($ICC_{deviant}$ = .96, $ICC_{prosocial}$ = .96) based on 22 observations coded over time.

In addition to deviant and prosocial modeling, verbal (e.g., "Indeed", "True", "No") and nonverbal (e.g., laughing, giving thumbs up, shaking head) reactions were coded (Piehler & Dision, 2007; Van de Bongardt et al., 2017). Reactions were coded

as reinforcement or as correction. Reactions were coded as reinforcement when the reaction indicated a positive evaluation of the other peer's behavior (e.g., "Indeed", laughing). Reactions were coded as deviant reinforcement when deviant behavior of the peer was reinforced and as prosocial reinforcement when prosocial behavior of the peer was reinforced. Corrections were coded when the reaction indicated a negative evaluation of the other peer's behavior (e.g., "No", shaking head).

Proportions of deviant and prosocial reinforcement were calculated relative to all coded reaction codes (i.e., reinforcements and corrections). Interrater reliability of the three independent coders was good concerning deviant and prosocial reinforcement ($ICC_{deviant}$ = .82, $ICC_{prosocial}$ = .87) based on 22 observations coded over time.

Dyadic mutuality. Dyadic mutuality was assessed at T1, based on the video-observations, by a combination of coding systems. Each member of the dyad was coded on three items: Responsiveness (i.e., the extent to which the student responded verbally and nonverbally to his or her peer), self-centeredness (i.e., the extent to which the student redirected the conversational flow to focus on personal ideas and experiences), and communicative efficiency (i.e., the appropriateness and competence of the messages send during the discussion; Piehler & Dishion, 2004a; Whalen, Henker, Collins, McAuliffe, & Vaux, 1979).

Additionally, each dyad as a whole was coded on three items: Reciprocity (i.e., verbal reciprocity such as engaging in a conversation-like interaction, and behavioral reciprocity such as eye-contact and posture orientation), shared attitudes and values (i.e., similar beliefs and attitudes about the discussed ideas) and affective valence (i.e., the emotional tone of the discussion and nonverbal behavior such as gestures, facial expression, and tone of voice (Piehler & Dishion, 2004a).

All items were rated on a 6-point Likert type scale (1 = *rarely or never*, 6 = *always or throughout*) and coded for the session as a whole. The item self-centeredness was reversed coded, so high values representing low self-centeredness. Subsequently, the 9 items (i.e., two times three individual items and three items of the dyad) were averaged to form a score on dyadic mutuality per dyad. Interrater reliability of the three independent coders was good (ICC = .73) based on 22 observations coded over time.

Analyses

First, we tested whether changes in students' modeling and reinforcement mediated the relation between R&W and students' perceived classroom peer context in M*plus* 8.2. We modeled multilevel mediation models which allowed us to analyze changes in modeling, reinforcement, and perceived classroom peer context at classroom level. In other words, we examined to what extent changes in modeling and reinforcement in the classroom mediated the relation between condition and the classroom peer context (Preacher, Zyphur, & Zhang, 2010). The mediators deviant and prosocial modeling were analyzed as parallel mediators in one model per outcome measure.

Likewise, the mediators deviant and prosocial reinforcement were analyzed as parallel mediators in one model per outcome measure. Baseline measures of the concerned mediators and outcome were added as covariates. Additionally, since the conditions differed significantly on ethnic background, this variable was added as a covariate. If ethnic background was a significant covariate it was retained in the model, otherwise it was dropped in favor of a more parsimonious model.

Second, we tested whether the relation between changes in modeling and reinforcement and students' perceived classroom peer context was moderated by dyadic mutuality, using multilevel analyses in M*plus* 8.2. At level 1, the individual level, within-dyad variation between modeling or reinforcement and the outcome was modeled. This relation was allowed to vary between individuals using a random slope. At level 2, the dyad level, dyadic mutuality was examined as predictor of the variation in the mean slope of modeling or reinforcement and the outcome. Deviant and prosocial modeling were analyzed in parallel in one model per outcome, as were deviant and prosocial reinforcement. Condition was added at level 2 as a predictor of the outcome to control for change in students' perceived classroom peer context explained by the condition in which the students participated. Baseline measures of modeling/reinforcement and the concerned outcome were added as covariates. Due to the estimation of cross-level interactions, no standardized fit indices were available. There is significant moderation when the slope between modeling or reinforcement and the outcome is dependent on the level of dyadic mutuality. In case of a significant moderation, the differing relations for dyads with low, average, and high levels ($M \pm 1$ *SD*) of dyadic mutuality between the concerning independent variable and outcome were graphically displayed.

Results

Preliminary analyses

We examined group differences post intervention on the perceived classroom peer context, modeling, and reinforcement using ANCOVAs, controlling for ethnicity and the corresponding baseline measure (see Table 1). No significant differences between the conditions were found. Nevertheless, we examined mediation by modeling and reinforcement as the absence of an intervention effect (i.e., a direct effect) does not exclude the presence of an indirect effect (O'Rourke & MacKinnon, 2018). Correlations between the variables are reported in the supplementary materials.

Table 1 Means and Standard Deviations of Students' Perceived Classroom Peer Context, Modeling, Reinforcement, and Dyadic Mutuality per Condition and Group Comparison at Post-test

	R&W		**Control**		**Post-test differences (ANCOVA)**		
	Baseline *M* (*SD*)	Post-test *M* (*SD*)	Baseline *M* (*SD*)	Post-test *M* (*SD*)	*F*	*p*	$\eta^2_{partial}$
Outcomes							
Comfort	4.35 (.80)	4.19 (1.04)	4.57 (.56)	4.36 (.76)	.23	.633	.002
Cohesion	4.27 (.72)	4.05 (.95)	4.36 (.65)	4.23 (.69)	.04	.850	.000
Conflict	2.58 (1.15)	2.43 (1.24)	2.31 (1.00)	2.39 (1.02)	.44	.508	.004
Victimization	1.29 (.67)	1.32 (.86)	1.18 (.60)	1.14 (.43)	1.32	.253	.011
Mediators							
Deviant modeling	.27 (.17)	.34 (.21)	.25 (.14)	.33 (.15)	.11	.742	.001
Prosocial modeling	.17 (.11)	.15 (.11)	.16 (.09)	.14 (.08)	.67	.417	.005
Deviant reinforcement	.30 (.28)	.38 (.33)	.29 (.21)	.43 (.27)	.91	.343	.007
Prosocial reinforcement	.46 (.34)	.39 (.34)	.45 (.28)	.33 (.28)	1.79	.183	.014
Moderator							
Dyadic mutuality	4.27 (.57)		4.51 (.53)				

Mediation analyses

Both changes in modeling and in reinforcement did not mediate the effect of the intervention on students' perceived classroom peer context (see Table 2). More specifically, the intervention did not predict changes in modeling or reinforcement (Path a). Changes in modeling or reinforcement also did not predict the outcomes, except concerning victimization (Path b). An increase in prosocial modeling was related to a decrease in experienced victimization.

Moderation analyses

Dyadic mutuality moderated the relation between changes in prosocial modeling and experienced victimization (see Table 3). The negative relation between changes in prosocial modeling and experienced victimization was stronger for dyads with higher levels of mutuality (B_{Low} = -0.79; $B_{Average}$ = -1.70; B_{High} = -2.60; see Figure 1). No other moderations by dyadic mutuality were found on either modeling or reinforcement.

Table 2 Mediation Effects Modeling and Reinforcement at Classroom Level

	Modeling						**Reinforcement**					
	Path a		**Path b**		**Indirect effect**		**Path a**		**Path b**		**Indirect effect**	
	B (SE)	95% CI	*B (SE)*	95% CI	Effect (SE)	95% CI	*B (SE)*	95% CI	*B (SE)*	95% CI	Effect (SE)	95% CI
Model 1 Comfort												
Deviant	-.00 (.03)	-.06; .06	.06 (.36)	-.64; .76	.00 (.00)	-.00; .00	-.06 (.07)	-.19; .07	-.26 (.43)	-1.11; .59	.02 (.03)	-.04; .07
Prosocial	.02 (.01)	-.00; .04	.41 (.92)	-1.40; 2.21	.01 (.02)	-.02; .04	.07 (.07)	-.07; .21	-.10 (.31)	-.71; .52	-.01 (.02)	-.05; .04
Model 2 Cohesion												
Deviant	-.00 (.03)	-.06; .06	-.31 (.41)	-1.11; .48	.00 (.01)	-.02; .02	-.06 (.06)	-.17; .05	-.21 (.34)	-.88; .45	.01 (.02)	-.03; .06
Prosocial	.01 (.01)	-.01; .03	.39 (.76)	-1.10; 1.88	.01 (.01)	-.02; .03	.06 (.05)	-.04; .16	.00 (.36)	-.71; .72	.00 (.02)	-.04; .04
Model 3 Conflict												
Deviant	.00 (.03)	-.06; .06	-.02 (.50)	-.99; .95	.00 (.00)	-.00; .00	-.07 (.06)	-.17; .04	.33 (.44)	-.53; 1.18	-.02 (.03)	-.08; .04
Prosocial	.02 (.01)	-.01; .04	-.51 (.91)	-2.30; 1.28	-.01 (.01)	-.04; .02	.06 (.05)	-.03; .16	.12 (.29)	-.45; .69	.01 (.02)	-.03; .04
Model 4 Victimization												
Deviant	-.00 (.03)	-.06; .06	-.15 (.48)	-1.09; .80	.00 (.00)	-.01; .01	-.05 (.08)	-.20; .11	.21 (.46)	-.69; 1.10	-.01 (.02)	-.05; .03
Prosocial	.01 (.01)	-.01; .04	-1.86 (.70)**	-3.23; -.49	-.03 (.02)	-.07; .02	.06 (.36)	-.90; 7.75	-.27 (.25)	-.76; .21	-.02 (.10)	-.22; .19

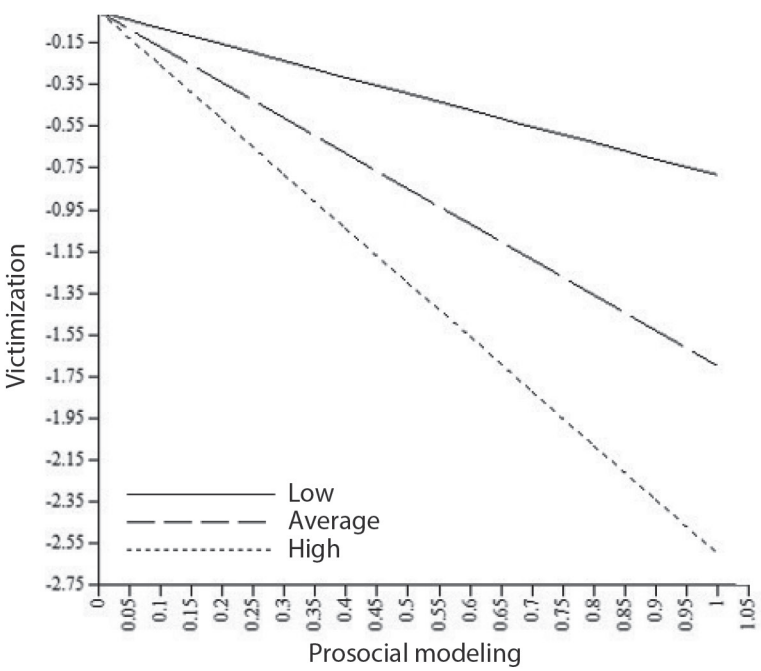
Note. Path a: Condition ⟶ Modeling/Reinforcement; Path b: Modeling/reinforcement ⟶ Outcome; * $p < .05$ ** $p < .01$.

Table 3 Moderation of Dyadic Mutuality of the Relation Between Modeling or Reinforcement and Perceived Classroom Peer Context at the Dyad Level

	Modeling		**Reinforcement**	
	B (*SE*)	95% CI	*B* (*SE*)	95% CI
Model 1 Comfort				
Deviant x mutuality	1.28 (2.48)	-3.58; 6.14	1.19 (.61)	-.01; 2.39
Prosocial x mutuality	2.80 (3.71)	-4.48; 10.07	-.31 (.57)	-1.43; .81
Model 2 Cohesion				
Deviant x mutuality	1.26 (1.15)	-.99; 3.50	.84 (.53)	-.19; 1.87
Prosocial x mutuality	3.22 (2.04)	-.79; 7.22	-.25 (.53)	-1.28; .79
Model 3 Conflict				
Deviant x mutuality	-2.43 (1.67)	-5.69; .84	-.03 (.61)	-1.23; 1.16
Prosocial x mutuality	-2.37 (1.55)	-5.40; .66	-.01 (.56)	-1.10; 1.08
Model 4 Victimization				
Deviant x mutuality	.06 (.54)	-.99; 1.12	.24 (.30)	-.35; .84
Prosocial x mutuality	-1.61 (.57)**	-2.73; -.48	-.02 (.20)	-.41; .37

Note. * $p < .05$ ** $p < .01$.

Figure 1. Moderation of dyadic mutuality of the relation between prosocial modeling and victimization.



Discussion

The present study extended previous research on peer influence by examining classmates' influences through observations within a randomized controlled trial of a universal intervention study, examining influence in not self-selected dyads, and focusing on both deviant and prosocial modeling and reinforcement. Contrary to our expectations, changes in classmates' modeling and reinforcement were not mechanisms of change in the universal school-based intervention R&W, that aims to improve positive peer relations. Overall, regardless of condition, we did find that increases in classmates' prosocial modeling was related to less victimization in the class, especially when the dyadic mutuality between the classmates was high.

The finding that classmates' modeling and reinforcement were not mechanisms of change in R&W might indicate that modeling and reinforcement were not intensely enough addressed during the intervention. More intensive attention to setting negative consequences for deviant behaviors and reinforcing prosocial behaviors in the classroom has been related to more positive behavior of students (Phillips Smith, Dumas, & Prinz, 2006). During the R&W lessons the trainer sets negative consequences for deviant behavior, reinforces prosocial behavior, and serves as a role model, as described in the intervention manual (Ykema, 2002). However, the intervention manual lacks guidelines for applying these techniques outside the intervention lessons. Hence, this classroom management approach may not have been implemented during regular lessons or implemented differently among teachers. This decreases the intensity and consistency with which modeling and reinforcement are addressed, limiting the opportunities for students to learn from prosocial models and refrain from deviant models (Phillips Smith et al., 2006).

Another explanation might be that the time frame of R&W is too short. Sijtsema and Lindenberg (2018) indicated in their review that studies showing a relation between peer influence and antisocial behavior used a time span from six months to one year, whereas the R&W intervention takes approximately four months. A longer time frame might be necessary to establish change in classmates' modeling and reinforcement and to have an influence on the classroom peer context as group dynamics take time to change.

The finding that an increase in prosocial modeling is related to a decrease in victimization is particular eminent for anti-bullying interventions and in line with the social learning theory (Bandura, 1977). Classmates showing prosocial modeling are more inclined to show affiliation and involve all classmates in classroom activities degrading the level of exclusion and rejection (Juvonen & Ho, 2008). Other students might imitate this prosocial behavior which results in less victimization. In contrast to Ando and colleagues (2005), we did not find a relation between deviant peer influences and victimization. Given that Ando and colleagues (2005) studied peer influences of friends and the present study influences of classmates, the difference in findings might

indicate that the processes through which friends influence each other differ from the processes through which classmates influence each other. Thus, for diminishing victimization in the classroom specifically, it seems especially important to focus on improving prosocial behaviors instead of reducing deviant behaviors, which is in line with suggestions of Busching and Krahé (2020).

The relation between prosocial modeling and victimization appears to be stronger when dyadic mutuality levels between classmates are higher. In interactions in which classmates are interested in each other, show affection, and are responsive, students appear to be more influenced by the prosocial behavior of their classmates and might be more likely to imitate this behavior, which is in line with previous research (e.g., Barry & Wentzel, 2006). Hence, in addition to stimulating prosocial behaviors in the class, attention should be given to improving classmates' mutuality in order to strengthen the positive effect of prosocial modeling on victimization. For instance, interventions could provide positive and fun exercises in the class in which classmates who do not interact on a daily basis work together. This might improve students' emotions toward each other and their expectations for future interactions (i.e., the relationship level of the classroom context) resulting in more positive mutual feelings and affection between students.

We did not find relations between deviant and prosocial modeling and reinforcement and interpersonal relations in the class (i.e. perceived levels of comfort, cohesion, and conflict). Perhaps, modeling and reinforcement in the classroom reflect the norms and shared cultural conventions in the class due to which these mechanisms exert their influence mainly on the group level – the level on which victimization is represented. Another possible explanation is that victimization might more strongly represent students' perceptions of the school context as dangerous (Goldstein et al., 2008), whereas interpersonal relations might more strongly represent students' feelings of social support in the classroom (Hopson et al., 2014). While perceptions of the school context appear to be influenced by experiencing and witnessing the behaviors of all classmates (Goldstein et al., 2008), perceived social support appears to be mostly influenced by friends (Bokhorst, Sumter, & Westenberg, 2010). Hence, maybe only modeling and reinforcement of friends in the class have an influence on perceived interpersonal relations in the class, regardless of the relationship quality.

The absence of a relation between modeling and reinforcement and interpersonal relations in the class could also have a methodological explanation. The questionnaire regarding interpersonal relations in the class consisted of items referring to the class and classmates in general, whereas the question regarding experienced victimization concerned the students themselves. For instance, when students indicated that there were conflicts in the class, they were not necessarily involved in these conflicts. In contrast, when students indicated experienced victimization by bullies they were victimized themselves. Even though the used questionnaire gives a general overview of the interpersonal relations in the class, it does not indicate to what extent students themselves are affected by the interactions and relations between classmates. Thus,

future research examining interpersonal relations in the class should add questions asking to which extent students are affected by the behaviors of other classmates.

Limitations and future directions

When considering the findings of our study, it is important to note some strengths and limitations. A strength of the study is the use of observations to assess classmates' influences and mutuality. Using observations we were able to directly code modeling, reinforcement, and dyadic mutuality without depending on subjective perspectives of students. Furthermore, we examined both deviant and prosocial peer influences in not self-selected dyads. This allowed us to examine negative as well as positive peer influences with reduced selection effects. Moreover, our study had an experimental design with a pre- and post-measurement enabling us to examine changes in modeling, reinforcement, and the perceived classroom peer context.

A limitation of the present study is the somewhat small sample size. Even though our sample size is rather large for an observation study, it might be that due to the relatively small sample size some relations failed to reach significance. Additionally, we did not examine classroom characteristics as predictors. Due to the limited number of clusters at classroom level the models were kept as simple as possible. However, classmates' influences might depend on characteristics of the general classroom context such as class size or gender composition. Future research could focus on classroom characteristics and examine whether these characteristics influence students' perceived classroom peer context. Moreover, the mediators and outcomes were measured at the same time point (i.e., post intervention). We have analyzed these time points because we expected changes in modeling and reinforcement to mediate intervention effects immediately after the intervention rather than intervention effects between post and follow-up measurements (Beauchaine & Slep, 2018). However, this approach limited the extent to infer causal order (Weeland et al., 2018). Furthermore, we measured victimization with only one item. Even though it is common in research concerning bullying to measure (types of) victimization with one item, it might be more reliable to use multiple items.

Conclusion

The present study showed that increases in prosocial modeling of classmates were related to decreases in students' experienced victimization, especially when dyadic mutuality between classmates was high. However, changes in deviant and prosocial modeling and reinforcement did not mediate the effect of R&W on the perceived classroom peer context. Whereas prosocial modeling was related to victimization,

deviant modeling had no effect on students' perceptions of the classroom peer context. Interventions that aim to improve how students' perceive the peer context in the class, and victimization in particular, should therefore focus on students' prosocial modeling. Altogether, this study underscores the importance for future research to focus equally on deviant and prosocial peer influences within different sorts of relationships as deviant and prosocial influences are not mere ends of the same continuum. Additionally, our study highlights the importance for interventions aiming to stimulate the classroom peer context to take classmate influences into account.

Supplementary material

Table S1 Correlations Between Variables Assessed at Baseline and Post Intervention

	1.	2.	3.	4.	5.	6.	7.	8.	9.	10.	11.	12.	13.	14.	15.	16.
Classroom peer context T1																
1. Comfort	-															
2. Cohesion	.45**	-														
3. Conflict	-.23**	-.51**	-													
4. Victimization	.15	-.42**	-.15	-												
Mediators T1																
5. Deviant modeling	.26**	-.15	-.27**	-.03	-											
6. Prosocial modeling	-.21*	.13	.22*	-.14	-.46**	-										
7. Deviant reinforcement	.13	-.06	-.08	-.01	.36**	-.26**	-									
8. Prosocial reinforcement	-.23**	.03	.17	.02	-.49**	.36**	-.64**	-								
Moderators T1																
9. Dyadic mutuality	-.22*	.00	.07	-.12	.04	.02	.05	.02	-							
Classroom peer context T2																
10. Comfort	-.37**	.30**	.33**	-.10	-.15	.20*	-.15	.14	-.16	-.25**	-					
11. Cohesion	-.38**	.27**	.37**	-.09	-.14	.28**	-.18	.10	.15	-.36**	73**	-				
12. Conflict	-.23**	-.07	.30**	.11	-.02	-.05	-.17	.09	.14	-.25**	-.36**					
13. Victimization	.01	-.15	-.11	.09	.19*	-.18*	-.02	-.23**	.08	.25**	-.12	-.07	-			
Mediators T2																
14. Deviant modeling	.27**	-.03	-.14	.03	.48**	-.40**	.15	-.25**	-.24**	.14	-.11	-.18	.14	-		
15. Prosocial modeling	-.12	.19*	.16	-.14	-.36**	.51**	-.21*	.26**	.03	-.09	.19*	.18	-.26**	-.54**	-	
16. Deviant reinforcement	.10	.08	-.05	-.08	.13	-.15	.06	-.15	-.07	.14	-.05	-.11	.13	.39**	-.20*	-
17. Prosocial reinforcement	-.15	.02	.08	.03	-.27**	.20*	-.14	.28**	.10	-.15	.11	.13	-.22*	-.43**	.18*	-.61**

Note. * $p < .05$; ** $p < .01$.

Mertens, E. C. A.,
Deković, M.,
Leijten, P.,
Van Londen,
Reitz, E.

Manuscript submitted for publication.

Author Contributions

EM, MD, MvL, and ER conceptualized the study. EM coordinated the data collection, analyzed the data and wrote the manuscript. All authors provided feedback on the study.

Acknowledgments

This study is financially supported by a grant from the Netherlands Organization for Health Research and Development (ZonMw Grant No. 531001106).

6

What to Do or Not to Do to Stimulate Students' Intrapersonal and Interpersonal Domains: Meta-Analysis of School-Based Intervention Components

Abstract

Many universal school-based interventions aim to stimulate students' intrapersonal (e.g., self-esteem) and interpersonal (e.g., school climate) domains. To improve our understanding of why some of these interventions yield stronger effects than others, we identified intervention components that are related to stronger or weaker intervention effects. We systematically searched four databases (i.e., PsycINFO, PubMed, ERIC, CENTRAL) for controlled evaluations of universal school-based interventions. In total, 104 included studies (529 effect sizes) reported on 99 unique interventions. Interventions showed small positive effects in the intrapersonal ($d = .19$) and interpersonal ($d = .15$) domains. Focusing on self-awareness and problem solving, using more active learning approaches, and using more extensive interventions predicted stronger intervention effects on aspects of both domains. In contrast, efforts to improve emotion regulation, assertiveness, cognitive coping, and using group discussions predicted weaker intervention effects. Our finding that commonly implemented components did not necessarily predict stronger intervention effects suggests the need to carefully select components for inclusion in interventions.

PROSPERO registration number: CRD42019137981

Keywords: Components, School-based intervention, Intrapersonal domain, Interpersonal domain, Students

What to Do or Not Do to Stimulate Students' Intrapersonal and Interpersonal Domains: Meta-Analysis of School-Based Intervention Components

Schools are expected to foster not only their students' cognitive development, but also their students' wellbeing. Schools should implement policies and practices striving to improve students' attitudes, values, and social support (Langford et al., 2014; World Health Organization, 1995). To this end, a range of universal school-based interventions have been developed to enhance students' intrapersonal and interpersonal domains as both domains are markers of positive development, psychosocial wellbeing, and preparedness for future social interaction (e.g., within occupational, romantic, and family domains; Barber, 2005; Shek & Leung, 2016). The intrapersonal domain refers to managing one's own feelings, emotions, and attitudes pertained to the individual self (Barber, 2005). The interpersonal domain refers to the ability to build and maintain positive relationships with others and to understand social situations, roles, and norms, and respond appropriately (Pellegrino & Hilton, 2012; Shek & Leung, 2016). Even though the two domains bidirectionally influence each other, they are considered as distinct domains. While the intrapersonal domain represents an individual's subjective psychological functioning, the interpersonal domain represents an individual's social functioning (Dufner, Gebauer, Sedikides, & Denissen, 2019). Students can acquire competencies in both domains by mastering relevant cognitive, affective, and social skills, such as the ability to identify emotions (intrapersonal domain) and perspective taking (interpersonal domain; Durlak, Weissberg, Dymnicki, & Taylor, 2011). They are at increased risk of developing problems in the intrapersonal domain, such as internalizing behavior, or in the interpersonal domain, such as aggression, when mastery of (some of) these skills lacks or falls behind (Modecki, Zimmer-Gembeck, & Guerrà, 2017; White, Jarrett, & Ollendick, 2013).

Children's intra- and interpersonal domains develop throughout their youth, but the importance of these skills becomes particularly pronounced in adolescence when adolescents consolidate their own identity and peer relationships become increasingly important. They spend less time at home and longer hours at school which provides them increasing opportunities and requirements to interact with others, such as peers, teachers, and romantic partners (Barber, 2005). This makes secondary school a potentially good target for interventions to foster the intra- and interpersonal domains.

In the present meta-analysis, we therefore examined the effects of universal secondary school-based interventions on students' intrapersonal and interpersonal domains.

School-based interventions addressing adolescents' intra- and interpersonal domains typically show small positive effects (e.g., effect sizes (Cohen's *d*) ranging from .03 to .24; Dray et al., 2017; Durlak et al., 2011; Jiménez-Barbero, Ruiz-Hernández, Llor-Zaragoza, Pérez-García, & Llor-Esteban, 2016). One way to increase the effectiveness of interventions is by studying which components are related to intervention effects. By identifying components associated with stronger or weaker intervention effects, existing and new interventions can be optimized. In addition, schools can make informed decisions about which intervention to implement, by selecting interventions based on the evidence base for the components. As a first step towards cataloguing potentially effective components, Boustani and colleagues (2015) provided an overview of the components that are most frequently included in effective school-based interventions (e.g., problem solving, psychoeducation). Although such a frequency count provides a useful overview, it does not show whether the effectiveness of interventions are actually related to the presence of the components. Furthermore, due to the focus on effective interventions, the overview does not identify potentially ineffective or iatrogenic components. Therefore, the present meta-analysis identified which components are related to stronger intervention effects and which components are related to weaker intervention effects.

In the literature, typically three types of components are distinguished: Content, instructional, and structural components. *Content components* are specific skills adolescents learn to promote positive outcomes, such as emotion regulation and problem solving (Boustani et al., 2015), i.e., "what do they learn." *Instructional components* are techniques and methods of information delivery used by the intervention facilitator, such as cognitive restructuring and modeling (Boustani et al., 2015), i.e., "how do they learn it." *Structural components* describe the structure of the intervention that might impact results, such as the number of sessions and whether or not parents are included in the intervention (Lee et al., 2014), i.e., "how is the intervention set up." By examining all three types of components, we are able to unravel whether each type of components is equally important for intervention success or whether especially one type of components appears to be more important than others.

There are various meta-analyses of intervention components that predict intervention effects (e.g., De Vries, Hoeve, Assink, Stams, & Asscher, 2015; Kaminski et al., 2008; Van der Put, Assink, Gubbels, & Boekhout van Solinge, 2018), but few meta-analyses have focused on components of school-based interventions. Meta-analyses that did examine components of school-based interventions focused on substance use, sexual risk behaviors (e.g., pregnancy, STD/HIV) and/or nutrition (see for a review of reviews Peters, Kok, Ten Dam, Buijs, & Paulussen, 2009). For example, Onrust, Otten, Lammers, and Smit (2016), focusing on substance use, found that in middle and high school components that sought to stimulate students' self-control

and problem solving, and components that included cognitive restructuring, adjusting social norms (e.g., peer education), and parental involvement predicted stronger substance use reductions. Hennessy and Tanner-Smith (2015), focusing on alcohol use, found that in secondary school components that included an individual and motivational enhancement approach were more effective in reducing alcohol use.

In the present meta-analysis, we examined which components are related to more (or less) effective school-based interventions addressing students' intra- and interpersonal domains. We focused on a broad range of outcomes to unravel which components are important for improving students' overall development in the two general domains, and which components are important for improving specific competencies and problems in the two domains. We studied relations between components and intervention effects across different populations and circumstances. Although the effectiveness of components can be dependent on characteristics of participants and circumstance, our meta-analysis provides an initial overview of potentially relevant components in general. We analyzed all three types of components (i.e., content, instructional, and structural) and tested whether interventions with a specific component showed larger (or smaller) effect sizes than interventions without that component, using multilevel meta-regression. This enabled us to identify not only which components were associated with stronger effects, suggesting potential effective components, but also components associated with weaker effects, suggesting potential ineffective components. Knowing what does *not* work is equally important as knowing what does work (e.g., Poulin, Dishion, & Burraston, 2001; Werch & Owen, 2002).

Concerning content components, based on the results of Onrust and colleagues (2016) and Boustani and colleagues (2015), we hypothesized that basic life skills and self-awareness would be related to stronger intervention effects on students' intra- and interpersonal domains. Basic life skills refers to abilities for adaptive and positive behavior to deal with demands and challenges of everyday life (World Health Organization, 1997). Several reviews suggest the importance of basic life skills, such as problem solving, assertiveness, and social skills, for a range of outcomes of effective school-based interventions (e.g., intra- and interpersonal domains, Boustani et al., 2015; drug use, Cuijpers, 2002). Self-awareness indicates a realistic and accurate assessment of one's strengths and norms, and is related to improvements on the interpersonal domain (e.g., Shek & Leung, 2016). Raising self-awareness, such as insight building and self-efficacy, is often used in effective interventions targeting the intrapersonal and interpersonal domains (Boustani et al., 2015).

For instructional components, we hypothesized that components using a more active learning approach, in which students interact with each other and perform tasks (e.g., practicing through role-play), would be related to stronger intervention effects. Active learning approaches have consistently been related to stronger effects. For instance, Kaminski and colleagues (2008) found in their meta-analysis that parenting interventions in which parents practiced the learned skills were more effective than

interventions that did not include practice. Similarly, Cuijpers (2002) concluded in his review of school-based drug interventions that interventions using more active methods (e.g., discussion) were more effective than interventions using more passive methods (e.g., didactic instruction).

Regarding structural components, the general assumption is that longer and extensive interventions are more effective than briefer and less extensive interventions (Yeager & Walton, 2011). The evidence, however, is conflicted. Some meta-analyses showed that longer and extensive interventions are indeed more effective. For instance, interventions showed stronger effects as the time span, number of sessions, and involved persons (i.e., whole school, parents) increased (De Vries et al., 2015; Ttofi & Farrington, 2011). Other meta-analyses, on the other hand, showed that briefer and less extensive interventions are more effective (i.e., "less is more"). For instance, interventions showed stronger effects when the time span was short, the number of sessions limited, and no additional services were provided (Cuijpers, 2002; Kaminski et al., 2008; Van der Put et al., 2018). Longer and extensive interventions require more time and effort to implement, which potentially takes away time and energy from the main goal (Bakermans-Kranenburg, Van IJzendoorn, & Juffer, 2003). Therefore, we hypothesized that briefer and less extensive interventions would be related to stronger intervention effects on students' intra- and interpersonal domains than longer and extensive interventions.

In summary, identifying components related to stronger or weaker intervention effects has important theoretical and practical implications. First, it expands our knowledge concerning interventions. We begin to unravel what is more important to change students' intra- and interpersonal domains: What they learn, how they learn it, or how the intervention is set up? Second, it enables schools to make informed decisions about which intervention to implement and existing and new interventions can be optimized; the present meta-analysis not only examined what interventions "should do", but also what they "should not do." This knowledge is a first step towards improving the effectiveness of school-based interventions addressing students' intra- and interpersonal domains, ultimately enhancing students' positive development, psychosocial wellbeing, and preparedness for future challenges.

Method

Inclusion and Exclusion Criteria

We sought to include evaluations of universal secondary school-based interventions targeting students' intrapersonal and interpersonal domains. Universal secondary school-based interventions were defined as interventions delivered to students during

regular school hours, targeting all students (Mychailyszyn et al., 2012; Peters et al., 2009). The intrapersonal domain was defined as feelings, emotions, and attitudes about the self (Barber, 2005) consisting of competencies (e.g., resilience, self-esteem, self-regulation, general wellbeing) and problems (e.g., internalizing behavior). The interpersonal domain was defined as the ability of an individual to build and maintain positive relationships with others and understanding social situations, roles and norms (Shek & Leung, 2016) consisting of competencies (e.g., social competence, sexual health, positive school climate) and problems (e.g., aggression, bullying).

Studies were eligible for review when (1) the intervention was implemented in a regular school (i.e., not in special education), (2) the intervention was implemented during regular school hours in a group setting, (3) the intervention was aimed at improving (subdomains of) the intra- and/or interpersonal domain (i.e., interventions primarily aiming to improve students' physical health (e.g., prevention of substance use, nutrition, pregnancy, STDs) or prevent mental disorders (e.g., depression) were excluded.), (4) the intervention was universal, so targeting all students, (5) the participants were in middle school or high school (Grades 6 – 12), (6) the study included a control group, (7) the study included a quantitative baseline and post intervention measurement of (subdomains of) the intrapersonal domain and/or interpersonal domain, (8) sufficient information concerning baseline and post intervention measurements was reported, or obtained after contact with the author, so that effect sizes could be calculated post intervention, corrected for baseline differences, (9) the study was written in English, and (10) the study was published as article, book, or book chapter. Research has shown that including unpublished studies does not reduce the possible impact of publication bias and is sometimes even counterproductive due to selection bias (Ferguson & Brannick, 2012).

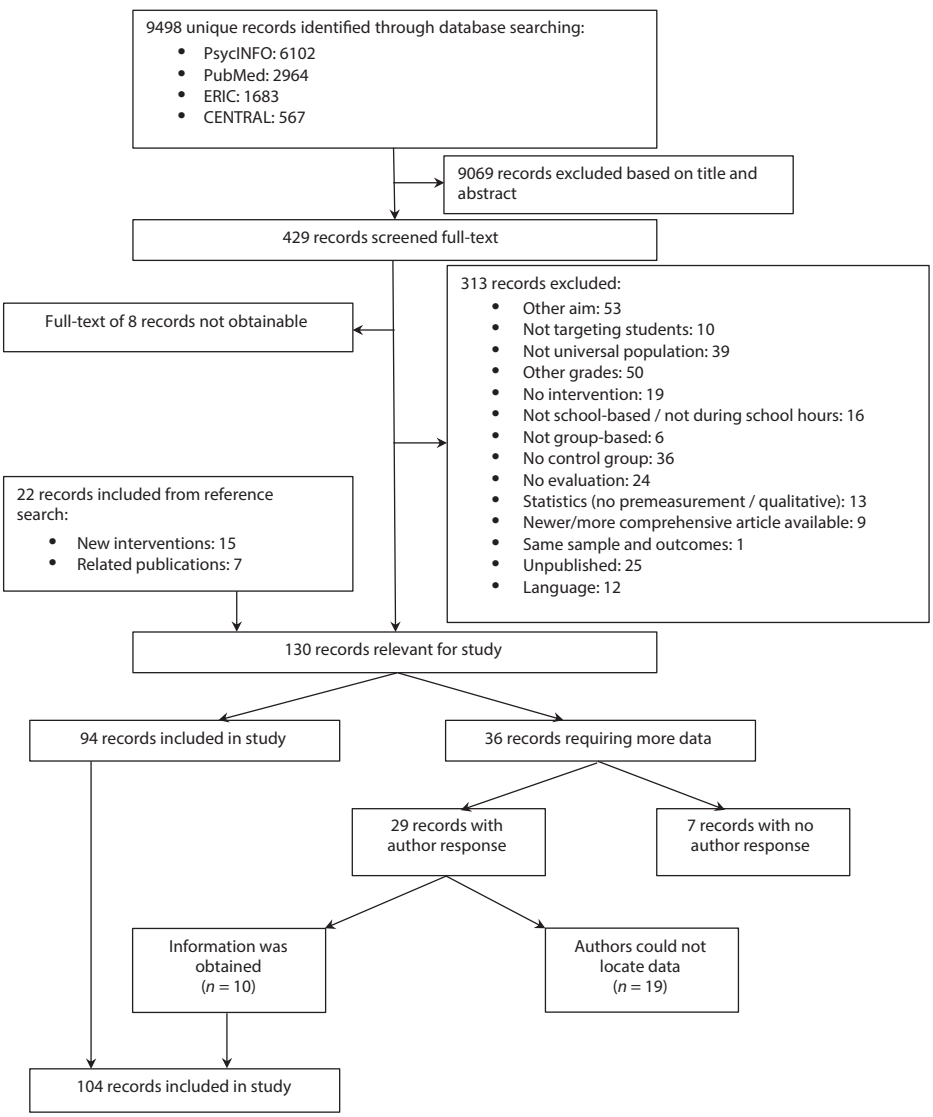
Literature Search

We searched four databases (i.e., PsycINFO, PubMed, ERIC, and CENTRAL). Search terms were used to elicit school-based interventions (e.g., school, class), interventions (e.g., prevention, intervention), adolescents (e.g., adolescent, youth), and intra- and interpersonal outcomes (e.g., self-esteem, social competence). Because these search terms led to an extremely high number of studies, we added some restrictions to the search, to avoid picking up interventions targeting other populations (e.g., preschool, clinical) or domains (e.g., substance use, lifestyle) than targeted in this study. This search (April 2019) resulted in 6,102 studies in PsycINFO, 2,964 studies in PubMed, 1,683 studies in ERIC, and 567 studies in CENTRAL. Removal of duplicates resulted in 9,498 unique studies. In addition, reference lists of included studies and identified relevant reviews and meta-analyses were searched. This resulted in 22 additional studies.

All studies identified by the search were first screened for eligibility based on their title and abstract. Based on this screening 9,068 studies (95%) were excluded.

Reliability of this first selection, based on 9% (800 studies) screened by two independent researchers, was substantial, with 98% agreement (Cohen's κ = .71; Landis & Koch, 1977). The remaining 429 studies were studied full-text. In this step another 310 studies (72%) were excluded. Reliability of this second selection, based on 10% (45 studies) screened by two independent researchers, was substantial, with 89% agreement (Cohen's κ = .76; Landis & Koch, 1977). Any disagreements between the researchers concerning inclusion were solved through discussion. See Figure 1 for the flow diagram.

Figure 1. Flow diagram.



Data Extraction

Studies were coded for information concerning the study (e.g., year of publication, country where study was conducted), sample (e.g., age, gender distribution), design and method (e.g., randomization, attrition analyses), intervention (e.g., intervention provider, aim of intervention), effect size data (e.g., outcome category), and intervention components (e.g., problem solving, practice, parental involvement). The intervention components were based on the reviews and meta-analyses by Boustani and colleagues (2015), Kaminski and colleagues (2008), Onrust and colleagues (2016), Peters and colleagues (2009) and Van der Put and colleagues (2018). An overview of all components and their definitions is presented in Appendix A. Sources cited in the study and other freely available materials, such as descriptions from the developer or websites, were retrieved for coding the components (Boustani et al., 2015; Kaminski et al., 2008). In cases where insufficient data were reported for calculating the effect size, the first author was contacted. When this author had not responded after a reminder, the second or last author was contacted and, if necessary, reminded. If the required data could not be obtained after this, the study was excluded from the meta-analysis (see Figure 1 for the flow diagram).

Of the included studies, 28% (30 studies) was coded independently by a second coder for reliability. The inter-rater reliability was moderate to excellent (Landis & Koch, 1977) with an average intra-class-correlation of .97 (*SD* = .05), ranging from .88 to 1.00, for continuous variables, and an average Cohen's kappa of .82 (*SD* = .11), ranging from .60 to 1.00, for categorical variables. Coding of the component 'Insight building' was not reliable with Cohen's kappa of .52. Disagreements between the two coders were discussed and solved unanimously.

Calculation and Analyses of Effect Sizes

Effect sizes were represented as Cohen's *d*, reflecting the standardized mean difference between the intervention and control condition, following the procedures of Lipsey and Wilson (2001). Effect sizes were calculated at post intervention (i.e., within 6 months after the intervention) and corrected for baseline differences. Positive effect sizes indicated better results for the intervention compared to the control condition. All effect sizes were adjusted using the Hedges' (1983) small sample correction prior to analyses. Outliers were examined and, when believed to be unrepresentative, winsorized by replacing outliers with the value of the lower or upper value of two standard deviants from the mean (Lipsey & Wilson, 2001).

Publication Bias

As commonly known, studies with nonsignificant or negative results are less likely to be published than studies with significant or positive results. The risk of publication bias was tested using a funnel plot. When the funnel plot was asymmetrical according to Egger's regression test (Egger, Schmidt, Schneider, & Minder, 1997) the trim-and-fill analysis (Duval & Tweedie, 2000a; 2000b) was used to adjust the effect for possible publication bias. This analysis estimates how many studies fall outside the symmetric part of the funnel plot and trims this outlying part. With the remaining symmetric funnel plot the true center of the funnel is estimated. The trimmed studies and their missing counterparts are replaced in the funnel representing imputed 'missing' effect sizes. Based on this filled funnel plot, the corrected mean is estimated resulting in an adjusted effect size. Tests to visualize and examine publication bias assume independence of effect sizes, which is not the case in multilevel meta-analyses. We took this violation into account by using the variance of the effect sizes as a moderator in Egger's regression test.

Analyses

We calculated an effect size for each reported measure of the intra- or interpersonal domain. To account for the clustering of effect sizes within a trial, we used multilevel meta-analytical models with three levels: Sampling variance around each effect size (level 1), variance between effect sizes within studies (level 2), and variance between studies (level 3; Assink & Wibbelink, 2016; Van den Noortgate, López-López, Marín-Martínez, & Sánchez-Meca, 2013).

The unit of analyses were the interventions rather than the publications, since we are interested in the effectiveness of the intervention compared to the control condition. When one publication reported on two interventions, both interventions were included and analyzed separately. When multiple publications reported on the same intervention, evaluated in different studies with different samples, their effect sizes were analyzed together, clustered within the same intervention. When multiple publications reported on the same intervention, evaluated in the same study with the same sample, we coded the most comprehensive publication; the less comprehensive publication was checked for additional information and their effect sizes were analyzed together, clustered within the same intervention.

The multilevel analyses were conducted in R using the metaphor package (Viechtbauer, 2010). First, the overall effects of universal school-based interventions on students' intrapersonal and interpersonal domains were estimated in separate models. Methodological rigor was assessed to examine how well the overall effect sizes reflected the effects of the intervention rather than methodological influences or biases (Lipsey & Wilson, 2001). Based on the Cochrane Risk of Bias 2.0 tool for

Cluster Randomized Trials (Higgins et al., 2016) randomization (random vs. quasi-random assignment) and completeness of outcome data (percentage of drop-out) were analyzed as covariates. Additionally, the type of comparison group (passive: No intervention/waitlist vs. active: Care as usual/other intervention) was examined as covariate to examine absolute versus relative effects of the interventions. Characteristics of methodological rigor that predicted the overall effect sizes were included as covariates in further analyses.

To analyze which components were associated with stronger or weaker intervention effects, moderation analyses were conducted. Moderation analyses were conducted only if both levels of the moderator (i.e., component present or not) contained at least three effect sizes (Crocetti, 2016).

Results

Descriptive Characteristics

The present meta-analysis included 104 publications reporting on 99 unique interventions. In total, 529 effect sizes were extracted from the publications comparing the intervention with the control condition on the intrapersonal domain ($k = 218$) or the interpersonal domain ($k = 311$). Four effect sizes were extreme outliers, more than four standard deviations above the mean. All were derived from the same study (Haynes & Avery, 1979) and believed to be unrepresentatively high. These four effect sizes were therefore winsorized.

The studies, published between 1979 and 2019 (Median publication year: 2013), were conducted in the USA ($k = 36$), Canada ($k = 2$), Europe ($k = 45$), Australia ($k = 7$), Asia ($k = 13$), and Africa ($k = 1$). In total, the studies comprised 97,884 participants with an average age of 13.70 years ($SD = 1.50$) at the start of the intervention and mean sex distribution of 49% boys ($SD = 16.43$). Of the studies reporting ethnicity (59%), participants represented mostly ethnic majority in 59% of the studies, mostly ethnic minorities in 28% of the studies, and mixed ethnic majority and minorities in 13% of the studies. Appendix B provides the key characteristics of the included studies.

Overall Effect Sizes

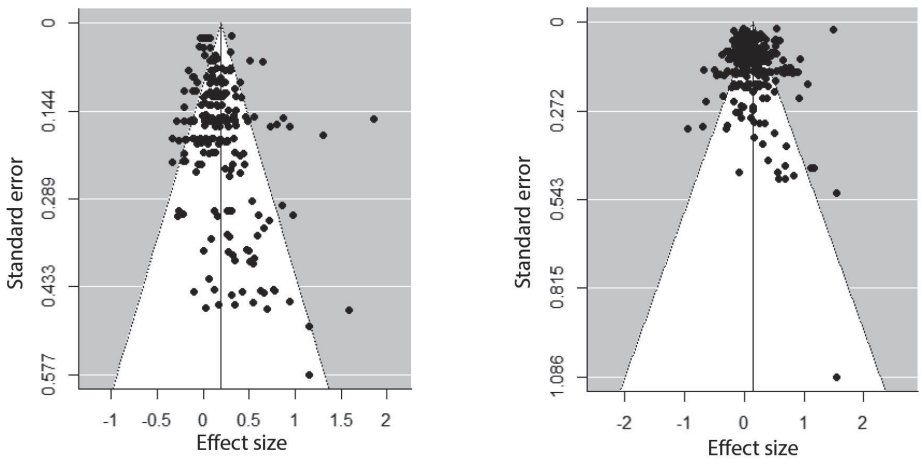
Interventions had a small positive effect on students' intrapersonal domain ($d = .19$, 95% CI [.13; .25]). More specifically, the positive intervention effects for self-esteem and self-regulation were somewhat stronger than for internalizing problems and wellbeing. No significant intervention effect was found for resilience (see Table 1). Interventions

also had a small positive effect on students' interpersonal domain (d = .15, 95% CI [.10; .19]). The magnitude of intervention effects was fairly similar for aggression, sexual health, social competence, and bullying. Interventions showed the strongest positive effects on school climate. However, this effect did not reach significance due to the small number of effect sizes for this subdomain.

Table 1 Effectiveness of Interventions Targeting the Intra- and Interpersonal Domains

Domains	Effect sizes *(k)*	Effect size	95% CI
Intrapersonal	**218**	**.19**	**.13; .25**
Resilience	13	.06	-.01; .14
Self-esteem	53	.25	.11; .39
Self-regulation	33	.21	.08; .33
Wellbeing	63	.13	.08; .19
Internalizing problems	50	.19	.10; .29
Interpersonal	**311**	**.15**	**.10; .19**
Sexual health	61	.16	.07; .26
Social competence	63	.16	.10; .23
School climate	17	.24	-.11; .58
Aggression	84	.10	.03; .17
Bullying	82	.13	.03; .24

Figure 2. Funnel plot of effect sizes concerning the intrapersonal (left) and interpersonal (right) domains.



Publication Bias

For both the intra- and interpersonal domains, the distribution of effect sizes appeared to be symmetrical (Egger's regression test: Intrapersonal $z = -.22$, $p = .826$; Interpersonal $z = .17$, $p = .862$; see Figure 2), indicating that there was low risk of publication bias.

Intervention Components Related to Intervention Effects

Preliminary analyses. Interventions targeting students' intrapersonal domain (see Figure 3) and those targeting students' interpersonal domain (see Figure 4) shared many commonly used components. Most commonly used content components are teaching students social skills, emotion regulation, and insight building. Most commonly used instructional components are implementing discussions, practice, and didactic instruction. The most commonly used structural component is additional individual guidance during the intervention.

Concerning methodological rigor, whether or not participants were randomized, drop-out rate, and type of comparison group were not related to effect sizes concerning the intrapersonal domain or the subdomains. Whether or not participants were randomized was related to effect sizes concerning the interpersonal domain; randomized studies yielded stronger effects. Percentage of drop-out was related to effect sizes concerning social competence; studies with lower drop-out rates yielded stronger effects. Whether or not participants were randomized and drop-out rates were also related to effect sizes concerning bullying; randomized studies and higher drop-out rates yielded stronger effects. Therefore, randomization and drop-out were added as covariates when it were significant predictors of the effect size in the moderation analyses concerning the interpersonal domain, social competence, and bullying.

Intrapersonal domain. Of the eight content components, none were significantly related to intervention effects on students' intrapersonal domain in general (see Table 2). However, there was a trend that teaching emotion regulation had weaker intervention effects on the intrapersonal domain overall, and specifically on self-esteem. For the subdomains, teaching assertiveness had weaker effects on internalizing problems. Furthermore, there was a trend that insight building had stronger effects on resilience.

Of the ten instructional components, practicing during the intervention had significantly stronger intervention effects on students' intrapersonal domain overall. None of the other components were significantly related to intervention effects on students' intrapersonal domain in general.

Of the five structural components, none were related to intervention effects on students' intrapersonal domain in general. For the subdomains, stronger effects were found on internalizing problems when the whole school was involved and when the intervention had more sessions.

Figure 3. Frequencies of components applied in interventions targeting the intrapersonal domain and subdomains

Figure 4. Frequencies of components applied in interventions targeting the interpersonal domain and subdomains.

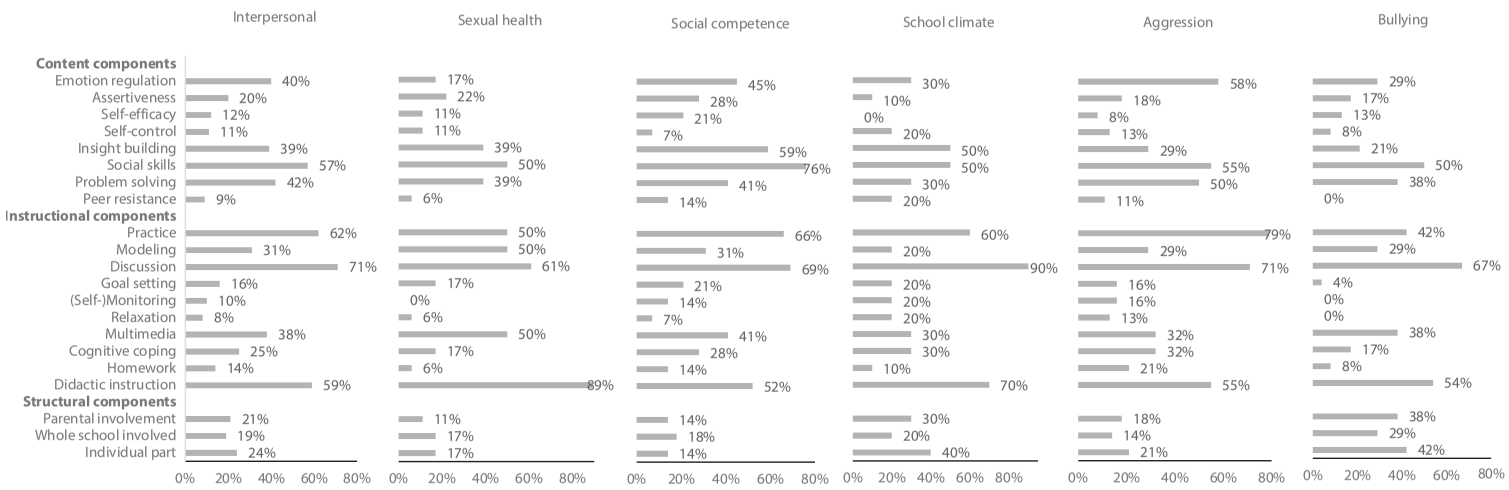


Table 2 *Effect Sizes of Interventions With and Without Components Targeting the Intrapersonal Domain*

Component		Intrapersonal			Resilience			Self-esteem			Self-regulation			Wellbeing			Internalizing		
		#ES	ES	*B*	#ES	ES	*B*	#ES	ES	*B*	#ES	ES	*B*	#ES	ES	*B*	#ES	ES	*B*
Content components																			
Emotion regulation	Yes	116	.14	-.11[†]	11	-	-	26	.13	-.24[†]	18	.20	-.03	27	.12	-.05	23	.17	-.03
	No	102	.25		2	-		27	.37		15	.23		36	.16		27	.21	
Assertiveness	Yes	61	.18	-.02	9	.12	.08	14	.25	-.00	9	.07	-.22	12	.15	.02	15	.04	-.21[*]
	No	157	.20		4	.04		39	.26		24	.29		51	.13		35	.26	
Self-efficacy	Yes	61	.20	.00	9	.07	-.02	19	.23	-.04	10	.13	-.14	12	.14	.02	10	.08	-.15
	No	157	.19		4	.09		34	.27		23	.27		51	.13		40	.23	
Self-control	Yes	18	.07	-.13	0	-	-	2	-	-	9	.15	-.08	5	.05	-.09	2	-	-
	No	200	.20		13	-		51	-		24	.23		58	.14		48	-	
Insight building	Yes	92	.18	-.03	7	.22	.18[†]	24	.20	-.11	15	.29	.15	30	.12	-.03	16	.28	.13
	No	126	.21		6	.04		29	.31		18	.14		33	.15		34	.15	
Social skills	Yes	99	.20	.02	9	.07	-.02	31	.31	.13	17	.23	.07	21	.09	-.07	21	.16	-.07
	No	119	.18		4	.09		22	.18		16	.16		42	.17		29	.23	
Problem solving	Yes	51	.24	.07	6	.19	.15	17	.34	.12	7	.30	.13	5	.19	.06	17	.20	.01
	No	167	.17		7	.04		36	.21		26	.17		58	.13		33	.19	
Peer resistance	Yes	8	.17	-.02	0	-	-	2	-	-	3	.27	.08	3	.13	-.01	0	-	-
	No	210	.20		13	-		54	-		30	.20		60	.13		50	-	

Table 2 *Continued.*

Component		Intrapersonal			Resilience			Self-esteem			Self-regulation			Wellbeing			Internalizing		
		#ES	ES	*B*	#ES	ES	*B*	#ES	ES	*B*	#ES	ES	*B*	#ES	ES	*B*	#ES	ES	*B*
Instructional components																			
Practice	Yes	148	.24	.12*	11	-	-	34	.34	.23	21	.24	.11	42	.14	.01	39	.22	.09
	No	70	.12		2	-		19	.11		12	.12		21	.13		11	.13	
Modeling	Yes	34	.16	-.05	0	-	-	13	.26	.00	5	.08	-.15	7	.14	.01	9	.08	-.13
	No	184	.20		13	-		40	.25		28	.23		56	.13		41	.21	
Discussion	Yes	143	.20	.01	10	.13	.09	40	.21	-.20	16	.26	.12	42	.14	.01	33	.21	.05
	No	75	.19		3	-[a]		13	.41		17	.14		21	.13		17	.16	
Goal setting	Yes	27	.15	-.06	6	.19	.15	8	.16	-.12	4	.01	-.24	9	.12	-.01	1	-	-
	No	191	.20		7	.04		45	.28		29	.25		54	.14		49	-	
(Self-)monitoring	Yes	29	.25	.07	0	-	-	5	.44	.21	1	-	-	10	.10	-.04	13	.35	.19
	No	189	.18		13	-		48	.23		32	-		53	.14		37	.16	
Relaxation	Yes	50	.17	-.03	2	-	-	4	.06	-.21	12	.19	-.03	18	.10	-.04	14	.19	-.01
	No	168	.20		11	-		49	.27		21	.22		45	.14		36	.20	
Multimedia	Yes	60	.16	-.06	2	-	-	21	.16	-.15	4	.18	-.03	24	.15	.04	8	.19	-.00
	No	158	.22		11	-		32	.31		29	.21		39	.12		42	.19	
Cognitive coping	Yes	90	.15	-.07	3	.04	-.11	12	.22	-.05	22	.13	-.18	20	.15	.03	28	.15	-.09
	No	128	.22		10	.15		41	.26		11	.31		43	.12		22	.23	
Homework	Yes	50	.21	.03	2	-	-	7	.34	.10	4	.17	-.05	21	.17	.05	16	.24	.07
	No	168	.19		11	-		46	.24		29	.21		42	.12		34	.18	
Didactic instruction	Yes	116	.19	-.01	4	.09	.02	28	.28	.05	16	.17	-.07	41	.14	.02	26	.23	.07
	No	102	.20		9	.07		25	.23		17	.25		22	.12		24	.16	

Table 2 *Continued.*

Component		Intrapersonal			Resilience			Self-esteem			Self-regulation			Wellbeing			Internalizing		
		#ES	ES	*B*	#ES	ES	*B*	#ES	ES	*B*	#ES	ES	*B*	#ES	ES	*B*	#ES	ES	*B*
Structural components																			
Parental involvement	Yes	6	.18	-.01	0	-	-	2	-	-	0	-	-	1	-	-	3	.29	.11
	No	212	.19		13	-		51	-		38	-		62	-		47	.18	
Whole school involved	Yes	20	.22	.02	1	-	-	10	.20	-.11	2	-	-	6	.11	-.03	3	.51	.36**
	No	198	.20		12	-		43	.31		31	-		57	.13		47	.15	
Individual part	Yes	48	.19	-.00	0	-	-	11	.23	-.03	11	.15	-.07	7	.12	-.01	16	.22	.04
	No	170	.19		13	-		42	.26		22	.22		56	.13		34	.18	
Number of sessions	I		.20	-.00		.11	-.01		.27	-.04		.19	.05		.13	-.03		.22	.20**
Number of components	I		.20	-.01		.09	.08		.26	-.04		.22	-.03		.13	-.01		.19	-.01

Note. ES = effect size; *B* = meta-regression coefficient; I = intercept; [a] did not converge

Differences in effect sizes between interventions with and without that component significant at level of † $p < .10$; * $p < .05$; ** $p < .01$.

In sum, interventions that included insight building, where students practiced during the sessions, that involved the whole school staff, and those that had more sessions showed stronger effects for stimulating subdomains of the intrapersonal domain than interventions that did not include these components. Interventions that taught emotion regulation and assertiveness showed weaker effects for stimulating subdomains of the intrapersonal domain than interventions that did not include these components.

Interpersonal domain. Of the eight content components, teaching problem solving was related to stronger intervention effects on students' interpersonal domain in general, specifically for interventions targeting bullying and school climate (see Table 3). In addition, there was a trend that insight building had stronger intervention effects on students' interpersonal domain overall. The other components were not related to intervention effects on the interpersonal domain in general. For the subdomains, insight building had stronger effects on social competence and bullying. Teaching emotion regulation and assertiveness had weaker effects on respectively bullying and aggression.

Of the ten instructional components, cognitive coping had weaker intervention effects on students' interpersonal domain overall, and specifically for interventions targeting bullying. The other components were not related to intervention effects on the interpersonal domain in general. Regarding the subdomains, using multimedia had stronger effects on social competence. In addition, there were trends that using didactic instruction and relaxation had stronger effects and using discussion had weaker effects on aggression. Of the five structural components, none were related to intervention effects on students' interpersonal domain in general. Concerning the subdomains, three components were related to stronger intervention effects. Interventions that included more sessions had stronger effects on bullying. Interventions that involved the whole school had stronger effects on bullying and school climate. Interventions with additional individual guidance had stronger effects on school climate and showed a trend that it had stronger effects on aggression. Furthermore, there was a trend that interventions that involved parents had stronger effects on school climate and that interventions that included more components had stronger effects on bullying.

In sum, interventions that taught insight building, and problem solving, used didactic instruction, relaxation, and multimedia, involved the whole school and parents, included additional individual guidance, more sessions, and more components showed stronger intervention effects for stimulating subdomains of the interpersonal domain than interventions that did not include these components. Interventions that taught emotion regulation, and assertiveness, and applied cognitive coping, and discussions showed weaker intervention effects for stimulating subdomains of the interpersonal domain than interventions that did not include these components.

Table 3 *Effect Sizes of Interventions With and Without Components Targeting the Interpersonal Domain*

Component		Interpersonal[a]			Sexual health			Social competence[b]			School climate			Aggression			Bullying[c]		
		#ES	ES	*B*	#ES	ES	*B*	#ES	ES	*B*	#ES	ES	*B*	#ES	ES	*B*	#ES	ES	*B*
Content components																			
Emotion regulation	Yes	127	.04	-.06	55	.00	-.19	31	.17	-.03	5	-.04	-.39	53	.10	-.01	52	.03	-.16†
	No	184	.11		6	.19		32	.20		12	.35		31	.10		30	.18	
Assertiveness	Yes	52	.09	.02	9	.23	.08	14	.09	-.14	3	.46	.26	15	-.05	-.19*	11	.23	.12
	No	259	.08		52	.15		49	.23		14	.21		69	.13		71	.11	
Self-efficacy	Yes	37	.11	.04	5	.28	.13	18	.16	-.01	0	-	-	8	-.04	-.15	5	.34	.23
	No	274	.07		56	.15		45	.16		17	-		76	.11		77	.11	
Self-control	Yes	25	.17	.10	5	-.03	-.22	2	-	-	3	.18	-.07	9	.24	.16	6	.38	.27
	No	286	.07		56	.19		61	-		14	.25		75	.08		76	.11	
Insight building	Yes	118	.12	.10†	20	.18	.03	41	.24	.16*	7	.09	-.29	20	.08	-.04	26	.56	.55**
	No	193	.03		41	.15		22	.08		10	.38		64	.11		56	.02	
Social skills	Yes	185	.08	.01	23	.10	-.12	51	.19	.01	9	.13	-.20	49	.11	.02	51	.19	.12
	No	126	.08		38	.22		12	.18		8	.33		35	.09		31	.08	
Problem solving	Yes	145	.21	.12*	39	.19	.04	24	.16	.00	6	.67	.63†	48	.13	.06	27	.30	.27**
	No	166	.10		22	.15		39	.16		11	.04		36	.07		55	.03	
Peer resistance	Yes	21	.08	.00	3	.18	.02	6	.21	.06	4	-.02	-.32	8	-.02	-.13	0	-	-
	No	290	.08		58	.16		57	.15		13	.30		76	.12		82	-	

Table 3 *Continued.*

Component		Interpersonal[a]			Sexual health			Social competence[b]			School climate			Aggression			Bullying[c]		
		#ES	ES	*B*	#ES	ES	*B*	#ES	ES	*B*	#ES	ES	*B*	#ES	ES	*B*	#ES	ES	*B*
Instructional components																			
Practice	Yes	187	.10	.04	22	.14	-.06	41	.19	.08	11	.25	.03	69	.11	.02	42	-.01	.13
	No	124	.06		39	.19		22	.11		6	.22		15	.09		40	-.14	
Modeling	Yes	95	.07	-.01	22	.15	-.02	13	.16	-.03	3	-.11	-.42	23	.15	.07	34	.10	-.05
	No	216	.08		39	.18		50	.20		14	.31		61	.08		48	.15	
Discussion	Yes	212	.08	-.02	27	.14	-.07	42	.21	.07	14	.26	.24	64	.06	-.13[†]	61	-.04	.14
	No	99	.09		34	.21		21	.14		3	.02		20	.20		21	-.18	
Goal setting	Yes	45	.10	.03	6	.13	-.05	15	.16	.00	5	.02	-.27	16	.04	-.07	2	-	-
	No	266	.07		55	.17		48	.16		12	.29		68	.11		80	-	
(Self-)monitoring	Yes	30	.12	.04	0	-	-	12	.13	-.04	5	-.02	-.32	13	.16	.07	0	-	-
	No	281	.07		61			51	.17		12	.30		71	.09		82		
Relaxation	Yes	20	.08	-.00	3	-.07	-.24	3	.06	-.15	4	.18	-.07	10	.29	.21[†]	0	-	-
	No	291	.08		58	.18		60	.21		13	.25		74	.08		82		
Multimedia	Yes	115	.09	.02	22	.23	.14	24	.28	.15[*]	6	.05	-.27	26	.08	-.03	35	.16	.05
	No	196	.07		39	.10		39	.13		11	.32		58	.11		47	.12	
Cognitive coping	Yes	84	-.02	-.13[*]	23	.07	-.11	49	.12	-.10	5	-.01	-.35	24	.09	-.02	18	.13	-.32[*]
	No	227	.11		38	.18		14	.21		12	.34		60	.11		64	.19	
Homework	Yes	30	.14	.07	3	.18	.02	5	.33	.15	2	-	-	15	.18	.09	4	.37	.26
	No	281	.07		58	.16		58	.18		15	-		69	.09		78	.11	
Didactic instruction	Yes	177	.12	.09[†]	58	.17	.07	27	.17	.02	13	.36	.44	44	.16	.13[†]	35	.14	.01
	No	134	.03		3	.10		36	.15		4	-.08		40	.03		47	.13	

Table 3 *Continued.*

Component		Interpersonal[a]			Sexual health			Social competence[b]			School climate			Aggression			Bullying[c]		
		#ES	ES	*B*	#ES	ES	*B*	#ES	ES	*B*	#ES	ES	*B*	#ES	ES	*B*	#ES	ES	*B*
Structural components																			
Parental involvement	Yes	62	.12	.05	16	.11	-.06	5	.03	-.18	6	.63	.57†	16	.10	-.01	19	.20	.12
	No	249	.07		45	.17		58	.20		11	.06		68	.10		63	.09	
Whole school involved	Yes	52	.09	.02	19	.06	-.13	6	.03	-.18	4	.98	.94**	9	.14	.03	13	.35	.32**
	No	259	.08		42	.19		57	.20		13	.04		75	.11		69	.03	
Individual part	Yes	56	.10	.02	7	.06	-.13	5	.13	-.06	6	.62	.63*	14	.24	.17†	24	.11	-.05
	No	255	.08		54	.19		58	.20		11	-.02		70	.07		58	.15	
Number of sessions	I		.07	.01		.11	-.14		.22	.16		.16	.16		.11	.04		-.09	.20**
Number of components	I		.07	.03		.16	-.03		.19	.02		.24	-.02		.10	.02		.16	.10†

Note. ES = effect size; *B* = meta-regression coefficient; I = intercept; [a] Analyses corrected for 'randomization', except the analysis concerning 'problem solving.'; [b] Analyses corrected for 'drop-out', except the analyses concerning 'problem solving', 'self-efficacy', 'didactic instruction', 'practice', 'goal setting', and '(self-)monitoring.'; [c] Analysis concerning 'insight building' corrected for 'drop-out', analyses concerning 'practice', 'discussion', and 'number of sessions' corrected for 'randomization.'; Differences in effect sizes between interventions with and without that component significant at level of † $p < .10$; * $p < .05$; ** $p < .01$.

Discussion

It is important to understand the intervention components that contribute to intervention effectiveness, or ineffectiveness, in order to guide intervention selection and implementation. Schools strive to improve their students' wellbeing, but their time and resources to invest in interventions are limited. This meta-analysis aimed to identify the intervention components that contribute to the effectiveness of universal secondary school-based interventions aiming to stimulate students intra- and interpersonal domains. In line with previous meta-analyses examining universal school-based interventions, we found small positive effects on students' intra- and interpersonal domains (e.g., Dray et al., 2017; Durlak et al., 2011; Jiménez-Barbero et al., 2016). Overall, none of the discrete components were consistently related to stronger or weaker effects on both students' intra- and interpersonal domains across the subdomains. In other words, components that were related to stronger or weaker intervention effects typically were so for more specific domains only. In terms of the type of components that matters most, content components seemed more relevant for stimulating the intrapersonal domain (e.g., self-esteem), and structural components seemed more relevant for stimulating the interpersonal domain (e.g., bullying). Importantly, components related to stronger intervention effects were not necessarily frequently implemented in interventions (e.g., in 10% - 19% of the interventions). Similarly, components related to weaker intervention effects were generally implemented frequently (e.g., in 40% - 53% of the interventions).

Content components teaching students self-awareness (i.e., insight building) and problem solving were related to stronger effects, whereas components teaching emotion regulation and assertiveness were related to weaker effects. These findings are in line with previous research that indicated teaching self-awareness and problem solving as potential effective components (e.g., Boustani et al., 2015). Teaching emotion regulation and assertiveness might be more relevant in different contexts than the secondary school context in which the interventions were implemented. For instance, emotion regulation might be more relevant when implemented in psychotherapy (e.g., Weisz et al., 2012) and teaching assertiveness might be more relevant for students at elementary schools (e.g., Onrust et al., 2016). In other words, content components might be differentially related to intervention effects in different contexts.

In general, instructional components that reflect an active learning approach were related to stronger intervention effects (e.g., relaxation, practice). This does not mean that interventions should only use active learning approaches and exclude more passive learning approaches. Interventions that used discussion as method delivery, an active learning component, were related to weaker intervention effects on aggression, whereas interventions using a didactic information delivery as method, a passive learning approach, were related to stronger intervention effects. These findings are in line with the meta-analysis of De Mooij, Fekkes, Scholte, and Overbeek (2020) that

showed that psychoeducation was related to stronger effects of Social Skills Training interventions. Using didactic instruction might fit better in the school context than using discussion. In a didactic instruction approach, the emphasis is on knowledge transfer between the teacher and the students, whereas a discussion approach is more dependent on the students and the social skills and cohesiveness of the group. Teachers might be less equipped to prevent a discussion from sidetracking than to teach psychological constructs (Horne, Stoddard, & Bell, 2007).

Teaching cognitive coping was related to weaker effects on interpersonal competences in general. This finding was somewhat surprising, given that cognitive coping is considered an effective component in other interventions as Cognitive Behavioral Therapy (Yovel, Mor, & Shakarov, 2014) with well-trained therapists (Kobak, Wolitzky-Taylor, Craske, & Rose, 2017). One possible explanation may be that school-based interventions, are often implemented by teachers who only received a short training (e.g., Challen, Machin, & Gillham, 2014) and have no to little experience in teaching cognitive coping.

The results concerning structural components showed that longer and more extensive interventions (e.g., involving parents and the whole school) were more effective for targeting system level outcomes such as school climate. Long-term and extensive interventions might be more effective when the intervention aims to increase students' feelings of safety at school. By targeting multiple systems in which the students are involved (e.g., school, family) teachers and parents might become more sensitive for problems students encounter, such as bullying or problematic relations with peers (Ttofi & Farrington, 2011) and a broad range of risk factors is addressed (Trip et al., 2015). For interventions targeting the individual level such as self-esteem, more extensive interventions were not related to stronger effects nor were less extensive interventions related to weaker effects. Based on these findings, less extensive interventions might be preferred to stimulate the intrapersonal domain due to the easier implementation (Bakermans-Kranenburg et al., 2003), while more extensive interventions may be better suited to stimulate the interpersonal domain.

Furthermore, our results showed that components related to stronger intervention effects were not necessarily commonly implemented. For instance, interventions that involved the whole school were related to stronger effects on internalizing problems, bullying and school climate. However, only 10% to 19% of the included interventions involved the whole school. In contrast, some components that were related to weaker effects are implemented more often. For example, teaching emotion regulation, included in 40% to 53% of the interventions, was related to weaker effect sizes on the intrapersonal domain in general, self-esteem, and bullying. Our frequency counts of components are in line with the frequency count by Boustani and colleagues (2015) of effective school-based interventions. These findings indicate that it is important to critically consider which components to include in an intervention and to not simply "do what previously has been done."

Several limitations merit attention. First, we tested associations between components and intervention effectiveness. Based on these associations, we cannot state whether specifically these components are (in)effective or whether other components confounded with that specific component accounted for the association. This meta-analysis should therefore be regarded as hypothesis generating; our results give future research indications which components are interesting to examine further. Future research should test causal individual and synergistic effects of components, and potential order effects of components. Second, the coding of components depended on the sufficiency of the intervention description in the included studies; if a component was not mentioned in the article, or other freely available information concerning the evaluated interventions, it was coded as not present. At the same time, components that are formally part of the intervention, and therefore reported and coded as such, may not necessarily be implemented. It might be that some components were thus coded as "present" while they were not actually implemented. Last, even though we included more than 500 effect sizes, some components (e.g., peer resistance, parental involvement) were less frequently implemented in interventions than other components (e.g., practice, discussion) resulting in better powered analyses for some components than for others.

In conclusion, when designing and implementing universal school-based interventions, and especially when no rigorous evidence base for the intervention is available, it is important to consider the evidence base of its included components. Some components are often implemented in interventions without being actually related to stronger intervention effects. In fact, some commonly implemented components (e.g., emotion regulation, discussion) were related to weaker intervention effects in our meta-analysis. Vice versa, some components that were related to stronger intervention effects (e.g., involvement of the whole school or parents) were only rarely included in interventions. Thus, it is essential to examine the evidence base of components before including it in an intervention, and to not solely focus on which components have been included in previous interventions. Another aspect of the evidence base of components to consider is which type of components is emphasized in the intervention. For interventions aiming to stimulate the intrapersonal domain several content components appear to be important, whereas interventions aiming to stimulate the interpersonal domain several structural components appear to be important. This meta-analysis provides an empirical foundation for the evidence base of components related to stronger and weaker effects for universal school-based interventions targeting the intra- and interpersonal domain.

Supplementary Materials

Appendix A – Definitions of Components

Content components = *specific skills adolescents learn to promote positive outcomes*	
Emotion regulation	Strategies to help youth identify and appropriately express emotions (including aggression)
Assertiveness	Exercises designed to promote the youth's ability to assert his or her needs appropriately with others
Self-efficacy	Techniques and training to enhance self-confidence and improve self-efficacy
Self-control	Strategies to help youth interrupt undesired behavioral tendencies (e.g., impulses) and refrain from acting on them.
Insight building	Activities specifically designed to help a youth achieve greater self-understanding and adjust attitudes
Social skills	Training youth how to communicate more effectively with others and providing constructive information, training, and feedback to improve interpersonal verbal or non-verbal functioning
Problem solving	Training in the use of techniques, discussions, or activities designed to bring about solutions to social, emotional, or behavioral problems
Peer resistance	Techniques or training to learn youth how to resist pressure from peers
Instructional components = *techniques and methods of information delivery used by the intervention facilitator*	
Practice	Practicing of a desired behavior during session (e.g., role-play)
Modeling	Demonstration to the youth of a desired behavior
Discussion	Discussion of topics within a group
Goal setting	The explicit selection of a therapeutic goal for the purpose of working toward achieving that goal
(Self-)monitoring	The repeated measurement of a target index (by the youth)
Relaxation	Techniques or exercises designed to induce physiological calming
Multimedia	The use of multimedia to bring or reinforce new knowledge or skills
Cognitive coping	Any techniques designed to alter interpretation of events or deal with stressful situations through examination of the youth's reported thoughts (e.g., cognitive restructuring)
Homework	Written, verbal, or behavioral assignments to complete between sessions
Didactic instruction	The formal (usually didactic) review of information (e.g., psychoeducation)
Structural components = *describe the structure of the intervention that might impact results*	
Parental involvement	Parents are directly or indirectly involved during the intervention
Whole school involved	The school staff is directly or indirectly involved during the intervention
Individual part	The intervention includes additional individual guidance or explicit individual progress through the intervention (e.g., expressive writing, internet-based intervention)
Number of sessions	Number of sessions of the intervention
Number of components	Number of components implemented in the intervention

Appendix B – Descriptives of Included Publications

Reference	Name intervention	Aim	Target	Category outcomes	Grades
Adler-Beader 2007	LoveU2: increasing your relationship smarts (RS adapted)	Healthy romantic relationships	Inter	Sexual health, social competence	9-12
Allara 2019	Diario della salute (my health diary)	Wellbeing and health	Intra	Wellbeing, aggression, school climate	7
Ando 2007	Adaptation of Going Places Program	Aggressive behavior	Inter	Self-regulation, aggression, social competence, school climate	7
Avery-leaf 1997	Dating violence prevention program	Dating violence	Inter	Sexual health	11, 12
Baker 2014	Respect	Sexual violence	Inter	Sexual health	9-12
Barkoukis 2016	Intervention against Cyberbullying	Cyberbullying	Inter	Social competence, bullying	7-11
Bonell 2017	Learning Together intervention	Bullying, aggression, and wellbeing	Inter	Wellbeing, aggression, bullying	7
Bosworth 2004	SMART Talk (based on BARN system)	Problem solving without violence	Inter	Self-esteem, aggression, social competence	6-8
Boulton 1996	Sticks and stones video	Bullying	Inter	Bullying	6-9
Bradley 2010	TestEdge	Stress, anxiety, wellbeing, and relationships	Intra and inter	Self-regulation, internalizing, wellbeing, social competence, school climate	10
Bull 2009	The fairplayer.manual	Bullying and relational aggression	Inter	Bullying	9-11
Burckhardt 2016	Strong Minds	Subjective wellbeing	Intra	Wellbeing	9, 10
Burckhardt 2017	Dialectical behavior therapy skills group	Mental health symptoms	Intra	Self-regulation, internalizing, aggression	10
Burckhardt 2017b	Acceptance and Commitment Therapy (ACT)	Wellbeing	Intra	Internalizing, wellbeing	10
Calear 2009	MoodGym	Anxiety and depression	Intra	Internalizing	8-10
Caplan 1992	The Positive Youth Development Program	Personal and social competence	Intra and inter	Self-esteem, self-regulation, wellbeing, social competence	6, 7
Caprara 2014	CEPIDEA	Prosocial behavior	Inter	Self-esteem, aggression, social competence	7
Carraro 2014	Play fighting	Aggressive behavior	Inter	Aggression	8
Castillo-Gualda 2017	INTEMO (Long version, 3 years)	Aggression	Inter	Wellbeing, aggression	7
Challen 2014	UK Resilience Program	Resilience	Intra	Internalizing, social competence	6
Chang 2013	Laughing Qigong Program	Stress	Intra	Self-esteem, wellbeing	7
Cheung 2010	Character education	Social competence	Inter	Social competence	8, 9

Reference	Name intervention	Aim	Target	Category outcomes	Grades
Coelho 2015	Positive Attitudes	Social-emotional competence	Intra and inter	Self-esteem, self-regulation, internalizing, social competence	7-9
Coker 2017	Green Dot violence prevention program	Sexual violence and interpersonal violence	Inter	Aggression	9-12
Connolly 2015	Respect in Schools Everywhere (RISE)	Bullying, sexual harassment, and date aggression	Inter	Internalizing, sexual health, bullying, school climate	7, 8
Constatine 2015	Sexuality Education Initiative	Sexual health	Inter	Self-esteem, aggression	9
Cross 2016	Cyber Friendly School	Cyberbullying	Inter	Bullying	8, 9
Daly 2015	Yoga	Emotion regulation	Intra	Self-regulation	NR
De Graaf 2016	Rock and Water	Sexual aggressive behavior and cognitions	Inter	Self-esteem, self-regulation, sexual health	9, 10
De Villiers 2012	Resilience program	Resilience	Intra	Resilience, self-esteem, self-regulation, social competence	6
Domino 2013	Take the lead	Social skills	Inter	Bullying	7
DuRant 1996	Violence prevention curriculum for adolescents	Violence use	Inter	Aggression	6-8
DuRant 1996	Conflict resolution: A curriculum for youth providers	Violence use	Inter	Aggression	6-8
Espelage 2013	Second Step: Student Success Through Prevention	Violence	Inter	Aggression, sexual health, bullying	6
Felver 2018	Learning to BREATHE	Wellbeing and learning	Intra	Resilience	9-12
Foshee 2005	Safe Dates	Dating violence	Inter	Aggression	8, 9
Frank 2016	Transformative life skills	Stress and social-emotional health	Intra	Self-regulation, wellbeing, aggression, school climate	6, 9
Freire 2018	Challenge: To Be +	Positive development	Intra and inter	Self-esteem, wellbeing	9
Garaigordobil 2004	Psychological intervention carried out with groups of adolescents	Emotional development	Intra	Self-esteem, self-regulation, internalizing, social competence	NR
Garaigordobil 2015a, Garaigordobil 2015b, Garaigordobil 2014	Cyberprogram 2.0	Interpersonal conflicts and self-esteem	Intra and inter	Self-esteem, internalizing, social competence, bullying, aggression	9, 10
Gardner 2004	Connections: Relationships and Marriage	Healthy romantic relations	Inter	Aggression, sexual health	11, 12
Ghahremani 2013	Youth empowerment SEminar (YES)	Emotional wellbeing	Intra	Self-regulation	7-12

Reference	Name intervention	Aim	Target	Category outcomes	Grades
Gianotta 2009	Expressive writing	Negative outcomes associated with peer-related problems	Inter	Self-regulation, bullying	7
Gigantesco 2015	Definizione di obiettivi e soluzione di problemi (establishing goals and problems solving)	Self-efficacy, psychological wellbeing, and life satisfaction	Intra	Self-regulation, wellbeing	9-11
Gollwitzer 2007	Vienna Social Competence Training (ViSC)	Class commitment, responsibility, and nonaggressive behavior in conflict	Inter	Aggression	6-8
Gouda 2016	Mindfulness-based stress education group program	Performance pressure	Intra	Self-esteem, self-regulation, internalizing, wellbeing, social competence	11
Haines, 1994	Stress inoculation training	Negative arousal in response to stress	Intra	Internalizing, wellbeing, aggression	9-12
Hains 1990	Cognitive intervention training program	Cope with stress and negative arousal	Intra	Self-esteem, internalizing, aggression	11, 12
Hains 1994	Stress inoculation training	Negative arousal in response to stress	Intra	Self-esteem, internalizing, wellbeing, aggression	NR
Haynes 1979	Communication skills training program	Self-disclosure and empathy	Intra	Self-esteem, social competence	11
Horn 2010	JES! Jugendpräventions-programm mit Expressivem Schreiben	Emotion regulation	Intra	Wellbeing	8
Huppert 2010	Mindfulness-based stress reduction	Mindfulness, resilience, and psychological wellbeing	Intra	Resilience, wellbeing	8
Ingram 2019	Stand up: Virtual reality to activate bystanders against bullying	Bullying	Inter	Aggression, social competence, bullying, school climate	7, 8
Jaycox 2006	Ending violence: A curriculum for educating teens on domestic violence and the las	Intimate partner violence	Inter	Sexual health	9
Jiménez-Barbero 2013	Count on me	Bullying and violence	Inter	Aggression	7, 8
Kasler 2013	Meaning of Life program (Israeli adaptation of the Laws of Life program)	Meaning in life	Intra	Self-esteem, wellbeing, school climate	10, 11
Kaveh 2014	Peer led training program	Self-esteem	Intra	Self-esteem	7
Khanna 2016	Nice Thinking (adjusted to Indian culture)	Gratitude and wellbeing	Intra	Wellbeing	7

Reference	Name intervention	Aim	Target	Category outcomes	Grades
Kiselica 1994	Stress inoculation training with assertiveness training	Anxiety and stress	Intra	Internalizing, wellbeing	9
Klingman 1993	Cognitive-behavioral oriented distress-coping training	Distress-coping	Intra	Wellbeing, social competence	8
Kozina 2018a Kozina 2018b	My FRIENDS	Anxiety	Intra	Internalizing, aggression	8
Lamke 1988	Cognitive-behavior modification program	Self-statements and self-esteem	Intra	Self-esteem	9
Macgowan 1997	Dating violence prevention program	Dating violence	Inter	Aggression	6-8
Menesini 2003	Peer support model	Bullying	Inter	Bullying	6-8
Muck 2018	Scientist practitioner program	Sexual violence	Inter	Sexual health	8, 9
Nash 2007	Empower youth program (and Usual school services)	Health, wellbeing, and optimism for future	Intra	Internalizing	6-8
Noggle 2012	Kripalu yoga	Overall wellbeing	Intra	Resilience, wellbeing, aggression	11, 12
Orpinas 1995	Second Step: A violence prevention curriculum	Violence	Inter	Self-esteem, aggression	6
Ortega-Barón 2019	Prev@cib	Bullying and cyberbullying	Inter	Bullying	7-10
Pacifici 2001	Dating and Sexual Responsibility	Dating violence	Inter	Sexual health	10
Proctor 2011	Strengths Gym	Build strengths and learn new strengths	Intra	Wellbeing, self-esteem	7, 8
Richardson 2009	BodyThink	Self-esteem	Intra	Self-esteem, bullying	7
Ruini 2006	Cognitive Behavioral Therapy	Mood disorder and psychobiological distress	Intra	Resilience, self-esteem, self-regulation, internalizing, wellbeing, aggression, social competence	NR
Ruini 2006	Well-being therapy (WBT)	Psychological wellbeing	Intra	Resilience, self-esteem, self-regulation, internalizing, wellbeing, aggression, social competence	NR
Ruini 2009	Well-being Therapy (WBT) with added cognitive-behavioral packages	Psychological wellbeing and optimal functioning	Intra	Self-regulation, internalizing, wellbeing, aggression, social competence	9, 10
Ruiz-Aranda 2012	INTEMO	Aggressive behaviors, psychosocial maladjustment, and mental health	Intra and inter	Self-esteem, internalizing, wellbeing, social competence	7
Sánchez-Jiménez 2018	Dat-e Adolescence	Dating violence	Inter	Self-esteem, aggression, sexual health	7-10

Reference	Name intervention	Aim	Target	Category outcomes	Grades
Schramm 2012	Relationship Smarts Plus	Healthy romantic relationships	Inter	Sexual health	8-12
Schultz 2001	Facing History and Ourselves	Perspective-taking, critical thinking, and more decisions	Inter	Self-esteem, aggression, social competence	8
Shek 2011	Positive Adolescent Training through Holistic Social Programs (P.A.T.H.S.)	Holistic youth development	Intra and inter	Resilience, self-esteem, self-regulation, social competence	7
Shinde 2018	Strengthening evidence base on school-based interventions for promoting adolescent health (SEHER)	School climate and health-promoting behaviors	Intra and inter	Internalizing, aggression, bullying, school climate	8
Shoshani 2013	Maytiv School Program	Mental health and empowerment	Intra	Self-esteem, wellbeing	7-9
Sibinga 2013	Mindfulness-based stress reduction program	Psychological symptoms and coping	Intra	Internalizing, aggression	7, 8
Simons-Morton 2005	Going Places Program	Social skills and problem behaviors	Inter	Aggression	6
Soliday 2004	Expressive writing Intervention	Positive functioning and stress	Intra	Internalizing, wellbeing	8
Solomontos-Kountouri 2016	ViSC social competence program (with added parental component)	Victimization and aggressive behavior	Inter	Aggression, bullying	7, 8
Sorrentino 2018	Tabby Improved Prevention and Intervention Program (TIPIP)	Cyberbullying and victimization	Inter	Bullying	NR
Stevens 2000	Flemish anti-bullying program	Bullying and victimization	Inter	Social competence, bullying	NR
Thomaes 2009	Self-affirmation intervention	Narcissistic aggression	Inter	Self-esteem, aggression	7, 8
Thompkins 2014	Violence Prevention Project	Conflict resolution skills	Inter	Social competence	9, 10
Tomyn 2016	Think Health and Wellbeing	Thinking style, self-esteem, and resilience	Intra	Resilience, self-esteem, internalizing	8
Trip 2015	Rational Emotive Behavioral Education (REBE)	Negative dysfunctional emotions and alternative to low frustration tolerance	Intra	Aggression, bullying	6
Trip 2015, Yanagida 2016	Viennese Social Competence (ViSC)	Bullying and aggressive behavior	Inter	Aggression, bullying	6, 7

Reference	Name intervention	Aim	Target	Category outcomes	Grades
Tunariu 2017	iNEAR	Positive identities, character strengths, and resilience	Intra	Wellbeing, social competence	6, 7
Van der Meulen 2010	Adjusted version of EQUIP program for Educators	Peer victimization	Inter	Sexual health, bullying, school climate	8, 9
Van Schoiack-Edstrom 2002	The Second Step, Middle school/Junior High program	Prosocial skills and impulsive-aggressive behavior	Inter	Aggression, social competence	6, 7
Williams 2015, Miller 2015	Start strong: Building healthy teen relationships	Healthy romantic relationships and dating violence	Inter	Aggression, bullying, sexual health, social competence	7
Williford 2013	KiVa antibullying program	Cyberbullying and victimization	Inter	Bullying	8, 9
Wong 2011	Restorative Whole-school Approach	Bullying	Inter	Self-esteem, social competence, bullying, school climate	7-9
Yom 2005	Educational Program for the Prevention of Sexual Violence	Sexual violence	Inter	Sexual health	6

Note. Intra = Intrapersonal domain; Inter = Interpersonal domain; NR = Not reported

7

General Discussion

General discussion

The focus of this dissertation was on the effectiveness of universal school-based interventions aiming to improve competencies and to prevent the development of problems in the intrapersonal (e.g., psychological wellbeing) and interpersonal (e.g., aggression) domain. The first aim of this dissertation was to study the effectiveness of these universal school-based interventions (i.e., what works). By conducting a Randomized Control Trial (RCT), I evaluated the effectiveness of a specific universal school-based intervention, Rock and Water (R&W; Ykema, 2002; 2018), for prevocational students (Chapter 3). The effectiveness of universal school-based interventions in general I examined by conducting a meta-analysis (Chapter 6). The second aim was to study whether heterogeneity in the context and in the student population affected intervention effectiveness (i.e., under what circumstances and for whom does it work?). More specifically, I focused on the extent to which intervention dosage, that is, the intervention's ecological width (i.e., the number of involved systems in an intervention) and intervention's time span (Chapter 3), and students' personality traits (Chapter 4) affected intervention effectiveness. The third aim was to study working mechanisms of universal school-based interventions (i.e., how does it work?). As mechanism of change (i.e., mediator), I examined whether classmates' modeling and reinforcement mediated the effect of R&W on the peer context in the classroom (an aspect of the interpersonal domain; Chapter 5). In the meta-analysis, I analyzed which components of interventions were associated with intervention effects in the intra- and interpersonal domains of universal school-based interventions in general (Chapter 6). In this chapter, I summarize and discuss the main findings, elaborate on strengths and limitations, make recommendations for future research and point out practical implications.

Aim 1: Effectiveness of Universal School-Based Interventions

The present dissertation is in line with previous research suggesting that prevocational students have an increased risk of developing psychosocial problems (e.g., Harakeh, De Looze, Schrijvers, Van Dorsselaer, & Vollebergh, 2012; Stevens & De Looze, 2018). My results of the RCT illustrated that during these two years prevocational students in the Control condition had a tendency to develop minor psychosocial problems, particularly in the interpersonal domain. More specifically, the students slightly decreased in psychological wellbeing (i.e., intrapersonal domain) and interpersonal relations in the class, and slightly increased in aggression, bullying, and victimization (i.e., interpersonal domain). Apparently, school policy as usual (i.e., Care As Usual; CAU) is not effective enough to prevent the development of psychosocial problems for prevocational students at this age (12 – 14 years). This finding emphasizes the necessity

of effective interventions for this group of students and highlights the importance of the current dissertation.

In general, universal school-based interventions showed some positive effects on improving competencies and decreasing problems in both the intra- and interpersonal domain. Overall, the general intervention effects were small, with somewhat stronger effects in the intrapersonal domain (Cohen's d = .19, 95%CI [.13; .25]) than in the interpersonal domain (Cohen's d = .15, 95%CI [.10; .19]). In the RCT evaluating R&W, I found a comparable pattern. R&W is roughly equally effective as other universal school-based interventions in the intrapersonal (Average Cohen's d = .23, range Cohen's d = .03 - .38) and interpersonal (Average Cohen's d = .15, range Cohen's d = .01 - .31) domain. Additionally, also R&W was more effective in the intrapersonal domain (significantly improving psychological wellbeing, sexual autonomy, and internalizing behavior, but not resilience) than in the interpersonal domain (significantly improving aggression, promising changes (but not significant) in interpersonal relations in the class and bullying, but no effect on externalizing behavior and victimization). Although most intervention effects were found in the intrapersonal domain, it is interesting to point out that in the interpersonal domain the intervention effect of R&W on aggression was relatively strong compared to the general intervention effect on aggression (R&W Cohen's d = .31 vs. meta-analysis Cohen's d = .10). Taken together, it appears a priority to examine how universal school-based interventions can be improved to meaningfully affect a broad range of competencies and problems in the interpersonal domain. This is especially important for interventions targeting prevocational students, as these students seemingly experience most problems in the interpersonal domain.

Notwithstanding the positive effects in the intrapersonal domain, it is noticeable that school-based interventions appear to be ineffective in stimulating students' resilience. Specifically, no intervention effects on resilience were found either for R&W or in the meta-analysis, even though these interventions often aim to improve resilience and many are described as "resilience-focused" (e.g., Dray et al., 2017). Improving students' resilience is important as it enables them to thrive and deal with challenges, stress, and adversity they experience in life (Campbell-Sills & Stein, 2007). More research is needed to study how interventions can improve resilience in order to help students to recover from stressful events and achieve their full potential.

Regarding secondary outcomes, R&W showed small positive effects on self-control and emotional self-regulation, showed promising (but not significant) changes in self-esteem, but showed no effect on students' self-reflection. I tested these four additional outcomes because the theory of R&W indicates these as important competencies that the intervention also aims to improve (Ykema 2002; 2018). The intervention effects of R&W on self-control, self-esteem, and emotional self-regulation were comparable in magnitude to those found in the meta-analysis (R&W Cohen's d = .22 - .29; Meta-analysis Cohen's d = .21 - .25). Besides these positive effects, R&W was not able to improve students' self-reflection. The correlations in the present study (not previously

reported in this dissertation) suggest that self-reflection might be a different type of competency since self-control, self-esteem, and emotional self-regulation correlated strongly together (r = .50 - .55), but weakly with self-reflection (r = -.15 - -.25). Perhaps self-reflection is cognitively too demanding for most prevocational students as it requires abstract thinking on the metacognition level (Sauter, Heyne, Blooöte, Van Widenfelt, & Westenberg, 2010). Alternatively, self-reflection might be stimulated through a different intervention method than self-control, self-esteem, and emotional self-regulation. Therefore, the next step is to examine how students' self-reflection can be stimulated and whether these four competencies might function as mediators given that the theory of the "Rock &Water house" suggests these competencies to be working mechanisms of the intervention (Ykema 2002; 2018). This puts the theory of the intervention to the test and increases understanding of the process of change which might be helpful for optimizing interventions (Kazdin, 2007).

Overall, the effectiveness of R&W is comparable with other universal school-based interventions addressing students' development in the intra- and interpersonal domains. It is difficult to determine the practical relevance of the improvements seen in the students. Even though the intervention effects are small, which is to be expected in universal interventions as these target a mainly healthy population (Nehmy & Wade, 2014), the potential impact can be large. If many individuals change their behavior just a bit, this could lead to large benefits in society (Greenberg & Abenavoli, 2017). Yet, the small intervention effects indicate that there is room to optimize universal school-based interventions in general and, more specifically, R&W.

Aim 2: Heterogeneity in Contexts and in Population

The second aim of this dissertation was to examine whether characteristics of the context (i.e., intervention dosage: Ecological focus and time span) and student population (i.e., personality traits) affected intervention effectiveness. Regarding the first aspect of intervention dosage, the findings showed that an intervention's ecological focus could affect intervention effects. R&W was only effective when the intervention had a narrow ecological focus, that is, when it was solely implemented during the intervention lessons. Hence, the circumstances under which an intervention is implemented can have a large influence on the intervention's effectiveness. More importantly, neglecting the context of an intervention could lead to drawing erroneous conclusions about the intervention's effectiveness. For instance, if I only examined R&W under "standard" conditions, I may have concluded that the intervention might be hardly worthwhile to implement.

The second aspect of intervention dosage, that is, the time span of intervention, also affected intervention effects. Students improved most during the first part of the intervention, whereas improvements in the second year were neglectable. This finding suggests that implementing only the first year of the intervention might be

sufficient to establish the effects in the assessed competencies and problems. Previous research also found that intervention effects decrease over time (e.g., Cuijpers, 2000; De Mooij, Fekkes, Scholte, & Overbeek, 2019) and illustrated that participants who benefit from an intervention often show these improvements early in the intervention (e.g., Lutz et al., 2014; Tadić et al., 2010). This seemingly typical decline in intervention effects may be related to participants' expectations and motivation since these are moderators of intervention effects (Ebert et al., 2013: Philips & Wennberg, 2014). Although not reported in one of the chapters in this dissertation, I assessed students' attitudes towards R&W and analyzed changes in their expectations and motivation to participate. The results showed that students' expectations and motivation to participate in the intervention indeed declined, which could have contributed to the decreases in intervention effects. Perhaps a 2-year intervention is simply too long to maintain the interest of students. Together, the findings regarding intervention dosage support previous research stating that more extensive and longer interventions are not necessarily more effective, but that sometimes "less is more" (e.g., Cuijpers, 2000; Durlak, Weissberg, Dymnicki, Taylor, & Schellinger, 2011).

Concerning characteristics of the student population, students' personality traits had generally little effect on intervention effects, confirming the universal nature of the intervention. However, three patterns of moderation did emerge. First, in line with the Risk moderation hypothesis (Spoth, Shin, Guyll, Redmon, & Azevedo, 2006), a compensatory effect was found for more vulnerable students. More specifically, students who were vulnerable to develop problems in the intra- or interpersonal domain – based on their levels of certain personality traits – benefitted most from the intervention. For instance, without intervention, students with high levels of Neuroticism improved less in psychological wellbeing than students with lower levels of this trait. However, when these vulnerable students received the intervention, they improved more in psychological wellbeing than less vulnerable students (i.e., students with low levels of Neuroticism). This pattern suggests that R&W might be able to prevent the development of problems indicating a true prevention effect (Nehmy & Wade, 2014). The second pattern suggested that high levels of Extraversion enabled students to benefit from an intervention in which many people are involved in the implementation. Extraverted students benefitted from R&W regardless of the number of involved people, whereas the involvement of more people (i.e., Standard and Plus conditions) seemed to counteract the intervention effects for less extraverted students. Perhaps the learning processes of extraverted students, characterized by cooperation and group work rather than internal processing of information (Felder, Felder, & Dietz, 2002), fits well with an intervention approach in which interactions with others are important, enabling them to benefit from the intervention. The third pattern indicated that personality traits affected intervention effects in the intrapersonal domain more than in the interpersonal domain. Personality traits can influence how one wants to feel and why (Hughes, Kratsiotis, Niven, & Holman, 2020) and might therefore be more related to one's feelings, emotions, and attitudes (i.e., the intrapersonal domain),

than their interactions and social perceptions (i.e., the interpersonal domain). Thus, an intervention's effectiveness may depend on students' vulnerability, number of people involved in the intervention, and the addressed domain. These results accentuate that there is no simple answer to the question "for whom is the intervention effective?".
The findings of the present study show the advantage of studying trajectories of change and highlight the complexity of evaluating "what works under what circumstances for whom". Modeling trajectories of change rather than mean differences at one moment in time creates the opportunity to analyze intervention effects during an intervention (Greenberg, & Abenavoli, 2017). Not only might these trajectories reveal critical points of change, they might also provide insights for opportunities to optimize interventions. Therefore, I stress the relevance of assessing participants during an intervention, as well as before and after an intervention. Such interim assessments can be conducted with a relatively low frequency, as in the present dissertation (2 measurement points), but also with a high frequency, for instance after each session. Furthermore, to add to the complexity of differential effectiveness, intervention effects were not only dependent on context, but also on the combination of context and characteristics of participants. In the Light condition (i.e., narrow ecological focus) students benefitted from R&W regardless of their level of Extraversion. In contrast, in the Standard and Plus conditions (i.e., a broader ecological focus) only Extraverted students were able to benefit from the intervention. Notwithstanding this complexity, it is essential to examine under what circumstances and for whom an intervention is effective in order to determine the generalizability of intervention effects to different circumstances and populations.

Aim 3: Working Mechanisms in Interventions

The first way in which I studied working mechanisms in interventions was through examining mechanisms of change (i.e., mediators). Classmates' modeling and reinforcement did not mediate the effect of R&W on the peer context in the classroom. Note that this conclusion has to be interpreted with caution as the mediators were only examined in a subsample and only in the first year of the intervention. Although the results could be due to research practices (e.g., conceptualization, coding), the finding that R&W was unrelated to modeling and reinforcement could also be an indication of action theory failure as the intervention is seemingly not able to successfully address the proposed mechanisms of change (O'Rourke & MacKinnon, 2018). In addition, modeling and reinforcement were generally unrelated to the peer context in the classroom, suggesting potential conceptual theory failure as well since the mechanisms of change are apparently not related to the outcomes (O'Rourke & MacKinnon, 2018). The exception is prosocial modeling which was related to less victimization in the class, especially when dyadic mutuality between classmates was high. Hence, prosocial modeling in the classroom could be a mechanism of change

for decreasing victimization in the class, amplified by dyadic mutuality. This means that interventions addressing classroom victimization, including R&W which was not effective in decreasing victimization, might be optimized by an increased focus on classmates' prosocial modeling and improving dyadic mutuality between classmates. For instance, to stimulate classmates' dyadic mutuality interventions could organize fun exercises in the class in which classmates have to work together. Even though the results of the current dissertation did not confirm the theory of R&W which suggests that classmates' modeling and reinforcement are mediators in the intervention, the results did provide useful theoretical insights concerning a potential mechanism of change that can be used to optimize interventions (Kazdin, 2007). Studying mechanisms of change is thus crucial for both theory and intervention development, and can be examined even if a direct effect of the intervention on an outcome is absent (O'Rourke & MacKinnon, 2018).

The second way in which I studied working mechanisms in interventions was through examining intervention components. The results of the meta-analysis indicated that components related to stronger intervention effects were not always commonly implemented, and conversely, some commonly implemented components were related to weaker intervention effects. Note, however, that these relations between components and intervention effects were based on correlations. Components were coded as either present or absent in the interventions included in the meta-analysis. Subsequently, the general effectiveness of interventions with this component is compared to the general effectiveness of interventions without this component. Correlational results are inherent to this type of meta-analysis and results should therefore be used for hypotheses generation rather than for drawing firm conclusions about the (in)effectiveness of individual components. Notwithstanding the correlational nature of these relations, the results suggest opportunities for optimizing interventions by adding components related to stronger effects and eliminating those related to weaker effects (Michie, Fixsen, Grimshaw, & Eccles, 2009). For instance, in line with previous research (Boustani et al., 2015; Cuijpers, 2000), my results showed that teaching students self-awareness and problem solving by means of active learning methods was related to stronger intervention effects. Hence, the results identified these components as potentially important components to include in interventions. As R&W already includes these components, it seems obvious that these components should be maintained in the program. In contrast, results from my meta-analysis also showed that teaching students assertiveness and emotion regulation, and using group discussion were related to weaker intervention effects. These findings suggest that decreasing the use of these components might optimize interventions. R&W also includes these components, thus it seems worthwhile to consider whether the use of these components should be decreased in order to improve its effectiveness.

However, optimizing interventions based on associations between components and intervention effects is less straightforward than it seems. Whether a component is related to intervention effects might depend on the circumstances of implementation

and characteristics of the population (Michie et al., 2009). The effectiveness of a component can be affected by interplay among components when combined for intervention. For instance, certain combinations or sequences of components might inflate, or reduce, their individual effectiveness (Collins, Murphy, & Strecher, 2007). In addition, participants' characteristics might affect a component's effectiveness. For instance, my meta-analysis identified emotion regulation as a potentially ineffective component, as it was associated with weaker intervention effects. Based on this finding one would suggest to eliminate this component from R&W. However, R&W improved students' emotional self-regulation, a desired intervention effect, making it counterintuitive to eliminate the teaching of emotion regulation from the program. While teaching students how to regulate their emotions might not be an effective component for the total student population, it could perhaps be effective for prevocational students given that this group of students probably requires a specific intervention approach. Initially, I aimed to take heterogeneity in the student population into account in my meta-analysis by examining whether the association between components and intervention effects was dependent on students' educational level. Unfortunately, this was not possible as students' educational level was often not reported in the papers. Thus, also when studying the effectiveness of components, it is crucial to examine under what circumstances and for whom the component is effective (Bonell, Fletcher, Morton, Lorenc, & Moore, 2012). Again, these findings highlight the complexity of generalizing findings from specific circumstances and populations to a larger overall conclusion (Rowe & Trickett, 2018).

Strengths and Limitations

The findings should be regarded in light of some general strengths and limitations of this dissertation. A strength of this dissertation is the use of two approaches to evaluate universal school-based interventions, that is, an RCT and a meta-analysis. On top of this dual approach, I also identified intervention components related to intervention effects of universal school-based interventions. I did this indirectly in the RCT by examining the involvement of the entire school staff (i.e., Standard condition) and of parents (i.e., Plus condition), representing the structural components of, respectively, a whole school approach and parental involvement. In the meta-analysis, I examined intervention components explicitly by analyzing which components were related to intervention effects. This elaborate approach enabled me to move beyond the evaluation of one specific universal school-based intervention and put the findings in a broader framework of this type of interventions.

Another strength is that I assessed intervention fidelity, which is essential information for drawing conclusions about the potential value of an intervention (Durlak & DuPre, 2008). In the RCT all aspects of implementation, as proposed by Durlak and DuPre (2008), were measured: Fidelity, dosage, quality, participants'

engagement, adjustments to the program, monitoring change in a control group, and program reach. Based on both R&W trainers' self-reports and observations by R&W experts, the intervention seemed implemented with moderate to high fidelity to the intervention manual indicating that in general all aspects of the intervention were administered.

A limitation concerns the extent to which the findings regarding R&W can be generalized. I evaluated the intervention only in Dutch schools and only for prevocational students. Due to differences between countries and student populations, one should be careful with generalizing the intervention effects beyond this specific context and population (Bonell et al., 2012). Thus, more research is needed to replicate my findings in different settings to determine the generalizability of the present results.

Furthermore, notwithstanding the strong research design of an RCT and the comparison of the intervention conditions to CAU, i.e., an active control group, the RCT in this dissertation had some limitations. One school dropped-out after randomization but before data collection started and was therefore replaced by another school. This slightly decreased the true RCT design of the study since the new school was placed in a condition rather than randomly assigned. Furthermore, allocation to conditions was conducted at school level, but the data were analyzed on the level of individual students. This is common practice in research examining school-based interventions (e.g., Bonell et al., 2018; Cross et al., 2016) as the interest is in change in individuals rather than in schools. I accounted for clustering at school level in the analyses to prevent underestimation of standard errors and exaggeration of statistically significant findings (Muthén & Muthén, 2017). An exception to these analyses are the analyses conducted for examining modeling and reinforcement as mediators. I analyzed these mechanisms at classroom level rather than the level of individual students since the research questions concerned changes in modeling and reinforcement in the classroom, hence, cluster-level variability (Preacher, Zyphur, & Zhang, 2010).

Future research

The present dissertation sheds light on the effectiveness of universal school-based interventions, in particular R&W, and raises new questions about intervention effects for vulnerable students, about working mechanisms, and about effectiveness of separate components. Can the beneficial effects for vulnerable students, found in the present study, be generalized to other subgroups of vulnerable students? Given that classmates' modeling and reinforcement were not mechanisms of change, what are other possible mechanisms of change? Are the components associated with stronger, and weaker, intervention effects indeed effective, or ineffective, components? And can school-based interventions be optimized by adding or deleting components? By

focusing on these questions future research can further deepen our understanding of universal school-based interventions and possibly increase their effectiveness.

Students who are vulnerable to develop problems are an important subgroup to study, especially in universal interventions. Universal interventions specifically aim to prevent the development of problems (Nehmy & Wade, 2014). As vulnerable students have an increased risk of developing problems, it is eminent that particularly those students are able to benefit from intervention. The present study found stronger intervention effects for students identified as vulnerable based on their levels of personality traits. However, it remains unclear to what extent these advantageous effects can be generalized to other subgroups of vulnerable students (Rowe & Trickett, 2018), such as students with elevated levels of problems at baseline or victimized students (e.g., Kaufman, Kretschmer, Huitsing, & Veenstra, 2018; Multisite Violence Prevention Project, 2008). In order to determine whether universal interventions are indeed able to particularly target the more vulnerable students, it is pivotal to examine intervention effectiveness for different subgroups of vulnerable students.

Furthermore, unraveling mechanisms of change is eminent in order to link intervention techniques with students' development (Michie et al., 2009), even when the intervention effects are small. In the present study, however, the mediators that were tested (i.e., modeling and reinforcement) did not appear to be mechanisms of change within R&W. In my future research I tend to examine self-control, self-reflection, self-esteem, and emotional self-regulation as mechanisms of change. The theory of the R&W house (Ykema 2002; 2018) states that R&W improves these competencies in students which, in turn, would enable these students to develop themselves in the intra- and interpersonal domains, suggesting mediation. The results of the present study imply that R&W is indeed able to improve students' self-control and emotional self-regulation, and showed promising (but not significant) changes in self-esteem (but not in students' self-reflection). Furthermore, these competencies have been related to competencies and problems in the intra- and interpersonal domain by previous research (e.g., Hughes et al., 2020; Mann, Hosman, Schaalma, & De Vries, 2004; Otten, Barker, Maughan, Arsenaeault, & Engels, 2010), making also the relation between the mediators and outcomes plausible. Hence, analyzing these potential mechanisms of change could contribute to understanding how interventions in general can foster students' development in both the intra- and interpersonal domain and, at the same time, examines the theory of the R&W house.

Although universal school-based interventions show positive effects on students' competencies and problems in the intra- and interpersonal domain, these effects are small, indicating the necessity to improve the effectiveness of these interventions. The results of my meta-analysis suggests that there is potential to improve these interventions as some commonly used components might be ineffective and other potentially effective components are rarely implemented. However, the relations between components and intervention effects in the meta-analysis are based on correlations. Future research needs to study these components explicitly to determine

which components are indeed (in)effective, under what circumstances and for whom (Bonell et al., 2012). Subsequently, existing interventions can be improved by adding effective and eliminating ineffective components (Michie et al., 2009), or new interventions can be designed based on the evidence-base for components (e.g., Weisz et al., 2012).

In addition, research should not only study how intervention components relate to immediate intervention effects, but also to long-term effects. Long-term intervention effects indicate that the intervention successfully and meaningfully changed students' development (Flay et al., 2005). However, many interventions do not succeed in changing one's development which is illustrated by extinction of intervention effects over time (e.g., Dray et al., 2017). To improve interventions' long-term effectiveness, it is pivotal to unravel which components contribute to long-term intervention effects and which components counteract these effects.

Practical Implications

Research. Effects of intervention are often studied either through evaluating complete intervention programs or through examining intervention components. These two approaches reflect two relatively separated research areas in which the first is a more traditional approach and the second an upcoming research approach (Collins et al., 2007). Both research approaches have their advantages and disadvantages. I recommend future research to combine these two research areas to maximize the advantages and minimize the disadvantages of the separate approaches.

The traditional approach (evaluating complete intervention programs) treats each intervention as a whole program with the main focus on "whether it works" (Collins et al., 2007). An RCT is considered the "golden standard" to evaluate interventions and seems indeed a solid design to test intervention effects. Interventions are developed, tested, adjusted and tested again in a new RCT (Collins et al., 2007). This research approach has led to high quality research examining sometimes interventions that have poor theory, have poorly understood effects, and are difficult to translate to other contexts (Bonell et al., 2012). As RCTs are time consuming and expensive, this process of intervention optimization is slow and inefficient. Research is focused on a specific intervention and the question remains to what extent the intervention effects can be generalized (Rowe & Trickett, 2018). At this moment, intervention databases contain many potentially effective interventions that still need to be evaluated by means of an RCT (e.g., 201 of the 233 interventions in the Database Effective Youth Interventions; Dutch Youth Institute (NJi), 2020). On top of this, new interventions are designed on a daily basis. As a result, multiple research teams are evaluating interventions with different levels of similarity without coordination between the teams (Bonell et al., 2012). To avoid waste of time and energy, it is therefore important that future research

only puts effort in evaluating those interventions that have a solid theoretical basis regarding the working mechanisms.

In contrast, the upcoming research approach (examining intervention components) treats interventions as packages build from components that can be studied in isolation, with the main focus on "how does it work" (Collins et al., 2007). Potentially meaningful components are identified based on developmental theories or mechanisms of change. Subsequently, the effectiveness of these components are examined in an experimental design, for instance by means of a factorial experiment (Collins, Dziak, Kugler, & Trail, 2014). The focus on theory and working mechanisms in such a research approach can facilitate generalization and adaptation of intervention to other contexts. A crucial requirement to enable generalization and adaptation is an accurate and detailed description of components. Future research can only build on previous research when components are described using consistent terminology and enough information (Michie et al., 2009). When components are identified as effective, they can be used to alter existing interventions or design new interventions. However, these altered or new interventions are not necessarily improved or effective. Combining intervention components might influence their effectiveness, for instance due to interplay among components or the sequence of the individual components (Collins et al., 2007). Therefore, it is essential that research also evaluates the intervention effects of these altered or newly designed interventions in a real-world setting.

I argue that intervention research can benefit from combining the more traditional research approach with the component research approach. The RCT conducted for this dissertation showed a conservative example of how the two approaches can be combined. By evaluating the intervention implemented with various levels of ecological focus, I indirectly examined the two structural components of a whole school approach and parental involvement. However, I would recommend future research to link the two approaches more explicitly and rigorously, for instance using the design proposed by Collins and colleagues (2007) consisting of three stages. First, the screening phase, relevant intervention components are identified and their effectiveness tested. Second, the refining phase, possible interactions between components and characteristics that might affect effectiveness are explored. Third, confirming phase, an intervention is build based on the evidence-base for components acquired in the previous two stages and evaluated. Hence, combining the two research approaches provides the opportunity to build interventions based on solid knowledge about working mechanisms (i.e., components) allowing for dissemination to different contexts while maintaining the necessary characteristics of the intervention (Bonell et al., 2012; Michie et al., 2009).

Schools and professionals. With the current dissertation I hope to contribute to increasing awareness of school staff and professionals that intervention effects might differ between contexts (e.g., intervention dosage), participants, in particular students (e.g., personality traits), and developmental domains (e.g., intrapersonal vs.

interpersonal domain). Interventions are often somewhat adjusted and implemented in different contexts or with different types of participants than for which the interventions were intended and evaluated (Durlak & DuPre, 2008). Given that the extent to which intervention effects can be generalized is often limited (Rowe & Trickett, 2018), schools and professionals are advised to monitor participants' change, for instance, by asking participants, teachers, or parents on a regular basis to complete a questionnaire. Monitoring progress enables schools and professionals to evaluate whether the intervention works in that specific context, population, and domain, or whether a different approach needs to be adopted (Flay et al., 2005).

Conclusion

Overall, universal school-based interventions showed small positive effects in enhancing students' competencies and preventing the development of problems in the intrapersonal and interpersonal domain. More specifically, R&W seems to hold some promise to target prevocational students. Given that the effectiveness of an intervention depends on characteristics of the context (e.g., intervention dosage) as well as of students (e.g., personality traits), small intervention effects can be expected in universal school-based interventions which are often implemented in a broad context and in a broad population. Nevertheless, it remains important to strive towards optimization of universal interventions. My meta-analysis indicated opportunities to improve these interventions as components related to stronger intervention effects were not necessarily often implemented and some commonly implemented components were associated with weaker intervention effects. Taken together, the present dissertation showed that overall universal school-based interventions are somewhat effective in fostering students' development in the intra- and interpersonal domains and indicated opportunities to further improve the effectiveness of these interventions.

References

A

Abidin, R. R. (1983). *Parenting Stress Index: Manual.* Charlottesville: Pediatric Psychology Press.

*Adler-Baeder, F., Kerpelman, J. L., Schramm, D. G., Higginbotham, B., & Paulk, A. (2007). The impact of relationship education on adolescents of diverse backgrounds. *Family Relations, 56,* 291-303,

*Allara, E., Beccaria, F., Molinar, R., Marinaro, L., Ermacora, A., Coppo, A., & Faggiano, F. (2018). A school-based program to promote well-being in preadolescents: Results from a cluster quasi-experimental controlled study. *The Journal of Primary Prevention, 40,* 151-170.

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.

Ando, M., Asakura, T., & Simons-Morton, B. (2005). Psychosocial influences on physical, verbal, and indirect bullying among Japanese early adolescents. *Journal of Early Adolescence, 25,* 268-297.

*Ando, M., Asakura, T., Ando, S., & Simons-Morton, B. (2007). A psychoeducational program to prevent aggressive behavior among Japanese early adolescents. *Health Education & Behavior, 34,* 765-776.

Asscher, J. J., Deković, M., Van den Akker, A. L., Manders, W. A., Prins, P. J. M., Van der Laan, P. H., & Prinzie, P. (2016). Do personality traits affect responsiveness of juvenile delinquents to treatment? *Journal of Research in Personality, 63,* 44-50.

Assink, M., & Wibbelink, C. J. M. (2016). Fitting three-level meta-analytic models in R: A step-by-step tutorial. *The Quantitative Methods for Psychology, 12,* 154-174.

*Avery-Leaf, S., Cascardi, M., O'Leary, K. D., & Cano, A. (1997). Efficacy of a dating violence prevention program on attitudes justifying aggression. *Journal of Adolescent Health, 21,* 11–17.

B

*Baker, C. K., Naai, R., Mitchell, J., & Trecker, C. (2014). Utilizing a train-the-trainer model for sexual violence prevention: Findings from a pilot study with high school students of Asian and Pacific Islander descent in Hawai'i. *Asian American Journal of Psychology, 5,* 106-115.

Bakermans-Kranenburg, M. J., Van IJzendoorn, M. H., & Juffer, F. (2003). Less is more: Meta-analyses of sensitivity and attachment interventions in early childhood. *Psychological Bulletin, 129,* 195-215.

Bandura, A. (1977). *Social learning theory.* Englewood Cliffs, NJ: Prentice-Hall.

Barber, B. K. (2005). Positive interpersonal and intrapersonal functioning: An assessment of measures among adolescents. In K. A. Moore, & L. H. Lippman (Eds.), *What do children need to flourish: Conceptualizing and measuring indicators of positive development* (pp. 147-161). New York, NY: Springer Science and Business Media.

*Barkoukis, V., Lazuras, L., Ourda, D., & Tsorbatzoudis, H. (2016). Tackling psychosocial risk factors for adolescent cyberbullying: Evidence from a school-based intervention. *Aggressive Behavior, 42,* 114-122.

Barry, C. M., & Wentzel, K. R. (2006). Friend influence on prosocial behavior: The role of motivational factors and friendship characteristics. *Developmental Psychology, 42,* 153-163.

Beauchaine, T. P., & Slep, A. (2018). Mechanisms of child behavior change in parent training : Comment on Weeland et al. (2018). *Development and Psychopathology, 30,* 1529-1534.

Berndt, T. J. (2002). Friendship quality and social development. *Current Directions in Psychological Science, 11,* 7-10.

Bishop, D. C., Pankratz, M. M., Hansen, W. B., Albritton, J., Albritton, L., & Strack, J. (2014). Measuring fidelity and adaptation: Reliability of a instrument for school-based prevention programs. *Issues and Solutions in Risky Behavior Prevention Research, 37,* 231-257.

Bokhorst, C. L., Sumter, S. R., & Westenberg, P. M. (2010). Social support from parents, friends, classmates, and teachers in children and adolescents aged 9 to 18 years: Who is perceived as most supportive? *Social Development, 19,* 417-426.

*Bonell, C., Allen, E., Warren, E., McGowan, J., Bevilacqua, L., Jamal, F., … Viner, R. M. (2018). Effects of the Learning Together intervention on bullying and aggression in English secondary schools (INCLUSIVE): A cluster randomised controlled trial. *Lancet, 392,* 2452-2464.

Bonell, C., Fletcher, A., Morton, M., Lorenc, T., & Moore, L. (2012). Realist randomized controlled trials: A new approach to evaluating complex public health interventions. *Social Science and Medicine, 75,* 2299-2306.

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.

*Bosworth, K., Espelage, D., DuBay, T., Daytner, G., & Karageorge, K. (2000). Preliminary evaluation of a multimedia violence prevention program for adolescents. *American Journal of Health Behavior, 24,* 268-280.

*Boulton, M. J., & Flemington, I. (1996). The effects of a short video intervention on secondary school pupils' involvement in definitions of and attitudes towards bullying. *School Psychology, 17,* 331-345.

Boustani, M. M., Frazier, S. T., Ecker, K. D., Bechor, M., Dinizulu, S. M., Hedemann, E. R., …, Pasalich D. S. (2015). Common elements of adolescent prevention programs: Minimizing burden while maximizing reach. *Administration and Policy in Mental Health, 42,* 209-219.

*Bradley, R. T., McCraty, R., Atkinson, M., Tomasino, D., Daugherty, A., & Arguelles, L. (2010). Emotion self-regulation, psychophysiological coherence, and test anxiety: results from an experiment using electrophysiological measures. *Applied Psychophysiology and Biofeedback, 35,* 261–283.

Brodersen, N. H., Steptoe, A., Boniface, D. R., & Wardle, J. (2007). Trends in physical activity and sedentary behaviour in adolescence: Ethnic and socioeconomic differences. *British Journal of Sports Medicine, 41,* 140-144.

Bronfenbrenner, U. (1979). *The ecology of human development.* Cambridge: Harvard University Press.

Bryan, L. A., Dersch, C., Shumway, S., & Arrendondo, R. (2004). Therapy outcomes: Client perception and similarity with therapist view. *The American Journal of Family Therapy, 32,* 11-26.

*Bull, H. D., Schultze, M., & Scheithauer, H. (2009). School-based prevention of bullying and relational aggression: The fairplayer.manual. *European Journal of Developmental Science, 3,* 312-317.

*Burckhardt, R., Manicavasagar, V., Batterham, P. J., & Hadzi-Pavlovic, D. (2016). A randomized controlled trial of strong minds: A school-based mental health program combining acceptance and commitment therapy and positive psychology. *Journal of School Psychology, 57, 41-52.*

*Burckhardt, R., Manicavasagar, V., Batterham, P. J., Hadzi-Pavlovic, D., & Shand, F. (2017). Acceptance and commitment therapy universal prevention program for adolescents: A feasibility study. *Child and Adolescent Psychiatry and Mental Health, 11,* 27-36.

*Burckhardt, R., Manicavasagar, V., Shaw, F., Fogarty, A., Batterham, P. J., Dobinson, K., & Karpin, I. (2018). Preventing mental health symptoms in adolescents using dialectical behaviour therapy skills group: A feasibility study. *International Journal of Adolescence and Youth, 23,* 70–85.

Busching, R., & Krahé, B. (2020). With a little help from their peers: The impact of classmates on adolescents' development of prosocial behavior. *Journal of Youth and Adolescence.* doi:10.1007/s10964-020-01260-8.

C

*Calear, A. L., Christensen, H., Mackinnon, A., Griffiths, K. M., & O'Kearney, R. (2009). The YouthMood Project: A cluster randomized controlled trial of an online cognitive behavioral program with adolescents. *Journal of Consulting and Clinical Psychology, 77,* 1021-1032.

Campbell-Sills, L., & Stein, M. B. (2007). Psychometric analysis and refinement of the Connor-Davidson Resilience Scale (CD-RISC): Validation of a 10-item measure of resilience. *Journal of Traumatic Stress, 20(6),* 1019-1028.

*Caplan, M., Weissberg, R. P., Grober, J. S., Sivo, P. J., Grady, K, & Jacoby, C. (1992). Social competence promotion with inner-city and suburban young adolescents: Effects on social adjustment and alcohol use. *Journal of Consulting and Clinical Psychology, 60,* 56-63.

*Caprara, G. V., Kanacri, B. P. L., Gerbino, M., Zuffiano, A., Alessandri, G., Vecchio, G., … Bridglall, B. (2014). Positive effects of promoting prosocial behavior in early adolescence: Evidence from a school-based intervention. *International Journal of Behavioral Development, 38,* 386-396.

*Carraro, A., Gobbi, E., & Moè, A. (2014). Brief report: Play fighting to curb self-reported aggression in young adolescents. *Journal of Adolescence, 37,* 1303–1307.

*Castillo-Gualda, R., Cabello, R., Herrero, M., Rodriguez-Carvajal, R., & Fernandez-Berrocal, P. (2017). A three-year emotional intelligence intervention to reduce adolescent aggression: The mediating role of unpleasant affectivity. *Journal of Research on Adolescence, 28,* 186-198.

Central Bureau for Statistics (CBS) (2017). *VO; leerlingen, onderwijssoort in detail, leerjaar* [Secondary education; students, education level in detail, schoolyear]. Retrieve from http://statline.cbs.nl/statweb/publication/?vw=t&dm=slnl&pa=80040ned&d1=0,3,6-8,14-15,19-21,55,57-59,61-64,69-70,72-74,76-79,84&d2=0&d3=0&d4=0-1,4&d5=l&d6=0&d7=7,(l-2)-l&hd=160216-1208&hdr=g4,g5,g1,g2,g3,g6&stb=t

Central Bureau for Statistics (CBS) (2020). *VO; leerlingen, onderwijssoort in detail, leerjaar.* [Secondary education; students, educational track in detail, schoolyear.] Retrieved from https://opendata.cbs.nl/statline/#/CBS/nl/dataset/80040ned/table?fromstatweb Accessed 10 June 2020.

*Challen, A. R., Machin, S. J., & Gillham, J. E. (2014). The UK Resilience Programme: A school-based universal nonrandomized pragmatic controlled trial. *Journal of Consulting and Clinical Psychology, 82,* 75-89.

*Chang, C., Tsai, G., & Hsieh, C. (2013). Psychological, immunological and physiological effects of a Laughing Qigong Program (LQP) on adolescents. *Complementary Therapies in Medicine, 21,* 660–668.

*Cheung, C., & Lee T. (2010). Improving social competence through character education. *Evaluation and Program Planning, 33,* 255-263.

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.

*Coelho, V. A., Marchante, M., & Sousa, V. (2015). "Positive attitude": A multilevel model analysis of the effectiveness of a social and emotional learning program for Portuguese middle school students. *Journal of Adolescence, 43,* 29-38.

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Hillsdale, NJ: Erlbaum.

*Coker, A. L., Bush, H. M., Cook-Craig, P. G., DeGue, S. A., Clear, E. R., Brancato, C. J., …, Recktenwald, E. A. (2017). RCT testing bystander effectiveness to reduce violence. *American Journal of Preventive Medicine, 52,* 566-578.

Collins, L. M., Dziak, J. J., Kugler, K. C., Trail, J. B. (2014). Factorial experiments. Efficient tools for evaluation of intervention components. *American Journal of Preventive Medicine, 47,* 498-504.

Collins, L. M., Murphy, S. A., & Strecher, V. (2007). The Multiphase Optimization Strategy (MOST) and the Sequential Multiple Assignment Randomized Trial (SMART). New methods for more potent eHealth interventions. *American Journal of Preventive Medicine, 32,* S112-S118.

*Connolly, J., Josephson, W., Schnoll, J., Simkins-Strong, E., Pepler, D., MacPherson, A., … Jiang, D. (2015). Evaluation of a youth-led program for preventing bullying, sexual harassment, and dating aggression in middle schools. *Journal of Early Adolescence, 35,* 403-434.

Connor-Smith, J. K., & Flachsbart, C. (2007). Relations between personality and coping: A meta-analysis. *Journal of Personality and Social Psychology, 93,* 1080-1107.

*Constantine, N. A., Jerman, P., Berglas, N. F., Angulo-Olaiz, F., Chou, C, & Rohrbach, L. A. (2015). Short-term effects of a rights-based sexuality education curriculum for high-school students: A cluster-randomized trial. *BMC Public Health, 15,* 293-305.

Crocetti, E. (2016). Systematic reviews with meta-analysis: Why, when, and how? *Emerging Adulthood, 4,* 3-18.

*Cross, D., Shaw, T., Hadwen, K., Cardoso, P., Slee, P., Roberts, C., … Barnes, A. (2016). Longitudinal impact of the Cyber Friendly Schools program on adolescents' cyberbullying behavior. *Aggressive Behavior, 42,* 166-180.

Cuijpers, P. (2002). Effective ingredients of school-based drug prevention programs. A systematic review. *Addictive Behaviors, 27,* 1009-1023.

D

*Daly, L. A., Haden, S. C., Hagins, M., Papouchis, N., & Ramirez, P. M. (2015). Yoga and emotion regulation in high school students: A randomized controlled trial. *Evidence-Based Complementary and Alternative Medicine, 2015,* 1-8.

Davidson, J. R. T., & Connor, K. M. (2017). *Connor-Davidson Resilience Scale (CD-RISC) manual.* Unpublished and partly accessible at www.cd-risc.com.

De Bruijn, D., Overmaat, M., Glaudé, M., Heemskerk, I., Leeman, Y., Roeleveld, J., & Van de Venne, L. (2005). Krachtige leeromgevingen in het middelbaar beroepsonderwijs: Vormgeving en effecten [Powerful learning environments in secondary vocational education: Organization and effects]. *Pedagogische Studiën, 82,* 77-95.

De Graaf, H., Meijer, S., Poelman, J., & Vanwesenbeeck, I. (2005). *Seks onder je 25e* [Sex under the age of 25]. Delft: Uitgeverij Eburon.

*De Graaf, I., De Haas, S., Zaagsma, M., & Wijsen, C. (2016). Effects of Rock and Water: An intervention to prevent sexual aggression. *Journal of Sexual Aggression, 22,* 4-19.

De Mooij, B., 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.

De Roos, S., & Sevens, G. (2018). Schoolbeleving [Perception of school]. In M. Boer, & G. Stevens (Eds.), *HBSC 2017: Gezondheid, welzijn en opvoeding van jongeren in Nederland [Health, wellbeing, and child rearing of youth in The Netherlands],* (pp. 43-61). Ridderkerk: Ridderprint BV.

De Vibe, M., Solhaug, I., Tyseen, R., Friborg, O., Rosenvinge, J. H., Sørlie, T., …, Bjørndal, A. (2015). Does personality moderate the effects of mindfulness training for medical and psychology students? *Mindfulness, 6,* 281-289.

*De Villiers, M., & Van den Berg, H. (2012). The implementation and evaluation of a resiliency programme for children. *South African Journal of Psychology, 42,* 93-102.

De Vries, S. L. A., Hoeve, M., Assink, M., Stams, G. J. J. M., & Asscher, J. J. (2015). Practitioner review: Effective ingredients of prevention programs for youth at risk of persistent juvenile delinquency – recommendations for clinical practice. *Journal of Child Psychology and Psychiatry, 56,* 108-121.

Deković, M., Asscher, J. J., Hermanns, J., Reitz, E., Prinzie, P., & Van den Akker, A. L. (2010). Tracing changes in families who participated in the Home-Start Parenting Program: Parental sense of competence as mechanism of change. *Prevention Science, 11,* 263-274.

Deković, M., Asscher, J. J., Manders, W. A., Prins, P. J. M., & Van der Laan, P. (2012). Within-intervention change: Mediators of intervention effects during Multisystemic Therapy. *Journal of Consulting and Clinical Psychology, 80,* 574-587.

DeNeve, K. M., & Cooper, H. (1998). The happy personality: A meta-analysis of 137 personality traits and subjective well-being. *Psychological Bulletin, 124,* 197-229.

Dishion, T. J. Spracklen, K. M., Andrews, D. W., & Patterson, G. R. (1996). Deviancy training in male adolescent friendships. *Behavior Therapy, 27,* 373-390.

Dishion, T. J., & Tipsord, J. M. (2011). Peer contagion in child and adolescent social and emotional development. *Annual Review of Psychology, 62,* 189-214.

Dishion, T. J., Andrews, D. W., & Crosby, L. (1995). Antisocial boys and their friends in early adolescence: Relationship characteristics, quality, and interactional process. *Child Development, 66,* 139-151.

Dishion, T. J., Crosby, L., Rusby, J. C., Shane, D., Patterson, G. R., & Baker, J. (1989). *Peer Process Code: Multidimensional system for observing adolescent-peer interaction.* (Available from the Oregon Social Learning Center, 160 E. 4th Ave., Eugene, OR 97401)

Dishion, T. J., Spracklen, K. M., Andrews, D. W., & Patterson, G. R. (1996). Deviancy training in male adolescent friendships. *Behavior Therapy, 27,* 373-390.

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.

Doehne, M., Von Grundherr, M., & Schäfer, M. (2018). Peer influence in bullying: The autonomy-enhancing effect of moral competence. *Aggressive Behavior, 44,* 591-600.

*Domino, M. (2013). Measuring the impact of an alternative approach to school bullying. *The Journal of School Health, 83,* 430-437.

Dray, J., Bowman, J., Campbell, E., Freund, M., Wolfenden, L., Hodder, R. K., McElwaine, K., …, Wiggers, J. (2017). Systematic review of universal resilience-focused interventions targeting child and adolescent mental health in the school setting. *Journal of the American Academy of Child and Adolescent Psychiatry, 56,* 813-824.

Duckworth, A. L., & Seligman, M. E. P. (2005). Self-discipline outdoes IQ in predicting academic performance of adolescents. *Psychological Science, 16,* 939-944.

Dufner, M., Gebauer, J. E., Sedikides, C., & Denissen, J. J. A. (2019). Self-enhancement and psychological adjustment: A meta-analytic review. *Personality and Social Psychology Review, 23,* 48-72.

Duncan, T. E., & Duncan, S. C. (2004). An introduction to Latent Growth Curve modeling. *Behavior Therapy, 35,* 333-363.

*DuRant, R. H., Treiber, F., Getts, A., McCloud, K., Linder, C. W., & Woods, E. R. (1996). Comparison of two violence prevention curricula for middle school adolescents. *Journal of Adolescent Health, 19,* 111-117.

Durlak, J. A. (2009). How to select, calculate, and interpret effect sizes. *Journal of Pediatric Psychology, 34,* 917-928.

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,* 327-350.

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.

Dutch Youth Institute (NJi) (2020). *Databank Effectieve Jeugdinterventies.* [Database Effective Youth interventions. Acknowledge interventions.] Retrieved from https://www.nji.nl/nl/Databank/Databank-Effectieve-Jeugdinterventies/Erkende-interventies Accessed 19 June 2020.

Duval, S., & Tweedie, R. (2000a). A nonparametric 'trim and fill' method of accounting for publication bias in meta-analysis. *Journal of the American Statistical Association, 95,* 89-99.

Duval, S., & Tweedie, R. (2000b). Trim and Fill: A simple funnel-plot based method of testing and adjusting for publication bias in meta-analysis. *Biometrics, 56,* 455-460.

E

Ebert, D. D., Gollwitzer, M., Riper, H., Cuijpers, P., Baumeister, H., & Berking, M. (2013). For whom does it work? Moderators of outcome on the effect of a transdiagnostic internet-based maintenance treatment after inpatient psychotherapy: Randomized Controlled Trial. *Journal of Medical Internet Research, 15,* doi:www.jmir.org/2013/10/e191

Egger, M., Smith, G. D., Schneider, M., & Minder, C. (1997). Bias in meta-analysis detected by a simple, graphical test. *BMJ, 315,* 629-634.

*Espelage, D. L., Low, S., Polanin, J. R., & Brown, E. C. (2013). The impact of a middle school program to reduce aggression, victimization, and sexual violence. *Journal of Adolescent Health, 53,* 180-186.

F

Farrell, A. D., Henry, D. B., Bettencourt, A. (2013). Methodological challenges examining subgroup differences: Examples from universal school-based youth violence prevention trials. *Prevention Science, 14,* 121-133.

Feingold, A. (2013). A regression framework for effect size assessments in longitudinal modeling of group differences. *Review of General Psychology, 17,* 111-121.

Felder, R. M., Felder, G. N., & Dietz, E. J. (2002). The effects of personality type on engineering student performance and attitudes. *Journal of Engineering Education, 91,* 3-17.

*Felver, J. C., Clawson, A. J., Morton, M. L., Brier-Kennedy, E., Janack, P., & DiFlorio, R. A. (2018). School-based mindfulness intervention supports adolescent resiliency: A randomized controlled pilot study. *International Journal of School & Educational Psychology.* doi:10.1080/21683603.2018.1461722

Ferguson, C. J., & Brannick, M. T. (2012). Publication bias in psychological science: Prevalence, methods for identifying and controlling, and implications for the use of meta-analyses. *Psychological Methods, 17,* 120-128.

*Filella, G., Ros-Morente, A., Oriol, X., & March-Llanes, J. (2018). The assertive resolution of conflicts in school with a gamified emotion education program. *Frontiers in Psychology, 9,* 2353.

Finkel, E. J., & Vohs, K. D. (2006). Introduction. Self and Relationships. In K. D. Vohs, & E. J. Finkel (Eds.), *Self and relationships: Connecting intrapersonal and interpersonal processes* (pp. 1-9). New York, NY: The Guilford Press.

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., …, Kainbacher, M. (2011). The bystander-effect: A meta-analytic review on bystander intervention in dangerous and non-dangerous emergencies. *Psychological Bulletin, 137,* 517-537.

Flay, B. R., Biglan, A., Boruch, R. F., Castro, F. G., Gottfredson, D., Kellam, S., Mościcki, E. K., …, Ji, P. (2005). Standards of Evidence: Criteria for efficacy, effectiveness and dissemination. *Prevention Science, 6,* 151-175.

Flay, B. R., Graumlich, S., Segawa, E., Burns, J. L., & Holliday, M. Y. (2004). Effects of 2 prevention programs on high-risk behaviors among African American youth. *Archives of Pediatric Adolescent Medicine, 158,* 377-384.

*Foshee, V. A., Bauman, K. E., Ennett, S. T., Suchindran, C., Benefield, T., & Linder, G. F. (2005). Assessing the effects of the dating violence prevention program "Safe dates" using random coefficient regression modeling. *Prevention Science, 6,* 245-258.

*Frank, J. L., Kohler, K., Peal, A., & Bose, B. (2016). Effectiveness of a school-based yoga program on adolescent mental health and school performance: Findings from a randomized controlled trial. *Mindfulness, 8,* 544-553.

*Freire, T., Lima, I., Teixeira, A., Araujo, M. R., & Machado, A. (2018). Challenge: To be+. A group intervention program to promote the positive development of adolescents. *Children and Youth Services Review, 87,* 173-185.

G

*Garaigordobil (2014). Effect of Cyberprogram 2.0 on reducing victimization and improving social competence in adolescence. *Revista de Psicodidáctica, 19,* 289-305.

*Garaigordobil, M. (2004). Effects of a psychological intervention on factors of emotional development during adolescence. *European Journal of Psychological Assessment, 20,* 66-80.

*Garaigordobil, M., & Martinez-Valderrey, V. (2015a). Effects of Cyberprogram 2.0 on "face-to-face" bullying, cyberbullying, and empathy. *Psicothema, 27,* 45–51.

*Garaigordobil, M., & Martinez-Valderrey, V. (2015b). The effectiveness of Cyberprogram 2.0 on conflict resolution strategies and self-esteem. *Journal of Adolescent Health, 57,* 229-234.

*Gardner, S. P., Giese, K., & Parrott, S. M. (2004). Evaluation of the Connections: Relationships and Marriage curriculum. *Family Relations, 53,* 521-527.

*Ghahremani, D. G., Oh, E. Y., Dean, A. C., Mouzakis, K., Wilson, K. D., & London, E. D. (2013). Effects of the Youth Empowerment Seminar on impulsive behavior in adolescents. *Journal of Adolescent Health, 53,* 139-141

*Giannotta, F., Settanni, M., Kliewer, W., & Ciairano, S. (2009). Results of an Italian school-based expressive writing intervention trial focused on peer problems. *Journal of Adolescence, 32,* 1377-1389.

*Gigantesco, A., Del Re, D., Cascavilla, I., Palumbo, G., De Mei, B., Cattaneo, C., … Bella, A. (2015). A universal mental health promotion programme for young people in Italy. *BioMed Research International, 2015,* 1-10.

Gille, E., Loijens, C., Noijons, J., & Zwitser, R. (2010). Leesvaardigheid. [Reading skills] *Resultaten PISA-2009. Praktische kennis en vaardigheden van 15-jarigen. [Results PISA-2009. Practical knowledge and skills of 15-year olds.]* (pp. 37-70). Arnhem: Cito.

Gillham, J. E., Reivich, K. J., Reres, D. R., Chaplin, T. M., Shatté, A. J., Samuels, B., Elkon, A. G. L., …, & Seligman, M. E. P. (2007). School-based prevention of depressive symptoms: A randomized controlled study of the effectiveness and specificity of the Penn Resiliency Program. *Journal of Consulting and Clinical Psychology, 75,* 9-19.

Gilman, R., & Anderman, E. M. (2006). The relationship between relative levels of motivation and intrapersonal, interpersonal, and academic functioning among older adolescents. *Journal of School Psychology, 44,* 375-391.

Goldberg, L. R. (1992). The development of markers for the Big-Five structure. *Psychological Assessment, 4,* 26-42.

Goldstein, S. E., Young, A., & Boyd, C. (2008). Relational aggression at school: Associations with school safety and social climate. *Journal of Youth and Adolescence, 37,* 641-654.

*Gollwitzer, M., Banse, R., Eisenbach, K., & Naumann, A. (2007). Effectiveness of the Vienna Social Competence Training on explicit and implicit aggression: Evidence from an Aggressiveness-IAT. *European Journal of Psychological Assessment, 23,* 150-156.

Gottfredson, D. C., & Gottfredson, G. D. (2002). Quality of school-based prevention programs: Results from a national survey. *Journal of Research in Crime and Delinquency, 39,* 3-35.

*Gouda, S., Luong, M. T., Schmidt, S., & Bauer, J. (2016). Students and teachers benefit from mindfulness-based stress reduction in a school-embedded pilot study. *Frontiers in Psychology, 7,* 590-607.

Greenberg, M. T., & Abenavoli, R. (2017). Universal interventions: Fully exploring their impacts and potential to produce population-level impacts. *Journal of Research on Educational Effectiveness, 10,* 40-67

H

*Hains, A. A, & Ellmann, S. W. (1994). Stress inoculation training as a preventative intervention for high school youths. *Journal of Cognitive Psychotherapy: An International Quarterly, 8,* 219–232.

*Hains, A. A. (1994). The effectiveness of a school-based, cognitive-behavioral stress management program with adolescents reporting high and low levels of emotional arousal. *School Counselor, 42,* 114-125.

*Hains, A. A., & Szyjakowski, M. (1990). A cognitive stress-reduction intervention program for adolescents. *Journal of Counseling Psychology, 37,* 79-84.

Harakeh, Z., De looze, M. E., Schrijvers, C. T. M., Van Dorsselaer, S. A. F. M., & Vollebergh, W. A. M. (2012). Individual and environmental predictors of health risk behaviours among Dutch adolescents: The HBSC study. *Public Health, 126,* 566-573.

Harter, S. (1988). *Manual for the Self-Perception Profile for Adolescents*. Denver, US: university of Denver.

Haviland, J. M., Davidson, R. B., Ruetsch, C., Gebelt, J. L., & Lancelot, C. (1994). The place of emotion in identity. *Journal of Research on Adolescence, 4,* 503-518.

*Haynes, L. A., & Avery, A. W. (1979). Training adolescents in self-disclosure and empathy skills. *Journal of Counseling Psychology, 26,* 526-530.

Hedges, L. (1983). A random effects model for effect sizes. *Psychological Bulletin, 93,* 388-395.

Hendrickx, M., Crombez, G., Roeyers, H., & Orobio de Castro, B. (2003). Psychometrische evaluatie van de Nederlandstalige versie van de Agressie Beoordelingsschaal van Dodge en Coie (1987) [Psychometric evaluation of the Dutch version of Dodge & Coie's (1987) aggression rating scale]. *Tijdschrift voor Gedragstherapie, 36,* 33-43.

Hennessy, E. A., & Tanner-Smith, E. E. (2015). Effectiveness of brief school-based interventions for adolescents: A meta-analysis of alcohol use prevention program. *Prevention Science, 16,* 463-474.

Higgins, J. P. T, Sterne, J. A. C., Savović, J., Page, M. J. Hróbjartsson, A., Boutron, I., Reeves, B., & Eldridge, S. A. (2016). A revised tool for assessing risk of bias in randomized trials. In J. Chandler, J. McKenzie, I. Boutron, & V. Welch (Eds.), Cochrane Methods. *Cochrane Database of Systematic Reviews, 10 (Suppl 1)*. doi:dx.doi.org/10.1002/14651858.CD201601.

Hinde, R. A. (1987). *Individuals, relationships and culture.* Cambridge: Cambridge University Press.
Hjemdal, O., Vogel, P. A., Solem, S., Hgen, K., & Stiles, T. C. (2011). The relationship between resilience and levels of anxiety, depression, and obsessive-compulsive symptoms in adolescents. *Clinical Psychology and Psychotherapy, 18,* 314-321.
Hofmann, V., & Müller, C. M. (2018). Avoiding antisocial behavior among adolescents: The positive influence of classmates' prosocial behavior. *Journal of Adolescence, 68,* 136-145.
Hopson, L. M., Schiller, K. S., & Lawson, H. A. (2014). Exploring linkages between school climate, behavioral norms, social supports, and academic success. *Social Work Research, 38,* 197-209.
*Horn, A. B., Pössel, P., & Hautzinger, M. (2010). Promoting adaptive emotion regulation and coping in adolescence: A school-based programme. *Journal of Health Psychology, 16,* 258–273.
Horne, A. M., Stoddard, J. L., & Bell, C. D. (2007). Group approaches to reducing aggression and bullying in school. *Group Dynamics: Theory, Research, and Practice, 11,* 262-271.
Huenecke, T. M., & Waas, G. A. (2010). Impact of discussion on peer evaluations: Perceptions of low achievement and effort. *The Journal of Experimental Education, 79,* 64-83.
Hughes, D. J., Kratsiotis, I. K., Niven, K., & Holman, D. (2020). Personality traits and emotion regulation: A targeted review and recommendations. *Emotion, 20,* 63-67.
Hukkelberg, S. S., & Ogden, T. (2013). Working alliance and treatment fidelity as predictors of externalizing problem behaviors in parent management training. *Journal of Consulting and Clinical Psychology, 81,* 1010-1020.
*Huppert, F. A., & Johnson, D. M. (2010). A controlled trial of mindfulness training in schools: The importance of practice for an impact on well-being. *The Journal of Positive Psychology, 5,* 264-274.
Huppert, J. D., Gorman, J. M., Bufka, L. F., Barlow, D. H., Shear, M. K., & Woods, S. W. (2001). Therapists, therapist variables, and Cognitive-Behavioral Therapy outcome in a multicenter trial for panic disorder. *Journal of Consulting and Clinical Psychology, 69,* 747-755.

I

*Ingram, K. M., Espelage, D. L., Merrin, G. J., Valido, A., Heinhorst, J., & Joyce, M. (2019). Evaluation of a virtual reality enhanced bullying prevention curriculum pilot trial. *Journal of Adolescence, 71,* 72-83.

J

*Jaycox, L. H., McCaffrey, D., Eiseman, B., Aronoff, J., Shelley, G. A., Collins, R. L., & Marshall, G. N. (2006). Impact of a school-based dating violence prevention program among Latino teens: Randomized controlled effectiveness trial. *Journal of Adolescent Health, 39,* 694-704.
*Jiménez-Barbero, J. A., Ruiz-Hernández, J. A., Llor-Zaragoza, L., Pérez-Garcia, M., & Llor-Esteban, B. (2016). Effectiveness of anti-bullying school programs: A meta-analysis. *Children and Youth Services Review, 61,* 165-175.
Juvonen, J., & Ho, A. Y. (2008). Social motives underlying antisocial behavior across middle school grades. *Journal of Youth and Adolescence, 37,* 747-756.

K

Kaminski, J. W., Valle, L. A., Filene, J. H., & Boyle, C. L. (2008). A meta-analytic review of components associated with parent training program effectiveness. *Journal of Abnormal Child Psychology, 36,* 567-589.

*Kasler, J., White, G. W., & Elias, M. J. (2013). Evaluation of the Meaning of Life program in Israel. *Peabody Journal of Education, 88,* 243-260.

Kaufman, T. M. L., Kretschmer, T., Huitsing, G., & Veenstra, R. (2018). Why does a universal anti-bullying program not help all children? Explaining persistent victimization during an intervention. *Prevention Science, 19,* 822-832.

*Kaveh, M. H., Hesampour, M., Ghahremani, L., & Tabatabaee, H. R. (2014). The effects of a peer-led training program on female students' self-esteem in public secondary schools in Shiraz. *Journal of Advances in Medical Education & Professionalism, 2,* 63-70.

Kazdin, A. E. (2007). Mediators and mechanisms of change in psychotherapy research. *Annual Review of Clinical Psychology, 3,* 1-27.

Kelly, E. B., Parra-Medina, D., Pfeiffer, K. A., Dowda, M., Conway, T. L., Webber, L. S., Jobe, J. B., …, Pate, R. R. (2010). Correlates of physical acitivity in Black, Hispanic, and White middle school girls. *Journal of Physical Activity and Health, 7,* 184-193.

*Khanna, P., & Singh, K. (2016). Effect of gratitude educational intervention on well-being indicators among North Indian adolescents. *Contemporary School Psychology, 20,* 305-314.

Kirby, K. N., Petry, N. M., & Bickel, W. K. (1999). Heroin addicts have higher discount rates for delayed rewards than non-drug-using controls. *Journal of Experimental Psychology: General, 128,* 78-87.

*Kiselica, M. S., Baker, S. B., Thomas, R. N., & Reedy, S. (1994). Effects of stress inoculation training on anxiety, stress, and academic performance among adolescents. *Journal of Counseling Psychology, 41,* 335-342.

Kline, R. B. (2005). *Principles and practice of structural equation modeling*. New York: Guilford Press.

Kline, R. B. (2015). Coming of Age. *Principles and practice of Structural Equation Modeling* (4th edition) (pp. 7-24). New York: The Guilford Press.

*Klingman, A., & Hochdorf, Z. (1993). Coping with distress and self harm: The impact of a primary prevention program among adolescents. *Journal of Adolescence, 16,* 121-140.

Kobak, K. A., Wolitzky-Taylor, K., Craske, M. G., & Rose, R. D. (2017). Therapist training on Cognitive Behavior Therapy for anxiety disorders using internet-based technologies. *Cognitive Therapy and Research*, 41, 252-265.

Koopman, M., Den Brok, P., Beijaard, D., & Teune, P. (2011). Learning processes of students in pre-vocational secondary education: Relations between goal orientations, information processing strategies and development of conceptual knowledge. *Learning and Individual Differences, 21,* 426-431.

*Kozina, A. (2018a). Can the "my friends" anxiety prevention programme also be used to prevent aggression? A six-month follow-up in a school. *School Mental Health, 10,* 500-509.

*Kozina, A. (2018b). School-based prevention of anxiety using the "My FRIENDS" emotional resilience program: Six-month follow-up. *International Journal of Psychology.* doi:10.1002/ijop.12553

Kuyper, H., Keuning, J., & Zijsling, D. (2010). *Basisrapport eerste meting van COOL5–18 in het derde leerjaar van het voortgezet onderwijs* [A report on the first wave of the COOL5–18]. Groningen: GION/Cito.

L

*Lamke, L. K., Lujan, B. M., & Showalter, J. M. (1988). The case for modifying adolescents' cognitive self-statements. *Adolescence, 23,* 967-974.

Landis, J. R., & Koch, G. G. (1977). The measurement of observer agreement for categorical data. *Biometrics, 33,* 159-174.

Langford, R., Bonell, C. P., Jones, H. E., Pouliou, T., Murphy, S. M., Waters, E., …, Campbell, R. (2014). The WHO Health Promoting School framework for improving the health and well-being of students and their academic achievement. *The Cochrane Databased of Systematic Reviews, 4.* doi:10.1002/14651858.CD008958.pub2

Lee, B. R., Ebesutani, C., Kolivoski, K. M., Becker, K. D., Lindsey, M. A. Brandt, N. E., … Chorpita, B.F. (2014). Program and practice elements for placement prevention: A review of interventions and their effectiveness in promoting home-based care. *American Journal of Orthopsychiatry, 84,* 244-256.

Liber, J. M., De Boo, G. M., Huizenga, H., & Prins, P. J. M. (2013). School-based intervention for childhood disruptive behavior in disadvantaged settings: A randomized controlled trial with and without active teacher support. *Journal of Consulting and Clinical Psychology, 81,* 975-987.

Lipsey, M. W. & Wilson, D. B. (2001). *Practical meta-analysis*. Thousand Oaks, CA: Sage Publications.

Longabaugh, R., & Magill, M. (2011). Recent advances in behavioral addition treatments: Focusing on mechanisms of change. *Current Psychiatry Report, 13,* 382-389.

Lutz, W., Hofmann, S. G., Rubel, J., Boswell, J. F., Shear, M. K., Gorman, J. M., … , & 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.

M

*Macgowan, M. J. (1997). An evaluation of a dating violence prevention program for middle school students. *Violence and Victims, 12,* 223–235.

Mann, M., Hosman, C. M. H., Schaalma, H. P., & De Vries, N. K. (2004). Self-esteem in a broad-spectrum approach for mental health promotion. *Health Education Research, 19,* 357-372.

*Menesini, E., Codecasa, E., Benelli, B., & Cowie, H. (2003) Enhancing children's responsibility to take action against bullying: Evaluation of a befriending intervention in Italian middle schools. *Aggressive Behavior, 29,* 1-14.

Mertens, E. C. A., Deković, M., Leijten, P., Van Londen, & Reitz, E. (2020). *What to do or not do to stimulate students' intrapersonal and interpersonal domains: Meta-analysis of school-based intervention components*. Manuscript submitted for publication.

Mertens, E. C. A., Deković, M., Van Londen, M., & Reitz, E. (2018). The effectiveness of Rock and Water in improving students' socio-emotional adjustment and social safety: Study protocol for a randomized controlled trial. *BMC Psychology, 6,* 1-12.

Mertens, E. C. A., Deković, M., Van Londen, M., Nye, E., & Reitz, E. (2020). *Solid as a rock, flexible as water? Effectiveness of a school-based intervention addressing students' intrapersonal and interpersonal domains.* Manuscript submitted for publication.

Michie, S., Fixsen, D., Grimshaw, J. M., & Eccles, M. P. (2009). Specifying and reporting complex behavior change interventions: The need for a scientific method. *Implementation Science, 4,* 40-45.

*Miller, S., Williams, J., Cutbush, S., Gibbs, D., Clinton-Sherrod, M., & Jones, S. (2015). Evaluation of the Start Strong initiative: Preventing teen dating violence and promoting healthy relationships among middle school students. *Journal of Adolescent Health, 56,* S14-S19.

Minges, K. E., Chao, A., Nam, S., Grey, M., & Whittemore, R. (2015). Weight status, gender, and race/ethnicity: Are there differences in meeting recommended health behavior guidelines for adolescents? *The Journal of School Nursing, 31,* 135-145.

Modecki, K. L., Zimmer-Gembeck, M. J., & Guerrà, N. (2017). Emotion regulation, coping, and decision making: Three linked skills for preventing externalizing problems in adolescence. *Child Development, 88,* 417-426.

*Muck, C., Schiller, E., Zimmermann, M., & Kartner, J. (2018). Preventing sexual violence in adolescence: Comparison of a scientist-practitioner program and a practitioner program using a cluster-randomized design. *Journal of Interpersonal Violence.* doi:10.1177/0886260518755488

Multisite Violence Prevention Project (2008). The Multisite Violence Prevention Project : Impact of a universal school-based violence prevention program on social-cognitive outcomes. *Prevention Science, 9,* 231-244.

Muthén, B., & Satorra, A. (1995). Complex sample data in Structural Equation Modeling. *Sociological Methodology, 25,* 267-316.

Muthén, L., K., & Muthén, B. O. (1998-2017). *Mplus user's guide.* 8th Edition. Los Angeles, CA: Muthén & Muthén.

Mychailyszyn, M. P., Brodman, D. M., Read, K. L., & Kendall, P. C. (2012). Cognitive-behavioral school-based interventions for anxious and depressed youth: A meta-analysis of outcomes. *Clinical Psychology, Science and Practice, 19,* 129-153.

Myerson, J., Baumann, A. A., & Green, L. (2014). Discounting of delayed rewards: (A) theoretical interpretation of the Kirby Questionnaire. *Behavioural Processes, 107,* 99-105.

Myerson, J., Baumann, A. A., & Green, L. (2017). Individual differences in delay discounting: Differences are quantitative with gains, but qualitative with losses. *Journal of Behavioral Decision Making, 30,* 359-372.

N

*Nash. K. A. (2007). Implementation and evaluation of the empower youth program. *Journal of Holistic Nursing, 25,* 26–36.

Nehmy, T. J., & Wade, T. D. (2014). Reduction in the prospective incidence of adolescent psychopathology: A review of school-based prevention approaches. *Mental Health and Prevention, 2,* 66-79.

*Noggle, J. J., Steiner, N. J., Minami, T., & Khalsa, S. B. S. (2012). Benefits of yoga for psychosocial well-being in a US high school curriculum: A preliminary randomized controlled trial. *Journal of Developmental and Behavioral Pediatrics, 33,* 193-201.

O

*O'Hearn, T. C., & Gatz, M. (1999). Evaluating a psychosocial competence program for urban adolescents. *The Journal of Primary Prevention, 20,* 119-144.

*O'Hearn, T., & Gatz, M. (2002). Going for the Goal: Improving youths' problem-solving skills through a school-based intervention. *Journal of Community Psychology, 30,* 281-303.

O'Mara, A. J., Marsh, H. W., Craven, R. G., & Debus, R. L. (2006). Do self-concept interventions make a difference? A synergistic blend of construct validation and meta-analysis. *Educational Psychologist, 41,* 181-206.

O'Rourke, H. P., & MacKinnon, D. P. (2018). Reasons for testing mediation in the absence of an intervention effect: A research imperative in prevention and intervention research. *Journal of Studies on Alcohol and Drugs, 79,* 171-181.

Onrust, S. A., Otten, R., Lammers, J., & Smit, F. (2016). School-based programmes to reduce and prevent substance use in different age groups: What works for whom? Systematic review and meta-regression analysis. *Clinical Psychology Review, 44,* 45-59.

*Orpinas, P., Parcel, G. S., McAlister, A., & Frankowski, R. (1995). Violence prevention in middle schools: A pilot evaluation. *Journal of Adolescent Health, 17,* 360-371.

*Ortega-Baron, J., Buelga, S., Ayllon, E., Martinez-Ferrer, B., & Cava, M. (2019). Effects of intervention program Prev@cib on traditional bullying and cyberbullying. *International Journal of Environmental Research and Public Health, 16,* 527-539.

Otten, R., Barker, E. D., Maughan, B., Arseneault, L., & Engels, R. C. M. E. (2010). Self-control and its relation to joint developmental trajectories of cannabis use and depressive mood symptoms. *Drug and Alcohol Dependence, 112,* 201-208.

P

*Pacifici, C., Stoolmiller, M., & Nelson, C. (2001). Evaluating a prevention program for teenagers on sexual coercion: A differential effectiveness approach. *Journal of Consulting and Clinical Psychology, 69,* 552-559.

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, Washington, DC.

Peters, L. W. H., Kok, G., Ten Dam, G. T. M., Buijs, G. J., & Paulussen, T. G. W. M. (2009). Effective elements of school health promotion across behavioral domains: A systematic review of reviews. *BMC Public Health, 9,* 182-196.

Philips, B., & Wennberg, P. (2014). The importance of therapy motivation for patients with substance use disorders. *Psychotherapy, 51,* 555-562.

Phillips Smith, E., Dumas, J. E., & Prinz, R. (2006). Prevention approaches to improve child and adolescent behavior and reduce deviant peer influence. In K. A. Dodge, T. J. Dishion, & J. Lansford (Eds.), *Deviant peer influences in programs for youth: Problems and solutions,* (pp. 296-311). New York: The Guilford Press.

Piehler, T. F., & Dishion, T. J. (2004a). *Peer dyadic mutuality rating system.* Unpublished coding manual (available from the Child and Family Center, University of Oregon).

Piehler, T. F., & Dishion, T. J. (2004b). *The Conversation topic code.* Unpublished coding manual (available from the Child and Family Center, University of Oregon).

Piehler, T. F., & Dishion, T. J. (2007). Interpersonal dynamics within adolescent friendships: Dyadic mutuality, deviant talk, and patterns of antisocial behavior. *Child Development, 78,* 1611-1624.

Pinquart, M. (2013). Do the parent-child relationship and parenting behaviors differ between families with a child with and without chronic illness? A meta-analysis. *Journal of Pediatric Psychology, 38,* 708-721.

Poulin, F., Dishion, T. J., & Burraston, B. (2001). 3-Year iatrogenic effects associated with aggregating high-risk adolescents win cognitive-behavioral preventive interventions. *Applied Developmental Science, 5,* 214-224.

Preacher, K. J., Zyphur, M. J., & Zhang, Z. (2010). A general multilevel SEM framework for assessing multilevel mediation. *Psychological Methods, 15,* 209-233.

*Proctor, C., Tsukayama, E., Wood, A. M., Maltby, J., Fox Eades, J., Linley, P. A. (2011) Strengths Gym: The impact of a character strengths-based intervention on the life satisfaction and well-being of adolescents. *The Journal of Positive Psychology, 6,* 377-388.

R

Ravens-Sieberer, U., & The European KIDSCREEN Group. (2006). *The KIDSCCREEN questionnaires. Quality of life questionnaires for children and adolescents – Handbook*. Lengerich: Pabst Science publisher.

*Richardson, S. M., Paxton, S. J., & Thomson, J. S. (2009). Is BodyThink an efficacious body image and self-esteem program? A controlled evaluation with adolescents. *Body Image, 6,* 75–82.

Rowe, H. L., & Trickett, E. J. (2018). Student diversity representation and reporting in universal school-based Social and Emotional Learning programs: Implications for generalizability. *Educational Psychology Review, 30,* 559-583.

Rubin, K. H., Bukowski, W. M., & Parker, J. G. (2006). Peer interactions, relationships, and groups. In N. Eisenberg (Ed.), *Handbook of child psychology: Volume III. Social, emotional, and personality development*. (6th edition, pp. 619-700). New York: Wiley.

*Ruini, C., Belaise, C., Brombin, C., Caffo, E., & Fava, G. A. (2006). Well-being therapy in school cettings: A pilot study. *Psychotherapy and Psychosomatics, 75,* 331-336.

*Ruini, C., Ottolini, F., Tomba, E., Belaise, C., Albieri, E., Visani, D., … Fava, G. A.(2009). School intervention for promoting psychological well-being in adolescence. *Journal of Behavior Therapy and Experimental Psychiatry, 40,* 522-532.

*Ruiz-Aranda, D., Castillo, R., Salguero, J. M., Cabello, R., Fernández-Berrocal, P., & Balluerka, N. (2012). Short- and midterm effects of emotional intelligence training on adolescent mental health. *Journal of Adolescent Health, 51,* 462-467.

S

Salmivalli, C., & Isaacs, J. (2005). Prospective relations among victimization, rejection, friendlessness, and children's self- and peer-perceptions. *Child Development, 76,* 1161-1171.

*Sanchez-Jimenez, V., Munoz-Fernandez, N., & Ortega-Rivera, J. (2018). Efficacy evaluation of "Dat-e Adolescence": A dating violence prevention program in Spain. *PloS One, 13,* doi:10.1371/journal.pone.0205802.

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., Van Widenfelt, B. M., & 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.

*Schramm, D. G., & Gomez-Scott, J. (2012). Merging relationship education and child abuse prevention knowledge: An evaluation of effectiveness with adolescents. *Marriage and Family Review, 48,* 792-808.

Schrijvers, C. T. M., & Schuit, A. J. (2010). *Middelengebruik en seksueel gedrag van jongeren met een laag opleidingsniveau* [Substance use and sexual behavior of youth with low educational levels]. Bilthoven: Rijksinstituut voor Volksgezondheid en Milieu (RIVM).

*Schultz, L. H., Barr, D. J., & Selman, R. L. (2001). The value of a developmental approach to evaluating character development programmes: An outcome study of Facing History and Ourselves. *Journal of Moral Education, 30,* 3-27.

Schum, L., & Stolberg, A. L. (2007). Standardization of the co-parenting behavior questionnaire. *Journal of Divorce and Remarriage, 47,* 103-132.

Senf, K., & Liau, A. K. (2013). The effects of positive interventions on happiness and depressive symptoms, with an examination of personality as a moderator. *Journal of Happiness Studies, 14,* 591-612.

Shek, D. T. L., & Leung, J. T. Y. (2016). Developing social competence in a subject on leadership and intrapersonal development. *International Journal on Disability and Human Development, 51,* 165-173.

*Shek. D. T. L., & Ma, C. M. S. (2011). Impact of the Project P.A.T.H.S. in the junior secondary school years: Individual growth curve analyses. *The Scientific World Journal, 11,* 253–266.

*Shinde, S., Weiss, H. A., Varghese, B., Khandeparkar, P., Pereira, B., Sharma, A., … Patel, V. (2018). Promoting school climate and health outcomes with the SEHER multi-component secondary school intervention in Bihar, India: a cluster-randomised controlled trial. *Lancet, 392,* 2465-2477.

Shiner, R., & Caspi, A. (2003). Personality differences in childhood and adolescence: Measurement, development, and consequences. *Journal of Child Psychology and Psychiatry, 44,* 2-32.

*Shoshani, A., & Steinmetz, S. (2014). Positive psychology at school: A school-based intervention to promote adolescents' mental health and well-being. *Journal of Happiness Studies, 15,* 1289-1311.

*Sibinga, E. M. S., Perry-Parrish, C., Chung, S., Johnson, S. B., Smith, M., & Ellen, J. M. (2013). School-based mindfulness instruction for urban male youth: A small randomized controlled trial. *Preventive Medicine, 57,* 799-801.

Sijtsema, J. J., & Lindenberg, S. M. (2018). Peer influence in the development of adolescent antisocial behavior: Advances from dynamic social network studies. *Developmental Review, 50,* 140-154.

*Simons-Morton, B., Haynie, D., Saylor, K., Crump, A. D., & Chen, R. (2005). The effects of the Going Places Program on early adolescent substance use and antisocial behavior. *Prevention Science, 6,* 187-197.

Solberg, M. E., & Olweus, D. (2003). Prevalence estimation of school bullying with the Olweus Bully/Victim Questionnaire. *Aggressive Behavior, 29,* 239-268.

*Soliday, E., Garofalo, J. P., & Rogers, D. (2004). Expressive writing intervention for adolescents' somatic symptoms and mood. *Journal of Clinical Child and Adolescent Psychology, 33,* 792-801.

*Solomontos-Kountouri, O., Gradinger, P., Yanagida, T., & Strohmeier, D. (2016). The implementation and evaluation of the ViSC program in Cyprus: Challenges of cross-national dissemination and evaluation results. *European Journal of Developmental Psychology, 13,* 737-755.

*Sorrentino, A., Baldry, A. C., & Farrington, D. (2018). The efficacy of the Tabby Improved prevention and intervention program in reducing cyberbullying and cybervictimization among students. *International Journal of Environmental Research and Public Health, 15,* 2536-2549

Spoth, R., Shin, C., Guyll, M., Redmond, C., & Azevedo, K. (2006). Universality of effects: An examination of the comparability of long-term family intervention effects on substance use across risk-related subgroups. *Prevention Science, 7,* 209-224.

Stevens, G., & De Looze, M. (2018). Welbevinden en psychische problemen [Wellbeing and psychological problems]. In M. Boer, & G. Stevens (Eds.), *HBSC 2017: Gezondheid, welzijn en opvoeding van jongeren in Nederland [Health, wellbeing, and child rearing of youth in the Netherlands],* (pp. 63-75). Ridderkerk: Ridderprint BV.

*Stevens, V., Bourdeaudhuij, I., & Oost, P. (2000). Bullying in Flemish schools: An evaluation of anti-bullying intervention in primary and secondary schools. *British Journal of Educational Psychology, 70,* 195-210.

Stoltz, S., Prinzie, P., De Haan, A., Van Londen, M., Orobio de Castro, B., & Deković, M. (2013). Child personality as moderator of outcome in a school-based intervention for preventing externalizing behavior. *European Journal of Personality, 27,* 271-279.

T

Tackett, J. L. (2006). Evaluating models of the personality-psychopathology relationship in children and adolescents. *Clinical Psychology Review, 26,* 584-599.

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.

Taylor, B. G., Mumford, E. A., Liu, W., & Stein, N. D. (2017). The effects of different saturation levels of the Shifting Boundaries intervention on preventing adolescent relationship abuse and sexual harassment. *Journal of Experimental Criminology, 13,* 79-100.

Taylor, Z. E., Conger, R. D., Robins, R. W., & Widaman, K. F. (2015). Parenting practices and perceived social support: Longitudinal relations with the social competence of Mexican-origin children. *Journal of Laina/o Psychology, 3,* 193-208.

Telzer, E. H., Van Hoorn, J., Rogers, C. R., & Do, K. T. (2018). Social influence on positive youth development: A developmental neuroscience perspective. *Advances in Child Development and Behavior, 54,* 215-258.

Ter Vrugte, J., De Jong, T., Wouters, P., Vandercruysse, S., Elen, J., & Van Oostendorp, H. (2015). When a game supports prevocational math education but integrated reflection does not. *Journal of Computer Assisted Learning, 31,* 462-480.

*Thomaes, S., Bushman, B. J., Orobio de Castro, B., Cohen, G. L., & Denissen, J. J. A. (2009). Reducing narcissistic aggression by buttressing self-esteem: An experimental field study. *Psychological Science, 20,* 1536-1542.

*Thompkins, A. C., Chauveron, L. M., Harel, O., & Perkins, D. F. (2014). Optimizing violence prevention programs: an examination of program effectiveness among urban high school students. *The Journal of School Health, 84,* 435-443.

Timmermans, A., Naaijer, H., Keuning, J., & Zijsling, D. (2017). *Cohortonderzoek COOL5-18 basisrapport meting VO-3 in 2014* [Cohort study COOL5-18 report measurement VO-3 in 2014]. Groningen: GION onderwijs/onderzoek.

*Tomyn, J. D., Fuller-Tyszkiewicz, M., Richardson, B., & Colla, L. (2016). A comprehensive evaluation of a universal school-based depression prevention program for adolescents. *Journal of Abnormal Child Psychology, 44,* 1621-1633.

*Trip, S., Bora, C., Sipos-Gug, S., Tocai, I., Gradinger, P., Yanagida, T., & Strohmeier, D. (2015). Bullying prevention in schools by targeting cognitions, emotions, and behavior: Evaluating the effectiveness of the REBE-ViSC program. *Journal of Counseling Psychology, 62,* 732-740.

Ttofi, M. M., & Farrington, D. P. (2011). Effectiveness of school-based programs to reduce bullying: A systematic and meta-analytic review. *Journal of Experimental Criminology, 7,* 27-56.

*Tunariu, A. D., Tribe, R., Frings, D., & Albery, I. P. (2017). The iNEAR programme: An existential positive psychology intervention for resilience and emotional wellbeing. *International Review of Psychiatry, 29,* 362–372.

V

Van de Bongardt, D., Reitz, E. Overbeek, G., Boislard, M., Burk, B., & Deković (2017). Observed normativity and deviance in friendship dyads' conversation about sex and the relations with youths' perceived sexual peer norms. *Archives of Sexual Behavior, 46,* 1793-1806.

Van den Noortgate, W., López-López, J. A., Marín'Martínez, F., & Sánchez-Meca, J. (2013). Three-level meta-analysis of dependent effect sizes. *Behavior Research Methods, 45,* 576-594.

*Van der Meulen, K., Granizo, L., & del Barrio, C. (2010). Using EQUIP for educators to prevent peer victimization in secondary school. *Journal of Research in Character Education, 8,* 61-76.

Van der Put, C. E., Assink, M., Gubbels, J., & Boekhout van Solinge, N. F. (2018). Identifying effective components of child maltreatment interventions: A meta-analysis. *Clinical Child and Family Psychology Review, 21,* 171-202.

Van Leeuwen, K. G., Mervielde, I., Braet, C., & Bosmans, G. (2004). Child personality and parental behavior as moderators of problem behavior: Variable- and person-centered approaches. *Developmental Psychology, 40,* 1028-1046.

Van Ness, P. H., Murphy, T. E., Araujo, K. L. B., Pisani, M. A., & Allore, H. G. (2007). The use of missingness screens in clinical epidemiologic research has implications for regression modeling. *Journal of Clinical Epidemiology, 60,* 1239-1245.

*Van Schoiack-Edstrom, L., Frey, K. S., & Beland, K. (2002). Changing adolescents' attitudes about relational and physical aggression: An early evaluation of a school-based intervention. *School Psychology Review, 31,* 201-216.

Verdurmen, J. E. E., Koning, I. M., Vollebergh, W. A. M., Van den Eijnden, R. J. J. M., & Engels, R. C. M. E. (2014). Risk moderation of a parent and student preventive alcohol intervention by adolescent and family factors: A cluster randomized trial. *Preventive Medicine, 60,* 88-94.

Verhulst, F. C., Van der Ende, J., & Koot, H. M. (1996). *Handleiding voor de CBCL/4-18*. Rotterdam: Sophia kinderziekenhuis, Erasmus MC.

Verhulst, F. C., Van der Ende, J., & Koot, H. M. (1997a). *Handleiding voor de Teacher's Report Form (TRF)*. Rotterdam: Afdeling Kinder- en Jeugdpsychiatrie, Sophia kinderziekenhuis / Erasmus MC.

Verhulst, F. C., Van der Ende, J., & Koot, H. M. (1997b). *Handleiding voor de Youth Self-Report (YSR)*. Rotterdam: Sophia kinderziekenhuis / Erasmus MC.

Viechtbauer, W. (2010). Conducting meta-analyses in R with the metafor package. *Journal of Statistical Software, 36,* 1-48.

W

Wampold, B. E., & Brown, G. S. (2005). Estimating variability in outcomes attributable to therapists: A naturalistic study of outcomes in managed care. *Journal of Consulting and Clinical Psychology, 73,* 914-923.

Wang, R. A. H., Nelson-Coffey, S. K., Layous, K., Jacobs Bao, K., Davis, O. S. P., & Haworth, C. M. A. (2017). Moderators of wellbeing interventions: Why do some people respond more positively than others? *Plos One, 12,* Doi:10.1371/journal.pone.0187601

Weeland, J., Chhangur, R. R., Jaffee, S. R., Van der Giessen, D., Matthys, W., Orobio de Castro, B., & Overbeek, G. (2018). *Development and Psychopathology, 30,* 1534-1540.

Weisz, J. R., Chorpita, B. F., Palinkas, F. A., Schoenwald, S. K., Miranda, J., Bearman, S. K., …, Research Network on Youth Mental Health (2012). Testing standard and modular designs for psychotherapy treating depression, anxiety, and conduct problems in youth. *Archives of General Psychiatry, 69,* 274-282.

Weisz, J. R., Chorpita, B. F., Palinkas, L. A., Schoenwald, S. K., Miranda, J., Bearman, S. K., Daleiden, E. L., …, & Research Network on Youth Mental Health (2008). Testing standard and modular designs for psychotherapy treating depression, anxiety, and conduct problems in youth. A randomized effectiveness trial. *Archives of General Psychiatry, 69,* 274-282.

Werch, C. E., & Owen, D. M. (2002). Iatrogenic effects of alcohol and drug prevention programs. *Journal of Studies on Alcohol and Drugs, 63,* 581-590.

Whalen, C. K., Henker, B., Collings, B. E., McAuliffe, S., & Vaux, A. (1979). Peer interaction in a structured communication task: Comparisons of normal and hyperactive boys and of methylphenidate (Ritalin) and placebo effect. *Child Development, 50,* 388-401.

Wheeler, S., & Richards, K. (2007). The impact of clinical supervision on counsellors and therapists, their practice and their clients. A systematic review of the literature. *Counselling and Psychotherapy Research, 7,* 54-65.

White, B. A., Jarrett, M. A., & Ollendick, T. H. (2013). Self-regulation deficits explain the link between reactive aggression and internalizing and externalizing behavior problems in children. *Journal of Psychopathology and Behavioral Assessment, 35,* 1-9.

*Williams, J., Miller, S., Cutbush, S., Gibbs, D., Clinton-Sherrod, M., & Jones, S. (2015). A latent transition model of the effects of a teen dating violence prevention initiative. *The Journal of Adolescent Health, 56,* S27-S32.

*Williford, A., Elledge, L. C., Boulton, A. J., DePaolis, K. J., Little, T. D., & Salmivalli, C. (2013). Effects of the KiVa antibullying program on cyberbullying and cybervictimization frequency among Finnish youth. *Journal of Clinical Child and Adolescent Psychology, 42,* 820-833.

Wilson, S. J., & Lipsey, M. W. (2007). School-based interventions for aggressive and disruptive behavior. Update of a meta-analysis. *American Journal of Preventive Medicine, 33,* S130-S143.

Wolters, C. A., & Daugherty, S. G. (2007). Goal structures and teachers' sense of efficacy: Their relation and association to teaching experience and academic level. *Journal of Educational Psychology, 99,* 181-193.

*Wong, D. S. W., Cheng, C. H. K., Ngan, R. M. H., & Ma, S. K. (2011). Program effectiveness of a Restorative Whole-school Approach for tackling school bullying in Hong Kong. *International Journal of Offender Therapy and Comparative Criminology, 55,* 846-862.

World Health Organization (1995). Global School Health Initiative. Retrieved from https://www.who.int/school_youth_health/gshi/en/

World Health Organization (1997). Life skills education in schools. Geneva: WHO/MNH/PSF/93.7A. Rev. 2.

Y

*Yanagida, T., Strohmeier, D., & Spiel, C. (2019). Dynamic change of aggressive behavior and victimization among adolescents: Effectiveness of the ViSC Program. *Journal of Clinical Child and Adolescent Psychology, 48,* S90–S104.

Yeager, D. S., & Walton, G. M. (2011). Social-psychological interventions in education: They're not magic. *Review of Educational Research, 81,* 267-301.

Ykema F. (2002). *Het Rots en Water perspectief. Een psychofysieke training voor jongens. Basisboek* [The Rock and Water perspective. A psychophysical training for boys. Basic book]. Amsterdam: Uitgeverij SWP; 2002.

Ykema F. (2014). *Rots en Water praktijkboek. Psychofysieke training voor jongens én meisjes* [Rock and Water practice book. Psychophysical training for boys and girls]. Amsterdam: Uitgeverij SWP.

Ykema F. (2018). *Rots en Water praktijkboek. Psychofysieke training voor jongens én meisjes* [Rock and Water practice book. Psychophysical training for boys and girls]. Amsterdam: 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.

*Yom, Y., & Eun, L. K. (2005). Effects of a CD-ROM educational program on sexual knowledge and attitude. *Computers, Informatics, Nursing, 23,* 214–219.

Yovel, I., Mor, N., & Shakarov, H. (2014). Examination of the core cognitive components of Cognitive Behavioral Therapy and Acceptance and Commitment Therapy: An analogue investigation. *Behavior Therapy, 45,* 482-494.

* Reference included in meta-analysis

Summary

The current dissertation focused on the effectiveness of universal school-based interventions in improving students' competencies and preventing problems in the intrapersonal (i.e., feelings, emotions, and attitudes about the self) and interpersonal (i.e., the ability to build and maintain positive relationships with others, to understand social situations, roles and norms, and to respond appropriately) domain. I examined 1) the effectiveness of a specific intervention, Rock and Water (R&W) through a Randomized Controlled Trial, and of universal school-based interventions in general through a meta-analysis, 2) whether intervention dosage and students' personality traits affected intervention effects, and 3) mechanisms of change and intervention components of universal school-based interventions.

Concerning the first aim, universal school-based interventions showed small positive effects in both the intra- and interpersonal domain. Intervention effects found in R&W were comparable in magnitude to the effects found for these interventions in general, with strongest intervention effects in the intrapersonal domain. Regarding specific competencies, universal school-based interventions, including R&W, appear unable to improve students' resilience even though many of these interventions aim to improve this competency.

With respect to the second aim, examining intervention dosage showed that R&W was only effective when few teachers were involved. Additionally, students improved most during the first part of the intervention. Hence, the first year might be sufficient to establish change in the assessed competencies and problems. Together these findings imply that for some interventions "less is more". Regarding students' personality, the effectiveness of R&W was only little affected by students' personality traits confirming the universal nature of the intervention. However, three moderation patterns did emerge. First, more vulnerable students – based on their levels of certain personality traits – seemed to benefit somewhat more from the intervention than less vulnerable students. Second, extraverted students, but not introverted students, were able to benefit from R&W when many people were involved. Third, personality traits seemed more influential in the intrapersonal domain than in the interpersonal domain. These findings emphasize the importance of examining characteristics of contexts and of participants for determining the extent to which intervention effects can be generalized.

Regarding the third aim, analyzing mechanisms of change showed that classmates' deviant and prosocial modeling and reinforcement did not mediate the effect of R&W on the peer context in the classroom. However, increasing prosocial modeling was related to decreasing victimization, indicating a potential working mechanism. Concerning intervention components, the meta-analysis showed that some components related to stronger effect sizes were rarely implemented (e.g., Individual guidance, Relaxation), whereas some commonly implemented components

were associated with weaker effect sizes (e.g., Emotion regulation, Discussions). This finding suggests opportunities to improve intervention effectiveness. However, optimizing interventions is complex as components' effectiveness could depend on characteristics of the context and of the population.

In sum, universal school-based interventions generally show small positive effects in the intra- and interpersonal domain, which is to be expected. These intervention effects appear dependent on intervention dosage and students' personality traits. Nevertheless, the small intervention effects indicate that there is potential to improve the effectiveness of universal school-based interventions.

Samenvatting

In dit proefschrift onderzocht ik de effectiviteit van universele schoolinterventies, die erop gericht zijn de bekwaamheid (*skills*) van leerlingen op het voorgezet onderwijs te vergroten en de belemmering (*struggles*) van deze leerlingen te verminderen in zowel het intrapersoonlijk als het interpersoonlijk domein. Het intrapersoonlijk domein betreft het persoonlijk functioneren (*Hoe zie je jezelf?*) en verwijst naar de subjectieve verwerking van gedrag, gedachten, en emoties van een individu (bijv. psychologisch welzijn, weerbaarheid). Het interpersoonlijk domein betreft het sociaal functioneren (*Hoe ga je met anderen om?*) en omvat het opbouwen en onderhouden van positieve relaties met anderen, het begrijpen van sociale situaties (rollen en normen) en het adequaat reageren (bijv. positieve relaties in de klas, geen agressie).

Bekwaamheid en belemmering in beide domeinen zijn belangrijke voorspellers van later psychosociaal welzijn. Het is daarom belangrijk de ontwikkeling van de leerlingen in deze domeinen te stimuleren. Juist de schoolomgeving is hiervoor geschikt, omdat jongeren veel tijd doorbrengen op school. Bovendien is de school, naast de academische ontwikkeling van leerlingen, ook verantwoordelijk voor de stimulering van persoonlijke attituden en waarden, en het sociale netwerk van leerlingen.

Het eerste doel van dit proefschrift was om de effectiviteit van universele schoolinterventies in het intra- en interpersoonlijk domein te onderzoeken (*Wat werkt?*). Het tweede doel was om te bestuderen of verschillen in context en tussen leerlingen invloed hebben op de effectiviteit (*Onder welke omstandigheden en voor wie werkt het?*). Het derde doel was het analyseren van mogelijk werkzame mechanismen (*Hoe werkt het?*).

Doel 1: Wat werkt?

Ik heb zowel de effectiviteit van alle goed onderzochte universele schoolinterventies onderzocht via een meta-analyse (Hoofdstuk 6), alsook de effectiviteit van een specifieke universele schoolinterventie, Rots en Water (R&W) via een gerandomiseerd onderzoek met een controlegroep en drie interventie condities (*Light*: een paar leerkrachten betrokken; *Standaard*: alle leerkrachten betrokken; en *Plus*: alle leerkrachten en ouders betrokken; RCT: Hoofdstuk 3).

Volgens de meta-analyse laten universele interventies vooral in het intrapersoonlijk domein (zelfvertrouwen, zelfregulatie, psychologisch welzijn, internaliserend gedrag), maar ook in het interpersoonlijk domein (seksuele gezondheid, sociale competentie, agressie, en pesten), kleine positieve effecten zien. R&W had vergelijkbare kleine positieve effecten, ten opzichte van het gebruikelijke schoolbeleid (*Controle* conditie), vooral in de *Light* conditie (zie doel 2). Wanneer we specifiek kijken naar de uitkomsten

bleek R&W in het intrapersoonlijk domein het psychologisch welzijn, seksuele autonomie, en internaliserend gedrag iets te verbeteren, en in het interpersoonlijk domein agressief gedrag iets te verminderen. Verder bleek R&W de zelfcontrole en emotieregulatie van leerlingen iets te vergroten.

Opvallend is dat universele schoolinterventies, inclusief R&W, niet in staat lijken om de specifieke intrapersoonlijke competentie *weerbaarheid* van leerlingen te verbeteren, terwijl dat vaak wel het doel is van deze interventies. Concluderend, aangezien de interventie effecten klein zijn, is het belangrijk om te onderzoeken hoe de effectiviteit van deze schoolinterventies vergroot kan worden.

Doel 2: Onder welke omstandigheden en voor wie werkt het?

Voor R&W heb ik gekeken of de *dosering* (de mate van intensiteit) van de interventie samenhing met de effectiviteit (Hoofdstuk 3). Daarnaast heb ik onderzocht of leerlingen met verschillende persoonlijkheidskenmerken anders reageren op de interventie (Hoofdstuk 4).

Met betrekking tot dosering van de interventie heb ik onderzocht of de ecologische focus (aantal mensen betrokken bij de interventie) en de duur van de interventie samenhing met de effectiviteit. De ecologische focus van de interventie heb ik onderzocht aan de hand van de drie interventie condities (*Light, Standaard,* en *Plus* – zie ook doel 1) die verschilden in het aantal mensen dat betrokken was bij R&W.

R&W was vooral effectief wanneer er weinig leerkrachten betrokken waren bij het uitvoeren van de interventie (alleen de leerkrachten die de R&W lessen gaven – *Light* conditie). In deze conditie werd R&W dus toegepast tijdens de interventielessen, maar niet tijdens reguliere lessen (*Standaard* en *Plus* conditie) of in de thuisomgeving (*Plus* conditie). Verder bleek dat leerlingen vooral tijdens het eerste gedeelte van de interventie (in het eerste schooljaar) verbetering lieten zien. Misschien is het voor het vergroten van competenties en het verminderen van problemen voldoende om alleen tijdens het eerste jaar en alleen tijdens de interventielessen (*Light* conditie) R&W te geven. Deze resultaten samen suggereren dat voor R&W mogelijk geldt dat *minder beter is* (“*less is more*”).

Daarnaast heb ik geanalyseerd of sommige leerlingen meer kunnen profiteren van R&W de andere leerlingen afhankelijk van hun persoonlijkheidskenmerken. Sommige persoonlijkheidskenmerken (zoals minder consciëntieus zijn, meer neurotisch, meer extravert) kunnen leerlingen bijvoorbeeld kwetsbaarder maken voor het ontwikkelen van problemen, waardoor ze wellicht ook meer potentie hebben om te verbeteren. Daarnaast zouden bepaalde persoonlijkheidskenmerken ook de overdracht van vaardigheden geleerd tijdens de interventielessen naar het dagelijks leven kunnen bevorderen.

Persoonlijkheidskenmerken van leerlingen bleken relatief weinig samen te hangen met de effectiviteit van R&W. Dit bevestigt het universele karakter van

R&W waarbij het de bedoeling is dat alle leerlingen evenveel kunnen profiteren van de interventie. Desalniettemin kwamen er drie patronen naar voren waarin persoonlijkheidskenmerken wel gerelateerd waren aan interventie effecten. Ten eerste, leerlingen die op basis van bepaalde persoonlijkheidskenmerken een verhoogde kans hadden om problemen te ontwikkelen (kwetsbare leerlingen) leken meer te profiteren van R&W dan minder kwetsbare leerlingen. Ten tweede, extraverte, maar niet introverte, leerlingen leken te kunnen profiteren van de interventie wanneer meer mensen erbij betrokken waren (alle leerkrachten en ouders – *Standaard* en *Plus* condities). Ten derde, het intrapersoonlijk domein bleek meer gerelateerd te zijn aan persoonlijkheidskenmerken dan het interpersoonlijk domein.

Het is dus belangrijk om eigenschappen van de omgeving en van de leerlingen te onderzoeken tijdens het evalueren van schoolinterventies. Alleen dan kan bepaald worden in hoeverre interventie effecten gegeneraliseerd kunnen worden naar een andere context of een andere doelgroep. Daarnaast is het belangrijk dat professionals die een interventie toepassen, zoals leerkrachten, zich ervan bewust zijn dat verschillende leerlingen anders kunnen reageren op dezelfde interventie.

Doel 3: Hoe werkt het?

Ik heb op twee manieren gekeken naar werkzame mechanismen van interventies. Voor specifiek R&W heb ik in de RCT via een observatiestudie de rol van klasgenoten onderzocht (Hoofdstuk 5). Voor interventies in het algemeen heb ik in de meta-analyse de rol van interventie componenten geanalyseerd (Hoofdstuk 6).

In de observatiestudie heb ik tweetallen van klasgenoten geobserveerd waarbij ik me heb gericht op modeling, bekrachtiging en de interactie tussen de tweetallen (dyadische wederkerigheid) en de relatie met de sociale context in de klas. Volgens de sociale leertheorie wordt gedrag beïnvloed via twee mechanismen, namelijk via *modeling* (het aanleren van nieuw gedrag door het observeren van leeftijdsgenoten) en *bekrachtiging* (het aanleren van nieuw gedrag door positieve feedback van leeftijdsgenoten). De invloed van deze twee mechanismen is mogelijk afhankelijk van de mate waarin de *interactie* tussen de klasgenoten positief is.

De resultaten tonen aan dat modeling en bekrachtiging geen werkzame mechanismen van R&W waren. R&W leek namelijk niet het deviante (negatief gedrag) en prosociale (positief gedrag) modeling en bekrachtiging van klasgenoten te veranderen. Bovendien waren modeling en bekrachtiging van klasgenoten niet gerelateerd aan de sociale context in de klas. Een uitzondering hierop was prosociale modeling. Een toename in prosociale modeling bleek namelijk gerelateerd te zijn aan een afname in gepest worden. Deze relatie was sterker naar mate de interactie tussen de leerlingen positiever was. Samen suggereren deze bevinding dat het voor anti-pest interventies de moeite waard kan zijn om prosociale modeling en positieve interactie tussen klasgenoten te stimuleren.

In de meta-analyse heb ik onderzocht welke onderdelen van interventies (interventie componenten) gerelateerd waren aan interventie effecten. Uit de meta-analyse van de interventie componenten bleek dat sommige componenten die gerelateerd waren aan sterkere interventie effecten zelden toegepast worden (bijv. individuele begeleiding, relaxatie). Daarentegen waren sommige vaak toegepaste componenten juist gerelateerd aan nadelige interventie effecten (bijv. emotieregulatie, discussies). Deze resultaten suggereren dat er mogelijkheden zijn om de effectiviteit van schoolinterventies te verbeteren. Echter, het verbeteren van interventies is complex omdat de effectiviteit van componenten afhankelijk kan zijn van de combinatie van componenten, van eigenschappen van de context en van eigenschappen van de doelgroep.

Conclusie

Universele schoolinterventies (ook R&W) laten over het algemeen kleine positieve effecten zien in het intra- en interpersoonlijk domein. Bij R&W zijn deze effecten afhankelijk van interventie dosering (ecologische focus en duur van de interventie) en persoonlijkheidskenmerken van leerlingen. De meta-analyse van interventie componenten toont aan dat er potentie is om de beperkte effectiviteit van universele schoolinterventies gericht op het intra- en interpersoonlijke domein te verbeteren.

Dankwoord

Dit proefschrift is tot stand gekomen dankzij de hulp en steun van vele anderen. Hier wil ik hen graag persoonlijk voor bedanken.

Allereerst wil ik mijn begeleiders bedanken voor hun toewijding, tijd, feedback en steun. Maja, jouw kritische vragen en scherpe geest hebben mij keer op keer uitgedaagd om keuzes te heroverwegen en te zoeken naar een alternatieve aanpak. Ellen, jouw nuchtere en zakelijke aanpak heeft me geholpen prioriteiten te stellen in drukke periodes. Monique, jouw praktijkgerichte insteek en ervaring met interventies hebben een extra dimensie toegevoegd aan mijn aanpak van het onderzoek en vertaling naar de praktijk.

I also want to thank my colleagues at the Centre for Evidence-Based Interventions at the University of Oxford for the warm welcome, interesting lectures, and pub nights. Frances, thank you for your time, ideas, and knowledge of meta-analyses. Elizabeth, when working together on the paper, your question "What do we want to schools to know?" always brought me back to the practical relevance of my research when I was too focused on my analyses. I also want to thank the DPhil students at the department for introducing me into Oxford's student life, in particular the life within the colleges. Ana and Janina, you made me feel at home in Oxford, thank you!

Daarnaast wil ik mijn co-auteur Patty bedanken. Jouw enthousiasme en expertise op het gebied van componenten van interventies werken aanstekelijk en zijn inspirerend.

Alle scholen, leerlingen, leerkrachten, ouders en het Rots en Water instituut wil ik bedanken voor hun inzet tijdens dit project. Zonder hen zou er geen data zijn geweest om te analyseren. Ook de studenten, in het bijzonder mijn student assistentes Natasha, Iris, Lissy en Milou, wil ik bedanken voor hun hulp bij de dataverzameling en het coderen van de observaties.

Verder wil ik al mijn collega's van de afdeling Orthopedagogiek Psychosociale Problemen bedanken. Jullie gezelligheid, openheid, en ideeën hebben ervoor gezorgd dat ik altijd met plezier naar mijn werk ben gegaan (en nog steeds ga).

Ook de PhDs wil ik bedanken voor hun ideeën, oplossingen, experimenteel leren in de pantry en bovenal gezelligheid. Rianne en Sanne, ik heb genoten van onze kopjes thee in de ochtend en de gifjes tijdens de digitale maanden aan het einde van mijn PhD. Lilly, mijn bondgenoot en "roomie" vanaf het allereerste begin, jij hebt mijn PhD traject verreikt en nog leuker gemaakt! Bedankt voor je luisterend oor, brainstormen, gezelligheid, en regelen van tafels en muurversiering. Uiteindelijk zijn ook bij mij de eerste grijze haren verschenen.

Tot slot wil ik graag mijn familie bedanken voor hun aanmoediging en steun. Mama, bedankt voor je aanhoudende steun en vertrouwen in mijn capaciteiten. Je hebt me altijd aangemoedigd door te zetten ook als ik zelf even iets minder positief was. Papa, bedankt voor je betrokkenheid, pogingen mijn wetenschappelijke, Engelse

artikelen te lezen, en lunchpakketjes met briefjes. Dankzij jou kon ik de uitgevallen controle school vervangen met een nieuwe school en zijn alle leerkracht vragenlijsten van die school ingevuld. Rob, ik heb genoten van het carpoolen waarbij ik je kon bijpraten over van alles of even bij kon slapen. Jeroen, dankzij jou kon ik op congressen pronken met mijn posters. Hans en Irmgard, bedankt voor jullie oprechte interesse en vragen gedurende mijn gehele traject. Lieve Pim, bedankt dat ik bij jou mijn project kon – of eigenlijk moest – loslaten, ook al deed ik dat soms met lichte tegenzin.

Curriculum Vitae

Esther Mertens was born January the 2nd 1990 in Dongen, the Netherlands. She obtained her bachelor degree in Child and Family Studies in 2013 at Utrecht University which she combined with the honors program (Von Humboldt College) and a semester studying at the University of Bologna. After the bachelor program, Esther completed both the research master Development and Socialization in Childhood and Adolescence and the clinical master Clinical Child and Family Studies (with honor) at Utrecht University. She obtained the certificate for Psychodiagnostics Assessment and received an award from The Dutch Society for Pedagogics and Education Experts (NVO) for her thesis on differential treatment response. In 2016, Esther started her four-year PhD project at the department of Clinical Child and Family Studies at Utrecht University, which led to the current dissertation. During her PhD project, she visited the Child Behaviour Research Clinic (University of Sydney, three weeks), the Laboratory for Youth and Mental health (Harvard University, one week), and the Centre for Evidence-Based Intervention (University of Oxford, three months). Currently, Esther continues to work at the department of Clinical Child and Family Studies as a post-doc focusing on intervention research.

Publication list

Papers in this dissertation

Mertens, E. C. A., Deković, M., Van Londen, M., & Reitz, E. (2018). The Effectiveness of Rock and Water in Improving Students' Socio-Emotional Adjustment and Social Safety: Study Protocol for a Randomized Controlled Trial. *BMC Psychology, 6,* https://doi.org/10.1186/s40359-018-0247-y

Mertens, E. C. A., Deković, M., Van Londen, M., Nye, E., & Reitz, E. *Solid as a Rock, Flexible as Water? Effectiveness of a School-Based Intervention Addressing Students' Intrapersonal and Interpersonal Domains*. Submitted for publication.

Mertens, E. C. A., Deković, M., Van Londen, M., & Reitz, E. *Personality as a Moderator of Intervention Effects of a School-Based Intervention.* Submitted for publication.

Mertens, E. C. A., Deković, M., Van Londen, M., & Reitz, E. *The Role of Classmates' Modeling and Reinforcement in a Universal School-Based Intervention Addressing Positive Peer Relations.* Submitted for publication.

Mertens, E. C. A., Deković, M., Leijten, P., Van Londen, M., & Reitz, E. *What to Do or Not to Do to Stimulate Students' Intrapersonal and Interpersonal Domains: Meta-Analysis of School-Based Intervention Components.* Submitted for publication.

Other papers

Mertens, E. C. A., Deković, M., Asscher, J. J., & Manders, W. A. (2017). Heterogeneity in response during Multisystemic Therapy: Exploring subgroups and predictors. *Journal of Abnormal Child Psychology, 45,* 1285-1295.

Reitz, E., **Mertens, E. C. A.,** Van Londen, M., & Deković, M., (2019). Veranderingen in sociale veiligheid, competentiebeleving en depressieve gevoelens van basisschoolkinderen die aan het interventieprogramma Rots en Water deelnemen: Een vergelijkingsstudie. *Tijdschrift Klinische Psychologie, 49,* 38-57.

Skills and *struggles* in the intra- and interpersonal domain are important predictors of adolescents' psychosocial wellbeing and their development into adulthood. The *intra*personal domain indicates one's feelings, emotions, and attitudes about the self. The *inter*personal domain indicates the ability to build and maintain positive relationships with others, to understand social situations, roles and norms, and to respond appropriately. A powerful context to address these two domains is the school context.

The overall aim of the current dissertation was to study whether and under what circumstances universal interventions in secondary schools can successfully stimulate students' competencies and prevent problems in the intra- and interpersonal domain. I conducted a meta-analysis to increase insights in these interventions in general and a Randomized Controlled Trial to gain a detailed understanding of a specific intervention, Rock and Water (R&W).

Findings indicate that universal school-based interventions generally showed small positive effects in both the intra- and interpersonal domain. Intervention effects of R&W were comparable in magnitude, especially in the intrapersonal domain, and appeared to be dependent of intervention dosage, suggesting '*less is more*', and students' personality traits. Furthermore, results of the meta-analysis implied that some components associated with stronger intervention effects were rarely implemented, while some components related to weaker effects were often implemented. Together these results indicate that there is potential to improve the effectiveness of universal school-based interventions.