



Solid as a rock, flexible as water? Effectiveness of a school-based intervention addressing students' intrapersonal and interpersonal domains

E.C.A. Mertens^{a,*}, M. Deković^a, M. Van Londen^a, E. Nye^b, E. Reitz^a

^a Utrecht University, Child and Adolescent Studies, Heidelberglaan 1, 3584 CS Utrecht, the Netherlands

^b University of Oxford, Department of Social Policy and Intervention, 32 Wellington Square, Oxford OX1 2ER, United Kingdom

ARTICLE INFO

Editor: Craig A. Albers

Keywords:

School-based intervention
Randomized controlled trial (RCT)
Intrapersonal domain
Interpersonal domain
Components

ABSTRACT

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. The aim of the present study was to examine, first, whether Rock & Water, a social emotional learning intervention that uses active forms of learning, is effective in improving students' competencies and preventing problems in the intra- and interpersonal domain, and second, whether intervention effects were influenced by the extent to which multiple systems are involved in the intervention. We conducted a randomized controlled trial with a sample of 7th grade students ($N = 1299$, $M_{\text{age}} = 12.38$, 54% boys). Students reported on outcomes of the intra- and interpersonal domains using digital questionnaires. The data were analyzed with Latent Growth Curve models. Results showed that the intervention was most effective when only a core team of teachers was involved in the intervention. The intervention improved several proximal outcomes (i.e., self-control and emotional self-regulation) and distal outcomes in students' intrapersonal and interpersonal domains. The intervention effects were strongest, albeit moderate, in the first year of the intervention. These results show that interventions with an active form of learning and implemented by a core team might be promising interventions for prevocational students, although effort should be put in increasing its effectiveness.

1. Introduction

Schools play an eminent role in fostering students' development in the intrapersonal and interpersonal domain (Pellegriano & Hilton, 2012). The *intrapersonal domain* refers to the ability to manage one's own feelings, emotions, and attitudes that pertain to the individual self. For instance, psychological wellbeing and internalizing behaviors fall within this domain (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. Interpersonal relations and aggression are examples that belong to this domain (Pellegriano & 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). When mastery of (some of) these skills falls behind, students can develop problems (Durlak et al., 2011; Modecki et al., 2017). For example, the ability to evaluate and regulate one's inner world and

* Corresponding author.

E-mail address: e.c.a.mertens@uu.nl (E.C.A. Mertens).

<https://doi.org/10.1016/j.jsp.2022.02.002>

Received 1 October 2020; Received in revised form 30 November 2021; Accepted 11 February 2022

Available online 24 February 2022

0022-4405/© 2022 The Authors. Published by Elsevier Ltd on behalf of Society for the Study of School Psychology. This is an open access article under the CC BY license (<http://creativecommons.org/licenses/by/4.0/>).

experiences facilitates the processing of behaviors, thoughts, and emotions resulting in positive personal functioning (Dufner et al., 2019; Finkel & Vohs, 2006). Additionally, the ability to plan one's own behavior and predict other's behavior facilitates socially appealing behavior that, in turn, helps build positive relations (Finkel & Vohs, 2006). Although the two domains influence each other, they are regarded as distinct domains. The intrapersonal domain reflects subjective personal functioning and is related to academic functioning and self-regulation. In contrast, the interpersonal domain reflects social functioning and is related to positive peer relations and conflict resolution (Dufner et al., 2019; Park et al., 2017; Pellegrino & Hilton, 2012). Therefore, schools should intentionally cultivate their students' competencies and prevent development of problems in both domains (Pellegrino & Hilton, 2012).

In the Dutch secondary education system, starting at age 12 (7th grade), there are three separate educational tracks, including the (a) preparatory vocational track (i.e., prevocational track), (b) preparatory college track, and (c) preparatory university track. Students are assigned to these three tracks based on their school results in primary school, a standardized test at the end of primary school, and the judgment of their 6th grade teacher. The *prevocational* track is the least advanced educational track for students who have a preference for 'hands on' learning. These prevocational students seem to be in particular need of school's stimulation of competencies and prevention of problems in the intra- and interpersonal domain given their reports of lower levels of wellbeing, more behavioral problems, and more problems with peers (Stevens & De Looze, 2018), as well as an increased risk for psychological problems (Schrijvers & Schuit, 2010) as compared to students in the other two educational tracks. Additionally, these students can be a more challenging group to involve in school-based interventions as they generally have shown less autonomy, less intrinsic motivation for school and (verbal) learning, and have lower reported cognitive capacities than students following the other two tracks (Timmermans et al., 2017). Although universal school-based interventions addressing competencies in the intra- and interpersonal domains generally show small positive effects (e.g., Durlak et al., 2011; Mertens et al., 2020), it remains unclear whether interventions are able to effectively target this specific group of students as effectiveness of interventions may differ between subgroups (e.g., Onrust et al., 2016).

Due to the sometimes challenging nature of targeting prevocational students, it seems pivotal for interventions that aim to target this group of students to use a 'hands on' learning approach that keeps them motivated and engaged. Research recommends the development of a learning environment characterized by a practical orientation that allows students to apply their newly learned skills in practice and set their own goals (Smit et al., 2014). Short moments of instructions or reflections are alternated with practicing new skills in a relevant context in which teachers function as a coach and role-model (De Bruijn et al., 2005; Smit et al., 2014). Teaching methods consist of activating stepwise, structured tasks to help students regulate their learning processes. Such an environment is suggested to foster students' understanding and use of the learned content (De Bruijn et al., 2005; Koopman et al., 2011). Hence, a promising intervention for prevocational students seems to be an intervention that combines a psychological approach (i.e., instruction and reflection) with active learning approaches (i.e., exercises, practice of skills, and role-play) that allows students to set their own goals and has a facilitator who acts as a coach and role-model.

A universal school-based intervention that aligns with the active learning style of prevocational students is Rock and Water (R&W; Ykema, 2002, 2018). R&W is a social emotional learning (SEL) intervention that uses active forms of learning. R&W aims to foster students' psychosocial development and social safety. The intervention focuses on psychosocial development by aiming to improve social skills and autonomy, theorizing that improving these competencies may coincide with increases in students' psychological wellbeing, resilience, and sexual autonomy and concomitant decreases in internalizing behavior. The intervention focuses on social safety by aiming to stimulate social and communication skills, create a safe and respectful environment, and provide alternative behaviors to the students, theorizing that this approach decreases externalizing behavior, aggression, and bullying, thereby resulting in more positive interpersonal relations among students. The applied teaching method is a combination of instruction, reflection, exercises to create physical awareness, and practicing skills. The trainer is positive, supporting, and serves as a role model.

R&W is based on the theory of the 'R&W house' (Ykema, 2002, 2018). This house is built on a foundation of four pillars: self-control, self-reflection, self-esteem, and emotional self-regulation. These four competencies are considered to be proximal outcomes of the intervention, forming the basic skills that enable students to develop themselves on the broader competencies. These broader competencies are the distal outcomes of the intervention represented in the five modules of the house. 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 focuses on communication with others, targeting behaviors such as positive social interactions between classmates. The fourth module aims to teach students to make decisions based on their own intuition and preferences. The fifth module emphasizes connectedness with others. Together the last two modules target more general feelings of psychological wellbeing.

Even though R&W is implemented in many countries (e.g., Australia, China, Singapore, France, the Netherlands), very little is known about the effectiveness of the intervention. Several small-scaled studies have indicated that after completing R&W participants felt more resilient, experienced a more positive identity, and used more active than passive coping styles (Ykema et al., 2006). More recently, De Graaf et al. (2016) examined the effectiveness of R&W on sexual aggression between intervention students and students in a waitlist control condition. They implemented seven R&W lessons of 90 min each in nine schools in the Netherlands. The intervention was implemented only for boys in the prevocational track in Grades 9–10. Results showed that self-reported coercive strategies and verbal manipulation decreased, whereas self-regulation and efficacy increased after completing R&W. 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 specifically its effectiveness when used with students in the prevocational track. Therefore, the first aim of the present study was 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 whether intervention effects were influenced by the extent to which multiple systems were 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, classroom, and school systems. Actively engaging these interrelated systems in the intervention may facilitate generalization and maintenance of new skills (Mahoney et al., 2021). For instance, Mertens

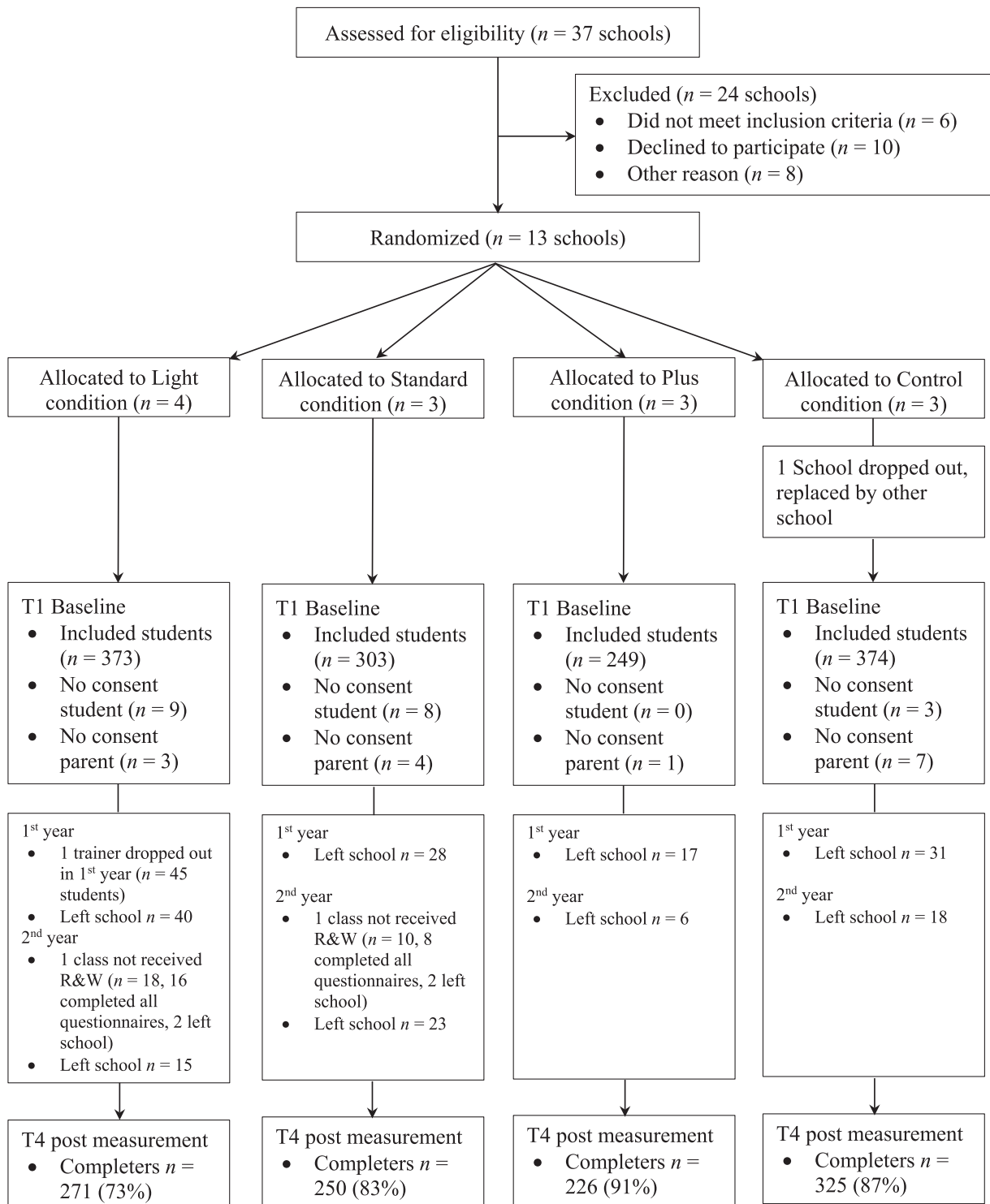


Fig. 1. Flow chart.

et al. (2020), in their meta-analysis examining components of universal school-based interventions, concluded that interventions involving the whole school and/or parents showed stronger effects on system level outcomes than interventions without these components. However, this positive effect of involving multiple systems in an intervention was not always found. For example, Durlak et al. (2011) found in their meta-analysis of SEL interventions 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 insight is required into how involving multiple systems in an intervention may influence its effectiveness (Kuosmanen et al., 2019). This knowledge will help close the science-to-practice gap as it informs schools how distinct components may affect intervention effects under real-world circumstances (Kuosmanen et al., 2019). This information is pivotal for schools in order to determine the extent of investment in interventions (e.g., number of teachers to train, whether or not to involve parents) that is necessary for interventions to work.

In sum, the present study consisted of two aims. First, we examined whether a universal school based SEL intervention (i.e., R&W) could stimulate prevocational students' development in the intrapersonal and interpersonal domains. We hypothesized that participants in the R&W conditions would show statistically significant improvements as compared to participants in the control condition on all assessed proximal and distal outcomes. Second, we examined whether three differently implemented versions of R&W (i.e., involving only a core team of teachers [R&W trainers], involving the entire teaching staff, or involving all teaching staff and parents) produced different intervention effects. We had no specific hypothesis concerning this second aim as research examining the involvement of multiple socio-ecological systems in interventions has resulted in inconsistent findings.

2. Method

2.1. Design and procedure

The effectiveness of R&W was examined in a randomized controlled trial consisting of four conditions. 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, a parent component was added to the Standard condition (i.e., 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).

Schools with a preparatory vocational education track (i.e., one of three educational tracks in the Dutch secondary school system) were eligible to participate in the present study. Some schools provide only the prevocational track, whereas other schools offer all tracks. However, each track has separate education programs, teachers, and classes. Regardless of the provided educational tracks within a school, the intervention was only implemented and examined in the prevocational track. Schools were excluded from this study if they had implemented R&W in the last 2 years or if they were special education schools. Thirteen schools throughout the Netherlands from 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 that showed interest in participating in the study (after randomization) and also met the inclusion criteria (see Fig. 1 for the flow chart).

Students completed digital questionnaires (a) before the intervention started (i.e., baseline; T1 = October/November 2017), (b) after completing the first year of R&W lessons (i.e., T2 = March/April 2018), (c) before the start of the R&W lessons in the second year of R&W (i.e., T3 = October 2018), and after the intervention (i.e., 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 Netherlands Trial Register number NL6371 (old number NTR6554; for protocol, see Mertens et al., 2018).

2.2. Participants

At baseline, the sample consisted of 1299 7th grade students. In the Netherlands, 7th grade generally corresponds with ages 12–13 years. In our sample, the students had an average age of 12.38 years ($SD = 0.62$). 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	$\eta^2_{\text{partial}/\phi}$
<i>N</i>	373	303	249	374			
Age, <i>M</i> (<i>SD</i>)	12.33 (0.57)	12.38 (0.66)	12.34 (0.60)	12.47 (0.64)	3.89	0.009	0.009
Boys, <i>n</i> (%)	170 (48%)	161 (56%)	131 (55%)	199 (57%)	7.38	0.061	0.077
Western background, <i>n</i> (%)	291 (82%)	115 (43%)	211 (91%)	198 (59%)	182.01	<0.001	0.392

There were no differences between the conditions regarding sex distribution, but there were differences in students' age and ethnic background (see Table 1). Students in the control condition were 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.

Overall, 15% of the data were missing. Little's MCAR test indicated that the data were not missing completely at random ($\chi^2(2339) = 2539.54, p = .002$). 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 those who remained in the study. Concerning the demographic variables, there were no differences on age ($F(1,1230) = 0.24, p = .626, \eta^2_{\text{partial}} = 0.000$), sex distribution ($\chi^2(1) = 0.1.36, p = .244, \phi = -0.033$), and ethnic background ($\chi^2(1) = 1.19, p = .276, \phi = 0.032$). Concerning the outcome variables, three MANOVAs, one per time point, showed that there were differences on the outcome variables at T1 ($F(39, 3591) = 1.45, p = .035, \eta^2_{\text{partial}} = 0.016$), but not at T2 ($F(39, 3363) = 0.72, p = .898, \eta^2_{\text{partial}} = 0.008$) and T3 ($F(26, 2032) = 0.74, p = .824, \eta^2_{\text{partial}} = 0.009$). 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 after Bonferroni correction ($p = .124$). When separately examining differences between students who dropped out and students who remained in the study, we did not find any differences on any of the demographic or outcome variables. Because Little's MCAR test tends to yield conservative results when applied to a large set of variables and because we found no differences between dropouts and completers on the demographic or any of the outcome variables, we regarded the missing data as missing at random (Van Ness et al., 2007).

2.3. Conditions and fidelity

2.3.1. R&W

R&W (Ykema, 2002, 2018) aims to foster competencies and to prevent problems in students' intra- and interpersonal domains through active forms of learning. This approach integrates play and exercises to guide students in how to make (physical) contact with others and to explore, respect, and set one's own and other's boundaries. Via games and exercises, body-awareness (e.g., muscle tension, breathing), nonverbal communication, and coping strategies of students are enhanced. For instance, 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. In another exercise, 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, according to the R&W theory. In yet another exercise, students play games together in which they try to tap the other student out of balance, aiming to teach students that sometimes it is better to stand your ground (i.e., as a rock) and sometimes it is better to move along with others (i.e., as water). In addition to these games, R&W also applies a behavioral approach. Students engage in role-playing to practice, for example, how to set and indicate boundaries, how to react calmly to provocations, and how to help students who get bullied. During and after the exercises, students receive feedback and reflect on the exercise. Feedback and reflection are guided by the trainer asking questions such as "Where was your breathing during the game?" and "How did you feel?". Trainers function as coaches, model and reinforce desired behaviors, and create a safe environment in which students are allowed to make mistakes. Skills central to the intervention, such as low breathing, standing strong, and setting boundaries, are discussed and practiced in multiple intervention lessons enabling students to continue practicing these skills. Furthermore, in the first intervention lesson, students set an individual physical exercise goal (e.g., number of push-ups, jumping rope) that they aim to achieve at the end of the intervention and work toward this goal in each intervention lesson. In 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 manualized program of 22 sessions in total that are implemented over 2 academic years (Ykema, 2002, 2018). In the first year, students received 14 R&W lessons and in the second year they received eight R&W lessons. The lessons were implemented weekly during 90-min physical education lessons. Trainers were (mostly) physical education teachers at the schools given that they have experience with teaching physical activities in class. During the lessons, students participated in physical exercises, games and role-play, practiced the new skills, reflected on the exercises, shared and discussed their thoughts with each other, summarized what was discussed during the lesson, and addressed how to use the learned skills in their daily lives (more information about the intervention in the study protocol, see Mertens et al., 2018).

2.3.2. Different versions of R&W

In the Light condition, a core team of teachers (i.e., R&W trainers) who completed the 3-day training to become certified R&W trainers implemented the intervention lessons. In the Standard condition, besides the teachers providing the actual R&W lessons (i.e., R&W trainers), the rest of the schools' teaching staff also received a 3-day introductory training to learn (a) the basic principles of R&W, (b) how to support the R&W trainers, and (c) how to apply the intervention techniques in their regular classes. For instance, teachers could do a short breathing exercise with students before the start of an exam. In the Plus condition, a parent component was added to the Standard version of intervention program. Parents (a) were invited to watch a documentary about R&W, (b) were invited to join a R&W lesson at the school, and (c) received weekly e-mails with information about the current week's R&W lesson 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.

2.3.3. Control condition

In the control condition, students received care as usual, which varied between schools. For example, in one school, students had a teacher as a 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 (i.e., 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 had a mentor, who also was a teacher, whom they could go to when they experienced difficulties; there also was an 'anti-bullying coordinator' at the school. This coordinator facilitated actions to prevent or stop bullying which could differ per situation.

2.3.4. Intervention fidelity

Intervention fidelity was assessed in two steps. First, we assessed the fidelity of the 90-min intervention lessons. Second, we assessed whether the additional teacher and parent components were actually implemented.

The fidelity of the R&W lessons were measured with two complementary methods: (a) self-reports of R&W trainers and (b) observations of 67 R&W lessons (Light = 28 lessons; Standard = 17 lessons; Plus = 22 lessons) by three R&W experts who completed the advanced R&W training and had multiple years of experience with providing the R&W intervention. R&W trainers completed a questionnaire about fidelity after every third lesson in the first year of the intervention and after every second lesson in the second year of the intervention. R&W experts completed a coding schema based on Bishop et al. (2014) during the observation. Reliability of the coding between R&W experts was not assessed due to lack of resources.

According to the self-reports, R&W trainers were generally able to complete lessons (65%) and did not deviate or only slightly deviated from the manual (72%). Trainers reported no significant differences on intervention fidelity across conditions. According to the observations, 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 adjusted 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 judgment of the R&W experts. Intervention fidelity concerning the R&W lessons did not differ between the conditions nor between the two years of implementation. 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.

To determine whether the Standard and Plus versions were implemented as intended, 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 ($F(2,24) = 5.73, p = .009$); in the Light condition only the R&W trainers (i.e., trained teachers) implemented R&W, whereas in the Standard and Plus conditions the R&W trainers reported that other teachers also applied the intervention techniques. Parents in the Plus condition who responded to the questionnaire ($n = 47$) reported that they read the weekly information sometimes (57%) or often (39%), and some parents indicated they watched the documentary about R&W (15%) and participated in a R&W lesson at their child's school (23%). Furthermore, parents in the Plus condition, as compared to parents in the Standard condition, reported to have talked more about R&W at home ($F(2,75) = 4.37, p = .016$) and used it more often with their child ($F(2,75) = 3.44, p = .037$). Hence, as planned, parents in the Plus condition were more involved in the intervention than in the Light and Standard conditions. Thus, both the teacher and parent components appeared to have been implemented.

2.4. Proximal outcomes

2.4.1. 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 et al., 2005) that consists 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 $\alpha = 0.62$ – 0.72).

2.4.2. Self-reflection

Students completed the Engage subscale from the Self-Reflection and Insight Scale (Sauter et al., 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 $\alpha = 0.53$ – 0.60), and subsequently, the 3 items that included a negative valence ("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 $\alpha = 0.74$ – 0.89).

2.4.3. Self-esteem

Students' level of global self-worth was self-reported using the subscale Global self-perception of the Self-Perception Profile (Harter, 1988). This subscale consists of 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 $\alpha = 0.73$ – 0.75).

2.4.4. Emotional self-regulation

Students completed the Difficulties in Emotion Regulation Scale (Anderson et al., 2016) to assess their 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 indicated higher levels of emotional self-regulation (Cronbach’s $\alpha = 0.88$ – 0.91).

2.5. Distal outcomes

2.5.1. Intrapersonal domain

2.5.1.1. Psychological wellbeing. To measure the presence of positive emotions, students completed the Psychological Wellbeing subscale 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 $\alpha = 0.76$ – 0.83).

2.5.1.2. Resilience. To assess students’ ability to bounce back from challenges that arise in life, students completed the Connor-Davidson Resilience Scale–Short Version (Davidson & Connor, 2017) that consists 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 $\alpha = 0.79$ – 0.92).

2.5.1.3. Sexual autonomy. To measure students’ coping skills in sexual situations, students completed 5 items from the Sex under 25 survey (e.g., “When I am with someone I like, I feel at ease”; De Graaf et al., 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 $\alpha = 0.53$) and adequate at T2, T3, and T4 (Cronbach’s $\alpha = 0.62$ – 0.65).

2.5.1.4. 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 $\alpha = 0.79$ – 0.86).

2.5.2. Interpersonal domain

2.5.2.1. Interpersonal relations in the class. Perceived interpersonal relations in the class were assessed using the Classroom Peer Context Questionnaire (Boor-Klip et al., 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 $\alpha = 0.80$ – 0.85).

2.5.2.2. 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 $\alpha = 0.65$ – 0.79).

2.5.2.3. 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 $\alpha = 0.65$ – 0.83).

2.5.2.4. 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*).

2.6. 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 (Appendix Table A1 provides an overview of the number of classrooms and trainers per school). We took clustering at the school level into account by applying the complex sample cluster feature of *Mplus* (version 8.2; Muthén & Muthén, 1998–2010). This is a conservative clustering procedure providing unbiased estimates of the standard errors (Muthén & Muthén, 1998–2010). Clustering at the class level was not taken into account as class composition was not stable over the years (e.g., Cross et al., 2016). This analysis approach was supported by examining the intraclass correlations (ICC). ICC at the school level and at the class level were generally not significant. Additionally, design effects ($d = 1 + (\text{average cluster size} - 1) * \text{ICC}$; Muthén, 2000) at the school level were generally larger than 2, indicating that clustering at the school level was a concern. In contrast, at the class level, design effects were generally smaller than 2, indicating that clustering at the class level was not of concern. 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, 1998–2010).

To examine the effectiveness of R&W, we tested a series of latent growth curve (LGC) models in *Mplus*, as suggested by Greenberg and Abenavoli (2017). LGC models estimate for each participant an individual growth curve based on their 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, 1998–2010). The slope is of main importance; when the intervention is effective as 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 0 and 3, respectively.

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 as compared to the control group, we examined the effectiveness of those conditions as 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 ($d = (\text{slope} * \text{duration})/\text{SD}$; 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

3. Results

3.1. 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 < 0.08, CFI > 0.90, SRMR < 0.10; 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 are presented in Appendix Table A2).

3.2. Effects of R&W

The standardized regression coefficients of the slope on the intervention conditions (as 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 (as compared to the control condition) in Table 5.

3.3. Proximal outcomes

Students in the Light conditions showed more beneficial trajectories of change for self-control and emotional self-regulation as compared to students in the control condition (see Fig. 2). Cohen's d effect sizes were 0.25 and 0.29, respectively. Students in the Light condition showed a steeper increase in self-control (mean slope = 1.75) than students in the control condition (mean slope = 1.44). Regarding emotional self-regulation, students in the Light condition improved over time (mean slope = 0.24), whereas students in the control condition slightly declined (mean slope = -0.10). For both outcomes, students improved most in the first year of the intervention. Furthermore, the results suggested a relevant difference (Cohen's $d = 0.22$), although not significant ($p = .060$), between the trajectory of change in self-esteem of students in the Light condition as compared to students in the control condition. Students in the Light condition remained relatively stable (mean slope = -1.68), whereas students in the control condition showed a small decrease in self-esteem (mean slope = -2.58). The intervention effect was strongest from T2 to T3, which was 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 Appendix Table A3 for the mean slopes of all conditions on the outcomes). Overall, the Light condition seemed to have the most beneficial growth trajectories for the proximal outcomes with most improvement early in the intervention.

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
	<i>M (SD)</i>	<i>M (SD)</i>	<i>M (SD)</i>	<i>M (SD)</i>	<i>M (SD)</i>	<i>M (SD)</i>	<i>M (SD)</i>	<i>M (SD)</i>	<i>M (SD)</i>	<i>M (SD)</i>	<i>M (SD)</i>	<i>M (SD)</i>	<i>M (SD)</i>	<i>M (SD)</i>	<i>M (SD)</i>	<i>M (SD)</i>
Proximal outcomes																
Self-control	3.33 (0.58)	3.53 (0.60)	3.52 (0.61)	3.54 (0.60)	3.49 (0.57)	3.66 (0.66)	3.56 (0.65)	3.61 (0.69)	3.41 (0.60)	3.53 (0.59)	3.43 (0.61)	3.53 (0.65)	3.49 (0.56)	3.59 (0.62)	3.57 (0.65)	3.58 (0.64)
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 (0.62)	3.28 (0.66)	3.24 (0.65)	3.28 (0.62)	3.28 (0.57)	3.33 (0.60)	3.29 (0.66)	3.29 (0.62)	3.25 (0.59)	3.33 (0.57)	3.18 (0.68)	3.12 (0.70)	3.31 (0.59)	3.29 (0.64)	3.22 (0.65)	3.22 (0.65)
Emotional self-regulation	3.83 (0.77)	3.94 (0.85)	3.98 (0.76)	3.97 (0.75)	3.85 (0.69)	3.99 (0.80)	3.91 (0.82)	3.91 (0.84)	3.87 (0.71)	3.88 (0.74)	3.81 (0.80)	3.83 (0.85)	3.93 (0.71)	3.94 (0.75)	3.90 (0.76)	3.92 (0.76)
Intrapersonal domain																
Psychological wellbeing	3.94 (0.62)	3.98 (0.69)	3.91 (0.71)	3.87 (0.78)	4.02 (0.61)	4.10 (0.71)	4.02 (0.73)	3.93 (0.77)	3.97 (0.62)	4.06 (0.63)	3.87 (0.72)	3.85 (0.80)	4.06 (0.59)	4.02 (0.68)	3.90 (0.75)	3.92 (0.78)
Resilience	2.36 (0.65)	2.49 (0.75)	2.15 (0.87)	2.47 (0.82)	2.43 (0.67)	2.70 (0.71)	2.56 (0.70)	2.53 (0.90)	2.43 (0.64)	2.57 (0.68)	2.47 (0.71)	2.51 (0.84)	2.51 (0.65)	2.68 (0.69)	2.57 (0.74)	2.61 (0.85)
Sexual autonomy	3.23 (0.44)	3.39 (0.48)	3.40 (0.44)	3.40 (0.46)	3.30 (0.44)	3.44 (0.52)	3.38 (0.54)	3.39 (0.51)	3.29 (0.40)	3.38 (0.46)	3.34 (0.45)	3.35 (0.57)	3.32 (0.44)	3.36 (0.49)	3.32 (0.50)	3.36 (0.52)
Internalizing behavior	0.53 (0.42)	0.41 (0.44)	0.44 (0.44)	0.42 (0.46)	0.44 (0.42)	0.33 (0.39)	0.40 (0.44)	0.37 (0.43)	0.49 (0.40)	0.45 (0.42)	0.51 (0.44)	0.48 (0.51)	0.43 (0.39)	0.39 (0.41)	0.45 (0.46)	0.42 (0.44)
Interpersonal domain																
Interpersonal relations in the class	3.90 (0.64)	3.81 (0.68)	3.83 (0.67)	3.73 (0.69)	3.98 (0.63)	3.87 (0.72)	3.81 (0.73)	3.80 (0.75)	4.07 (0.62)	4.00 (0.63)	3.92 (0.69)	3.84 (0.81)	3.91 (0.59)	3.78 (0.65)	3.77 (0.67)	3.75 (0.72)
Externalizing behavior	0.59 (0.38)	0.47 (0.39)	0.48 (0.39)	0.45 (0.41)	0.54 (0.39)	0.43 (0.38)	0.50 (0.44)	0.42 (0.42)	0.55 (0.35)	0.46 (0.36)	0.55 (0.38)	0.49 (0.45)	0.49 (0.34)	0.42 (0.37)	0.45 (0.39)	0.43 (0.39)
Aggression	1.79 (0.56)	1.71 (0.59)	1.69 (0.59)	1.66 (0.66)	1.85 (0.61)	1.84 (0.73)	1.94 (0.73)	1.83 (0.81)	1.73 (0.46)	1.72 (0.60)	1.80 (0.58)	1.74 (0.81)	1.78 (0.56)	1.76 (0.65)	1.81 (0.67)	1.85 (0.78)
Bullying	1.13 (0.48)	1.18 (0.61)	1.15 (0.55)	1.21 (0.64)	1.13 (0.41)	1.18 (0.60)	1.15 (0.58)	1.19 (0.64)	1.10 (0.41)	1.15 (0.47)	1.20 (0.68)	1.20 (0.64)	1.08 (0.36)	1.18 (0.59)	1.22 (0.71)	1.22 (0.71)
Victimization	1.32 (0.77)	1.43 (0.96)	1.30 (0.77)	1.31 (0.86)	1.28 (0.73)	1.35 (0.87)	1.22 (0.77)	1.22 (0.70)	1.36 (0.88)	1.37 (0.85)	1.40 (0.96)	1.40 (0.93)	1.37 (0.94)	1.56 (1.16)	1.41 (1.04)	1.34 (0.95)

Table 3
Model fit indices of outcomes.

	Model fit statistics		
	RMSEA	CFI	SRMR
Proximal outcomes			
Self-control ¹	0.057	0.936	0.077
Self-reflection	0.060	0.912	0.067
Self-esteem	0.068	0.925	0.073
Emotional self-regulation	0.057	0.931	0.075
Intrapersonal domain			
Psychological wellbeing	0.062	0.907	0.067
Resilience	0.060	0.796	0.084
Sexual autonomy ¹	0.053	0.862	0.073
Internalizing behavior ¹	0.061	0.934	0.068
Interpersonal domain			
Interpersonal relations in the class	0.053	0.918	0.069
Externalizing behavior	0.059	0.919	0.070
Aggression	0.055	0.888	0.069
Bullying	0.050	0.635	0.068
Victimization	0.058	0.696	0.069

Note. ¹Variance 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 4
Intervention effects over time of R&W conditions as compared to control condition.

	Light			Standard			Plus		
	β	SE	p	β	SE	p	β	SE	p
Proximal outcomes									
Self-control ¹	0.12	0.03	< 0.001	0.00	0.04	0.951	0.02	0.04	0.532
Self-reflection	0.01	0.04	0.803	-0.02	0.06	0.761	0.04	0.03	0.238
Self-esteem	0.18	0.10	0.060	0.09	0.11	0.394	-0.05	0.06	0.327
Emotional self-regulation	0.17	0.05	< 0.001	0.07	0.05	0.163	-0.01	0.02	0.676
Intrapersonal domain									
Psychological wellbeing	0.15	0.06	0.009	0.08	0.08	0.287	0.04	0.04	0.228
Resilience	0.09	0.73	0.902	0.18	0.27	0.507	-0.10	0.31	0.736
Sexual autonomy ¹	0.18	0.04	< 0.001	0.08	0.06	0.144	0.06	0.03	0.074
Internalizing behavior ¹	-0.13	0.04	< 0.001	-0.07	0.04	0.048	0.01	0.03	0.812
Interpersonal domain									
Interpersonal relations in the class	0.05	0.03	0.056	-0.05	0.07	0.499	-0.04	0.04	0.278
Externalizing behavior	-0.20	0.12	0.107	-0.13	0.11	0.223	-0.02	0.07	0.756
Aggression	-0.19	0.09	0.030	0.01	0.10	0.933	-0.01	0.05	0.822
Bullying	-0.12	0.06	0.056	-0.12	0.09	0.149	-0.05	0.09	0.624
Victimization	-0.01	0.04	0.817	-0.04	0.06	0.527	0.03	0.04	0.530

Note. ¹Variance 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.

3.4. Distal outcomes

3.4.1. Intrapersonal domain

Students in the Light condition showed a more beneficial trajectory of change for psychological wellbeing, sexual autonomy, and internalizing behavior as compared to students in the control condition (see Fig. 3). Cohen's *d* effect sizes ranged from 0.26 to 0.38. In the Light condition, students remained stable on psychological wellbeing (mean slope = 0.08), whereas students in the control condition decreased (mean slope = -0.25). The intervention effect was strongest from T2 to T3, which was between the first and second year of the intervention. On sexual autonomy, students in the Light condition showed a steeper increase (mean slope = 0.73) than students in the control condition (mean slope = 0.32). Students improved most in the first year of the intervention. For internalizing behavior, students in the Light and Standard conditions showed a steeper decline (mean slope_{Light} = -0.60; mean slope_{Standard} = -0.53) than students in the control condition (mean slope = -0.41). In both the Light and Standard conditions, the strongest improvements were again in the first year. No intervention effects were found for resilience and for students in the Plus condition (see

Table 5
Effect sizes of change over time in R&W conditions as 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
Proximal 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
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

Note. ¹Variance 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.

Appendix Table A3 for the slopes of all conditions on the 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 (mean slope_{Light} = -0.60; mean slope_{Standard} = -0.53). Overall, students in the Light condition seemed to have a more beneficial change over time in the intrapersonal domain, with the intervention's primary impact early in the intervention.

3.4.2. Interpersonal domain

Students in the Light condition showed a more beneficial trajectory of change for aggression as compared to students in the control condition (see Fig. 3). Cohen's *d* effect size was 0.31. Students in the Light condition remained relatively stable over time (mean slope = 0.09), whereas students in the control condition showed an increase in aggression (mean slope = 0.56). The strongest intervention effect was found in the first year of the intervention. Furthermore, the results suggested a relevant difference (Cohen's *d* = 0.24), albeit not significant ($p = .056$), between the trajectory of change in bullying of students in the Light condition as compared to students in the control condition. Students in the Light condition showed a less steep increase in bullying (mean slope = 2.34) than students in the control condition (mean slope = 2.80; see Fig. 3). The intervention effect was strongest in the first year. No intervention effects were found for interpersonal relations in the class, externalizing behavior, or victimization. Additionally, no intervention effects were found for students in the Standard and Plus conditions (see Appendix Table A3 for the slopes of all conditions on the outcomes in the interpersonal domain). Overall, the Light condition also seemed to have a more beneficial change over time in the interpersonal domain with, again, the strongest intervention effect in the first year of the intervention.

4. Discussion

The purpose of the present study was, first, to examine whether an SEL intervention using active forms of learning could positively affect prevocational students, a challenging group to involve in interventions, and, second, to determine whether different versions of this intervention yielded different intervention effects. The intervention, R&W, appeared to be moderately effective in stimulating the proximal outcomes (i.e., self-control, emotional self-regulation, and a trend for self-esteem) and in fostering some aspects of students' intra- and interpersonal domains, but only when the intervention was implemented solely by a core team of teachers. R&W Light was particularly 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 tendency to function as a buffer against increasing aggression and bullying. Intervention effects were generally moderate (when significant effects were found they ranged from 0.24 to 0.38) in comparison to the mean distribution of effect sizes for universal interventions (Tanner-Smith et al., 2018). Additionally, our observed effect sizes are consistent with effect sizes reported for other universal school-based interventions addressing the intra- and/or interpersonal domain. For instance, Durlak et al. (2011) reported meta-analysis effect sizes of 0.23 for positive attitudes toward the self and others, 0.24 for positive social behaviors, 0.22 for conduct problems, and 0.24 for emotional distress. Mertens et al. (2020) found meta-analysis effect sizes of 0.19 and 0.15 on, respectively, the intrapersonal and interpersonal domains in general.

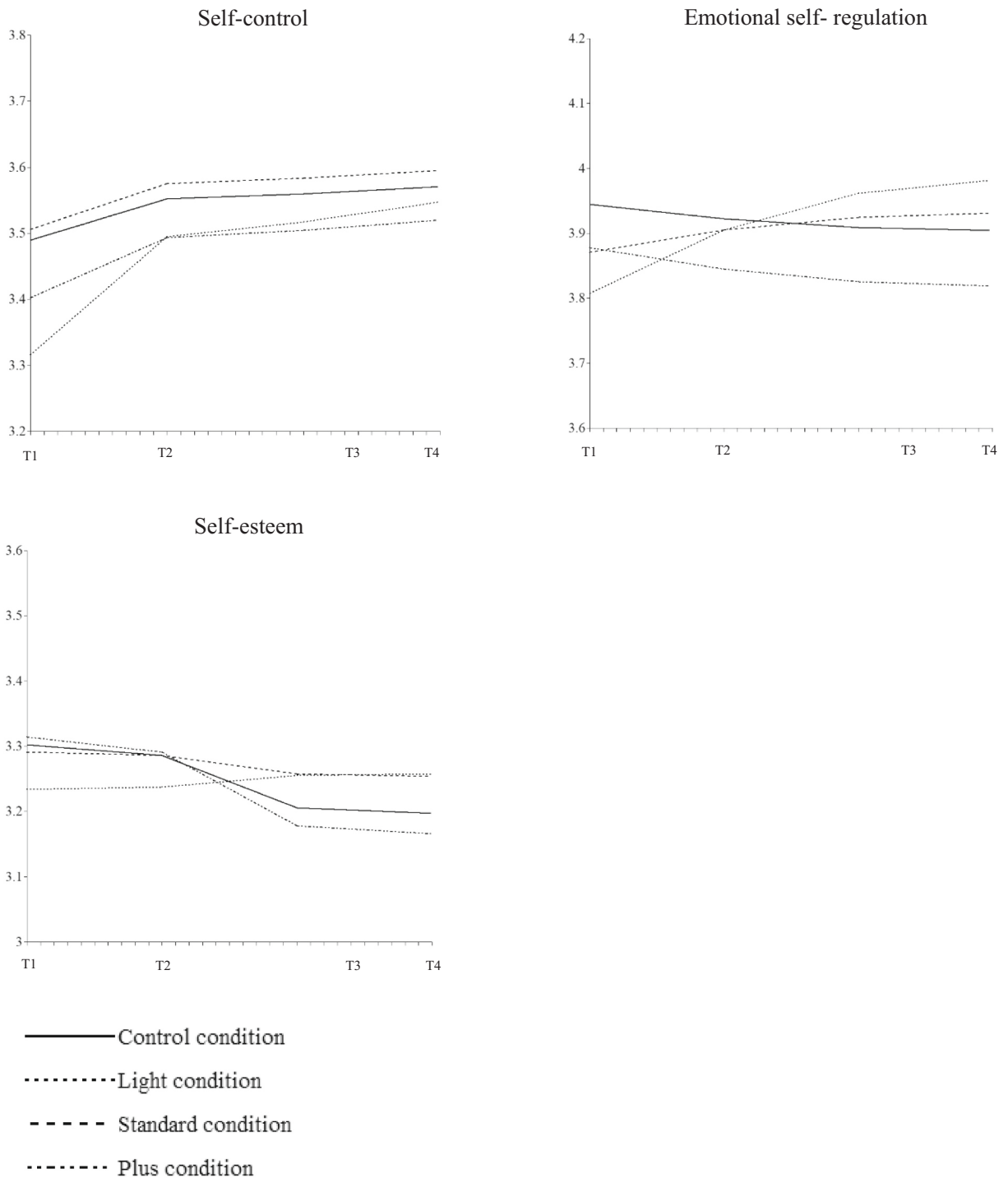
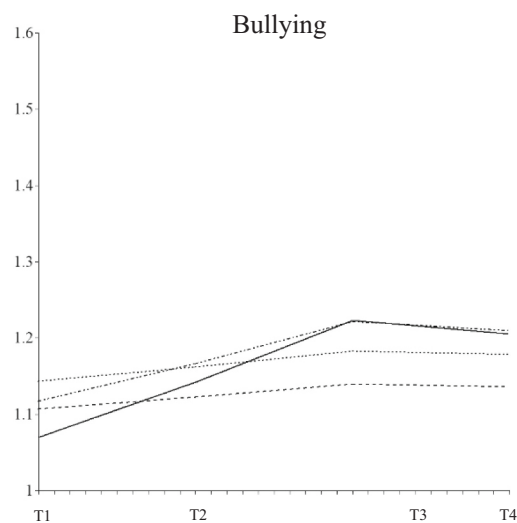
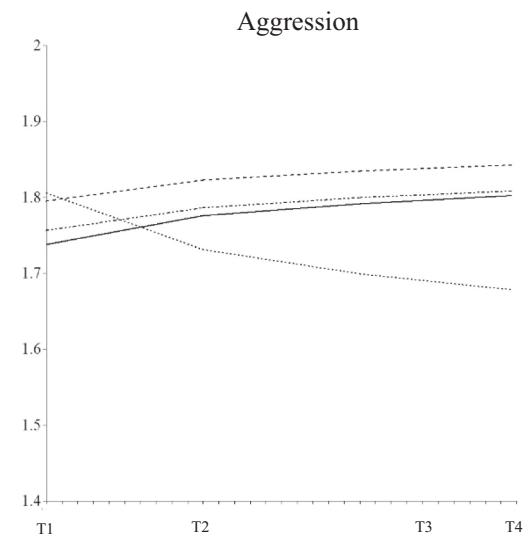
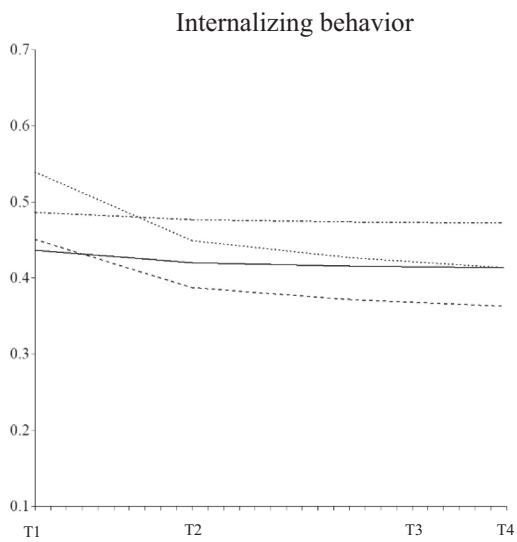
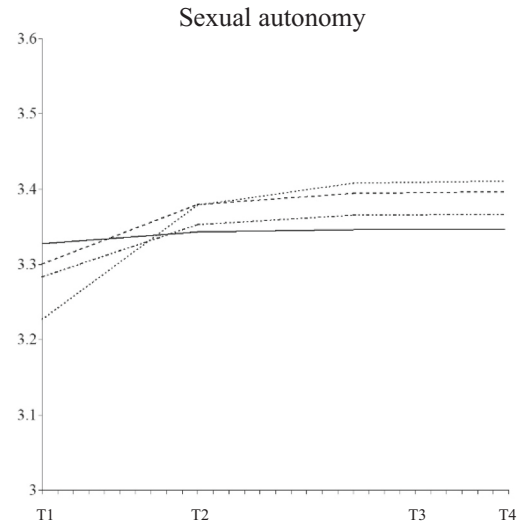
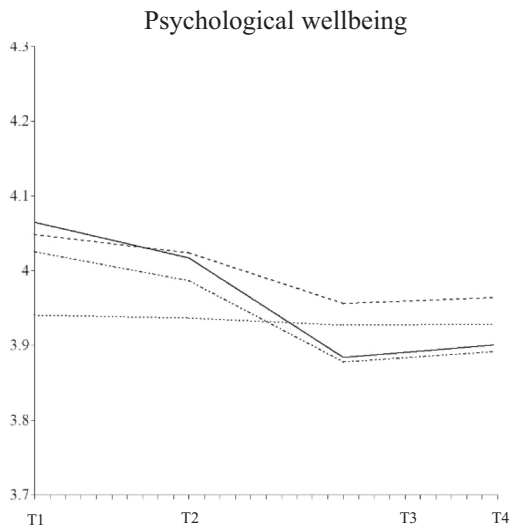


Fig. 2. Estimated growth trajectories of the intervention and control conditions concerning proximal outcomes.

It appears that using an active learning approach is a method that fits well with interventions targeting prevocational students; the alternation between psychological instruction or reflection and practical exercises or games may increase students' engagement in the intervention (Ter Vrugte et al., 2015), which, in turn, supports more optimal benefits from the intervention.

Regarding the proximal 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), thus indicating the intervention's potential. However, no intervention effect was found on students' self-reflection, which is one of the



— Control condition
 Light condition
 - - - Standard condition
 - · - · Plus condition

(caption on next page)

Fig. 3. Estimated growth trajectories of the conditions concerning the distal outcomes in the intra- and interpersonal domains.

pillars of the R&W house. Also previous interventions have appeared to be ineffective in improving self-reflection of prevocational students. For instance, [Ter Vrugte et al. \(2015\)](#) 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. It is possible that (early) adolescence is not the best developmental phase to measure self-reflection given that metacognitive skills accelerate during adolescence ([Barber, 2005](#)). Hence, adolescents' metacognitions enabling them to meaningfully reflect on thoughts, feelings, and behaviors ([Sauter et al., 2010](#)) might still be in the early stages of development; stimulating self-reflection may be more fruitful later in life when these metacognitions are more developed. Nonetheless, as self-reflection has been indicated as a potential effective component in secondary 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, when and how self-reflection can best be stimulated in these students.

Concerning the distal outcomes, most intervention effects were found in students' intrapersonal domain. Only one intervention effect (i.e., aggression, and one trend for 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 practical approach is used when addressing that domain, whereas 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 (e.g., "How did the bully feel when they said 'stop' together?"). This more verbal approach is cognitively more demanding and could possibly hinder prevocational students from benefitting from the intervention in the interpersonal domain.

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 et al., 2003](#); [Cuijpers, 2002](#); [De Mooij et al., 2020](#)). These meta-analyses have suggested that short-term interventions with a modest number of sessions are preferred. Research also 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 potentially related to a decrease in students' motivation, as motivation is frequently found to be a moderator of intervention effects ([Philips & Wennberg, 2014](#)). Thus, implementing only the first year of the intervention may minimize the burden on the students without jeopardizing the effectiveness of the intervention.

The second aim was to determine the potential influence of involving multiple socio-ecological systems in the intervention on the intervention's effectiveness. The results showed that the condition in which 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 consistent with the results of the meta-analysis by [Durlak et al. \(2011\)](#) that showed that the positive effect of involving more people and systems in interventions was not always found. Moreover, this finding underscores the importance of systematically examining how inclusion of different components relates to intervention effects. These insights can enable the optimization of intervention implementation and guide the investment of resources for schools implementing an intervention.

Implementing an intervention with a narrow socio-ecological focus (e.g., implementing the intervention program without additional teacher and/or parent components) might benefit from trainers' feelings of responsibility for implementation. In an intervention with a narrow socio-ecological focus, only a few people are involved, and thus they are solely responsible for properly and effectively implementing the intervention. These people may become "program champions" who provide support and solve problems and difficulties regarding the intervention, which can strengthen intervention implementation ([Durlak, 2016](#)). In an intervention with a broad socio-ecological focus (e.g., adding a teacher and/or parent component to the intervention program), 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 enact 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 an entire teaching staff and a parent component may have an increased risk of sending mixed messages to the students due to the large number of people who 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](#)), but the application of intervention techniques during regular lessons or at home is not described nor structured. This lack of structure can be especially challenging with techniques based on an active learning 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 thus confuse students and push the intended intervention message to the background, thus reducing the chance of students to benefit from the intervention. Stating explicit guidelines and clear goals for all people involved in the intervention is critical for high quality implementation and may facilitate integration within and across systems ([Kuusmanen et al., 2019](#)).

The results of the present study should be interpreted in light of its strengths and limitations. Strengths of the study were the multiple measurement points, the three conditions differing in the width of socio-ecological focus, and the large sample size. This enabled us to examine change in students over two years with different levels of socio-ecological focus. A limitation existed in our lack of information about the content of the involvement of the entire teaching staff and the parent component. Although we know that the teacher and parent components were implemented, we know relatively little about *how* intervention techniques were applied by the teachers and parents. This additional information would have provided insight in possible mixed intervention messages. Hence, future research should also measure how intervention techniques are used in specific components to examine whether differential implementation affects intervention effects. This knowledge can shed light on whether or not (and to what extent) the content of components 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. Conversely, using first time trainers ensured that trainers in our study were comparable in their experience with the intervention. Last, we examined intervention effects immediately following the intervention. Future research should analyze follow-up data to examine the long-term effectiveness of the intervention.

5. 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 SEL intervention using an active learning approach can positively affect prevocational students. The intervention was especially effective in targeting students' intrapersonal domain and showed the strongest, albeit moderate, effects in the first year. In the interpersonal domain, the intervention potentially functions as a buffer – R&W showed a tendency to lessen the increase in levels of aggression and bullying. Furthermore, the present study showed that intervention effects do not necessarily increase with a longer intervention or a broader socio-ecological implementation; our strongest intervention effects were shown in the first year and when the intervention was implemented with a core team of teachers only.

Our findings have important implications for practice. First, even though the intervention effects were small, these improvements may make important contributions to the students' positive development in the long run. Additionally, our findings suggest that an active learning approach might be a promising intervention method to engage students who are challenging to target in intervention. Hence, it appears that schools aiming to stimulate competencies (especially in the intrapersonal domain) of a challenging group of students can best implement an SEL intervention with active forms of learning. Second, the results suggesting that "less is more" indicate that it might *not* always be worthwhile for schools to invest in implementing a long and extensive intervention. As our results seem to suggest, a shorter intervention period may be sufficient to establish change and the involvement of more teachers and parents may not always have a beneficial effect. Instead, it might be more valuable for schools to invest in high quality training for a subgroup of staff and a high quality of implementation with a narrow socio-ecological focus.

Trial registration

Netherlands Trial Register number NL6371 (old number: NTR6554). Registered on the 3rd of July 2017.

Declaration of Competing Interest

None.

Acknowledgements

This study is financially supported by a grant from The Netherlands Organization for Health Research and Development, grant number 531001106.

Appendix

Table A1
Overview of number of classrooms and trainers per school.

Condition	School	# classrooms	# trainers
Light	1	2	4
	2	5	3
	3	8	2
	4	8	5
Standard	5	10	9
	6	6	6
	7	2	6

(continued on next page)

Table A1 (continued)

Condition	School	# classrooms	# trainers
Plus	8	8	5
	9	4	5
	10	3	4
Control	11	12	–
	12	1	–
	13	7	–

Table A2

Standardized factor loadings of the time points on the slope.

	Factor loadings on slope			
	T1	T2	T3	T4
Proximal outcomes				
Self-control ¹	0	0.66	0.76	0.85
Self-reflection	0	0.54	0.64	0.59
Self-esteem	0	0.09	0.43	0.49
Emotional self-regulation	0	0.41	0.64	0.69
Intrapersonal domain				
Psychological wellbeing	0	0.19	0.61	0.54
Resilience	0	0.19	0.04	0.11
Sexual autonomy ¹	0	0.69	0.84	0.79
Internalizing behavior ¹	0	0.62	0.73	0.79
Interpersonal domain				
Interpersonal relations in the class	0	0.36	0.59	0.69
Externalizing behavior	0	0.28	0.17	0.22
Aggression	0	0.38	0.59	0.55
Bullying	0	0.28	0.47	0.60
Victimization	0	0.36	0.63	0.60

*Note.*¹Variance 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. Factor loading reflect the average change in the observed variables from one time point to the next time point. These factor loadings indicated that students showed generally the largest change in the outcomes from T1 to T2. After T2 the average change leveled off. For instance, students changed on sexual autonomy from T1 to T2 with an increase of 0.69 units, from T2 to T3 with an increase of 0.15 units, and from T3 to T4 with an increase of 0.05 units. Hence, students showed the largest increase in sexual autonomy from T1 to T2, the second largest increase from T2 to T3, and the smallest increase from T3 to T4.

Table A3

Standardized slopes of trajectories of the conditions.

	Light			Standard			Plus			Control		
	Mean slope	SE	p	Mean slope	SE	p	Mean slope	SE	p	Mean slope	SE	p
Proximal outcomes												
Self-control ¹	1.75**	0.73	0.017	1.51	0.78	0.050	1.44	0.74	0.051	1.44	0.76	0.056
Self-reflection ²	0.46	0.55	0.407	0.44	0.66	0.503	0.54	0.59	0.362	0.43	0.56	0.444
Self-esteem	–1.68 [†]	1.27	0.187	–2.13	1.38	0.122	–1.98	1.04	0.056	–2.58	1.48	0.081
Emotional self-regulation	0.24**	0.82	0.769	0.10	1.15	0.934	–0.18	1.30	0.893	–0.10	1.01	0.919
Intrapersonal domain												
Psychological wellbeing	0.08**	1.13	0.947	–0.12	1.43	0.935	–0.26	1.46	0.860	–0.25	1.07	0.812
Resilience	0.55	0.75	0.462	0.61	0.63	0.328	0.52	0.69	0.449	0.63	0.78	0.416
Sexual autonomy ¹	0.73**	0.62	0.240	0.50	0.61	0.408	0.50	0.65	0.446	0.32	0.63	0.613
Internalizing behavior ¹	–0.60***	0.57	0.289	–0.53*	0.61	0.382	–0.32	0.57	0.571	–0.41	0.77	0.594
Interpersonal domain												
Interpersonal relations in the class	–1.64 [†]	1.00	0.102	–1.71	0.96	0.074	–1.93	1.05	0.067	–2.03	1.22	0.096
Externalizing behavior	–0.05	0.30	0.860	–0.04	0.61	0.947	0.03	0.24	0.894	0.03	0.25	0.908
Aggression	0.09*	1.14	0.940	0.31	0.55	0.568	0.54	0.90	0.548	0.56	0.90	0.530
Bullying	2.34 [†]	1.83	0.203	1.65	1.36	0.225	– ³	–	–	2.80	1.67	0.093
Victimization	0.29	0.74	0.695	0.37	1.17	0.750	1.02	3.83	0.790	0.80	7.08	0.910

*Note.*¹Variance 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; ²Variance of the baseline measurement of the outcome variable was fixed to zero for convergence; ³Slope could not be estimated due to

a negative residual variance of the slope. Trajectory intervention condition more beneficial than trajectory control condition with $\dagger p < .10$ * $p < .05$ ** $p < .01$. ^a Trajectory Light condition more beneficial than Standard condition.

References

- Anderson, L. M., Reilly, E. E., Gorrell, S., Schaumberg, K., & Anderson, D. A. (2016). Gender-based differential item function for the difficulties in emotion regulation scale. *Personality and Individual Differences*, *92*, 87–91.
- Bakermans-Kranenburg, M. J., 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.
- 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). Springer Science and Business Media.
- Bishop, D. C., Pankratz, M. M., Hansen, W. B., Albritton, J., Albritton, L., & Strack, J. (2014). Measuring fidelity and adaptation: Reliability of an instrument for school-based prevention programs. *Issues and Solutions in Risky Behavior Prevention Research*, *37*, 231–257.
- Boor-Klip, H. J., Segers, E., Hendrickx, M. M. H. G., & Gillissen, A. H. N. (2016). Development and psychometric properties of the classroom peer context questionnaire. *Social Development*, *25*, 370–389.
- Bronfenbrenner, U. (1979). *The ecology of human development*. Harvard University Press.
- 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.
- 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.
- 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., Heemskerck, I., Leeman, Y., Roeleveld, J., & Van de Venne, L. (2005). Krachtige leeromgevingen in het middelbaar beroepsopleiding: 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 [seks under the age of 25]*. Uitgeverij Eburon.
- De Graaf, I., De Haas, S., Zaagsma, M., & Wijssen, 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.
- 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.
- 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.
- Durlak, J. A. (2016). Programme implementation in social and emotional learning: Basic issues and research findings. *Cambridge Journal of Education*, *4*(3), 333–345. <https://doi.org/10.1080/030574X.2016.1142504>
- Durlak, J. A., Weissberg, R. P., Dymnicki, A. B., Taylor, R. D., & Schellinger, K. B. (2011). The impact of enhancing students' social and emotional learning: A meta-analysis of school-based universal interventions. *Child Development*, *82*, 405–432.
- Feingold, A. (2013). A regression framework for effect size assessments in longitudinal modeling of group differences. *Review of General Psychology*, *17*, 111–121.
- 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). 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., Vogrinic, C., Kastenmüller, A., Frey, D., Heene, M., Wicher, M., & Kainbacher, M. (2011). The bystander-effect: A meta-analytic review on bystander intervention in dangerous and non-dangerous emergencies. *Psychological Bulletin*, *137*, 517–537.
- 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.
- Harter, S. (1988). *Manual for the self-perception profile for adolescents*. University of Denver.
- Kline, R. B. (2005). *Principles and practice of structural equation modeling*. Guilford Press.
- 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. <https://doi.org/10.1016/j.lindif.2011.01.004>
- Kuosmanen, T., Clarke, A., & Barry, M. M. (2019). Promoting adolescents' mental health and wellbeing: Evidence synthesis. *Journal of Public Mental Health*, *18*(1), 73–83. <https://doi.org/10.1108/JPMH-07-2018-0036>
- Lutz, W., Hofmann, S. G., Rubel, J., Boswell, J. F., Shear, M. K., Gorman, J. M., Woods, S. W., & Barlow, D. H. (2014). Patterns of early change and their relationship to outcome and early treatment termination in patients with panic disorder. *Journal of Consulting and Clinical Psychology*, *82*, 287–297.
- Mahoney, J. L., Weissberg, R. P., Greenberg, M. T., Dusenbury, L., Jagers, R. J., Niemi, K., Schlinger, M., Schlund, J., Shriver, T. P., VanAusdal, K., & Yoder, N. (2021). Systemic social and emotional learning: Promoting educational success for all preschool to high school students. *American Psychologist*, *76*(7), 1128–1142. <https://doi.org/10.1037/amp0000701>
- Mertens, E. C. A., Deković, M., Leijten, P., Van Londen, M., & Reitz, E. (2020). Components of school-based interventions stimulating students' intrapersonal and interpersonal domains: A meta-analysis. *Clinical Child and Family Psychology Review*, *23*, 605–631. <https://doi.org/10.1007/s10567-020-00328-y>
- Mertens, E. C. A., Deković, M., 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). Article 36 <https://doi.org/10.1186/s40359-018-0247-y>
- Modecki, K. L., Zimmer-Gembeck, M. J., & Guerra, N. (2017). Emotion regulation, coping, and decision making: Three linked skills for preventing externalizing problems in adolescence. *Child Development*, *88*, 417–426.
- Muthén, B. (2000). Methodological issues in random coefficient growth modelling using a latent variable framework: Application to the development of heavy drinking. In J. Rose, L. Chassin, L. C. Presson, & J. Sherman (Eds.), *Multivariate applications in substance use research: New method for new questions* (pp. 113–140). Erlbaum.
- Muthén, L. K., & Muthén, B. O. (1998–2010). *Mplus user's guide* (6th ed.). Muthén & Muthén.
- 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. <https://doi.org/10.1016.cpr.2015.11.002>
- Park, D., Tsukayama, E., Goodwin, G., Patrick, S., & Duckworth, A. L. (2017). A tripartite taxonomy of character: Evidence for intrapersonal, interpersonal, and intellectual competencies in children. *Contemporary Educational Psychology*, *48*, 16–27.
- Pellegrino, J. W., & Hilton, M. L. (2012). *Education for life and work: Developing transferable knowledge and skills in the 21st century*. National Academy of Sciences.

- Philips, B., & Wennberg, P. (2014). The importance of therapy motivation for patients with substance use disorders. *Psychotherapy, 51*, 555–562.
- Ravens-Sieberer, U., & The European KIDSCREEN Group. (2006). *The KIDSCREEN questionnaires. Quality of life questionnaires for children and adolescents – Handbook*. Pabst Science Publisher.
- Satorra, A., & Bentler, P. M. (2010). Ensuring positiveness of the scaled difference chi-square test statistic. *Psychometrika, 75*, 243–248.
- Sauter, F. M., Heyne, D., Blöte, A. W., 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.
- 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]*. Rijksinstituut voor Volksgezondheid en Milieu (RIVM).
- 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.
- Smit, K., De Brabander, C. J., & Martens, R. L. (2014). Student-centred and teacher-centred learning environment in pre-vocational secondary education: Psychological needs, and motivation. *Scandinavian Journal of Educational Research, 58*(6), 695–712. <https://doi.org/10.1080/00313831.2013.821090>.
- Solberg, M. E., & Olweus, D. (2003). Prevalence estimation of school bullying with the Olweus bully/victim questionnaire. *Aggressive Behavior, 29*, 239–268.
- Stevens, G., & 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). Ridderprint BV.
- Tadić, A., Helmreich, I., Mergl, R., Hautzinger, M., Kohlen, 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.
- Tanner-Smith, E. E., Durlak, J. A., & Marx, R. A. (2018). Empirically based mean effect size distributions for universal prevention programs targeting school-aged youth: A review of meta-analyses. *Prevention Science, 19*, 1091–1101. <https://doi.org/10.1007/s11121-018-0942-1>.
- 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.
- Timmermans, A., Naaijer, H., Keuning, J., & Zijlsling, D. (2017). *Cohortonderzoek COOL5-18 basisrapport meting VO-3 in 2014 [cohort study COOL5-18 report measurement VO-3 in 2014]*. GION (onderwijs/onderzoek).
- 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.
- 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]*. 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]*. 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.