

Implementation of PATHS Through Dutch Municipal Health Services: A Quasi-Experiment

Ferry X. Goossens, Trimbos Institute (Netherlands Institute of Mental Health and Addiction), Utrecht, the Netherlands

Evelien M. J. C. Gooren, Faculty of Psychology and Education, VU University Amsterdam, the Netherlands

Bram Orobio de Castro, Department of Developmental Psychology, Utrecht University, the Netherlands

Kees W. van Overveld, Seminary for Remedial Education, Utrecht, the Netherlands

Goof J. Buijs, The Netherlands Institute for Health Promotion (NIGZ), Woerden, the Netherlands

Karin Monshouwer, Trimbos Institute (Netherlands Institute of Mental Health and Addiction), Utrecht, the Netherlands, and Department of Interdisciplinary Social Science, Utrecht University, the Netherlands

Simone A. Onrust, Trimbos Institute (Netherlands Institute of Mental Health and Addiction), Utrecht, the Netherlands

Theo G. W. M. Paulussen, TNO Quality of Life (Netherlands Organization of Applied Scientific Research), Leiden, the Netherlands

Vol. 6 (2) 2012

Focus:

Evidence-based Developmental Prevention of Bullying and Violence in Europe

Editorial (p. 165)

Guest Editorial: The Future of Research on Evidence-based Developmental Violence Prevention in Europe – Introduction to the Focus Section Manuel Eisner / Tina Malti (pp. 166 – 175)

Violence Prevention in Austrian Schools: Implementation and Evaluation of a National Strategy Christiane Spiel / Petra Wagner / Dagmar Strohmeier (pp. 176 – 186)

Clinical Significance of Parent Training for Children with Conduct Problems Martin Forster / Åsa Kling / Knut Sundell (pp. 187 – 200)

From Clinical-Developmental Theory to Assessment: The Holistic Student Assessment Tool Gil Noam / Tina Malti / Martin Guhn (pp. 201 – 213)

Preventing Child Behavior Problems at Preschool Age: The Erlangen-Nuremberg Development and Prevention Study Friedrich Lösel / Mark Stemmler (pp. 214 – 224)

Introducing, Researching, and Disseminating the Incredible Years Programmes in Wales Judy Hutchings (pp. 225 – 233)

- ▶ **Implementation of PATHS Through Dutch Municipal Health Services: A Quasi-Experiment** Ferry X. Goossens / Evelien M. J. C. Gooren / Bram Orobio de Castro / Kees W. van Overveld / Goof J. Buijs / Karin Monshouwer / Simone A. Onrust / Theo G. W. M. Paulussen (pp. 234 – 248)

Effectiveness of a Universal School-Based Social Competence Program: The Role of Child Characteristics and Economic Factors Tina Malti / Denis Ribeaud / Manuel Eisner (pp. 249 – 259)

The Impact of Three Evidence-Based Programmes Delivered in Public Systems in Birmingham, UK Michael Little / Vashti Berry / Louise Morpeth / Sarah Blower / Nick Axford / Rod Taylor / Tracey Bywater / Minna Lehtonen / Kate Tobin (pp. 260 – 272)

Successful Bullying Prevention Programs: Influence of Research Design, Implementation Features, and Program Components Bryanna Hahn Fox / David P. Farrington / Maria M. Ttofi (pp. 273 – 282)

Tackling Cyberbullying: Review of Empirical Evidence Regarding Successful Responses by Students, Parents, and Schools Sonja Perren / Lucie Corcoran / Helen Cowie / Francine Dehue / D'Jamila Garcia / Conor Mc Guckin / Anna Sevcikova / Panayiota Tsatsou / Trijntje Völlink (pp. 283 – 292)

KiVa Antibullying Program: Overview of Evaluation Studies Based on a Randomized Controlled Trial and National Rollout in Finland Christina Salmivalli / Elisa Poskiparta (pp. 293 – 301)

Knowing, Building and Living Together on Internet and Social Networks: The ConRed Cyberbullying Prevention Program Rosario Ortega-Ruiz / Rosario Del Rey / José A. Casas (pp. 302 – 312)

Empowering Students Against Bullying and Cyberbullying: Evaluation of an Italian Peer-led Model Ersilia Menesini / Annalaura Nocentini / Benedetta Emanuela Palladino (pp. 313 – 320)

Identity Centrality and In-Group Superiority Differentially Predict Reactions to Historical Victimization and Harm Doing Rezarta Bilali (pp. 321 – 337)

A Farewell to Innocence? African Youth and Violence in the Twenty-First Century Charles Ugochukwu Ukeje / Akin Iwilade (pp. 338 – 350)

Open Section



Implementation of PATHS through Dutch Municipal Health Services: A Quasi-Experiment

Ferry X. Goossens, Trimbos Institute (Netherlands Institute of Mental Health and Addiction), Utrecht, the Netherlands

Evelien M. J. C. Gooren, Faculty of Psychology and Education, VU University Amsterdam, the Netherlands

Bram Orobio de Castro, Department of Developmental Psychology, Utrecht University, the Netherlands

Kees W. van Overveld, Seminary for Remedial Education, Utrecht, the Netherlands

Goof J. Buijs, The Netherlands Institute for Health Promotion (NIGZ), Woerden, the Netherlands

Karin Monshouwer, Trimbos Institute (Netherlands Institute of Mental Health and Addiction), Utrecht, the Netherlands, and Department of Interdisciplinary Social Science, Utrecht University, the Netherlands

Simone A. Onrust, Trimbos Institute (Netherlands Institute of Mental Health and Addiction), Utrecht, the Netherlands

Theo G. W. M. Paulussen, TNO Quality of Life (Netherlands Organization of Applied Scientific Research), Leiden, the Netherlands

Only a limited number of effectiveness studies have been performed to study the benefits of efficacious behavior problems prevention programs for children when implemented through national health service systems. This study uses a quasi-experimental design to test the effectiveness of the school-based PATHS prevention program (Providing Alternative Thinking Strategies) when implemented through Dutch municipal health services by health promotion professionals. A sample of 1,294 children was followed for two years: 674 children attending nine schools providing PATHS and 620 children in nine comparison schools. We hypothesized finding an intervention effect of PATHS in terms of a significant reduction in teacher- and student-rated externalizing and internalizing problem behaviors, and a significant improvement in teacher-, student-, and peer-rated social skills and emotional skills. In fact, the results show low levels of program implementation and no intervention effects on problem behavior or social and emotional skills, suggesting that it is hard to reproduce positive intervention effects where an efficacious social-emotional prevention program is implemented through a national health service. More research is needed on the specific conditions required to implement efficacious programs effectively.

It has been shown that school-based prevention programs aimed at social and emotional learning can be efficacious (Durlak et al. 2011). Usually these programs are evaluated in efficacy trials characterized by optimal conditions, such as well-trained and carefully supervised intervention personnel, and ample resources, and by involvement of the program developers in the implementation process and research (Flay et al. 2005). The better the program im-

plementation and the greater the program fidelity, the stronger the effects are (Beelmann and Raabe 2009; Eisner 2009; Salmivalli, Kaukiainen, and Voeten 2005; Wilson and Lipsey 2007). However, efficacy trials have high internal validity but weaker external validity, for example actual utilization under normal community conditions (Weisz et al. 1995). Very few effectiveness studies have examined the benefits of efficacious prevention programs on a large scale

Acknowledgements: This study was funded by grant number 2430.0031 from the Dutch Health Care Research Organization (ZonMw). Furthermore, we are grateful to all who participated in this study (health promotion professionals, teachers, children, and their parents).

Authors' contributions: F. Goossens carried out the study, analyzed the data, and wrote the manuscript. E. Gooren coordinated the acquisition of data, cleaned the data, and revised the manuscript. B. Orobio de Castro revised the manuscript thoroughly. C. van Overveld (who is employed at the Dutch licensee of PATHS) trained and supported the municipal health

services and revised the manuscript. G. Buijs contributed to the design of the study, developed the implementation materials, and revised the manuscript. K. Monshouwer and S. Onrust revised the manuscript. T. Paulussen designed the study and revised the manuscript thoroughly. All authors have given final approval of the version to be published.

(Dodge 2009). Given that to date few efficacious prevention programs have been successfully implemented on a large scale (Elliot and Mihalic 2004), and their population impact on the prevention of problem behavior is very modest (Dodge 2009), more research on this topic is needed.

1. The PATHS Program

PATHS is a comprehensive school-based prevention program that aims to enhance social and emotional competence and reduce behavioral and mental problems of children from kindergarten to sixth grade. PATHS is based on five conceptual models that form a theoretical framework focusing on a broad range of protective and risk factors for internalizing and externalizing problem behavior. The curriculum provides one to three sessions each week focusing on self-control, emotional understanding, positive self-esteem, relationships, and interpersonal problem-solving skills. Lessons include didactic instruction, role playing, class discussion, modeling by teachers and peers, social and self-reinforcement, and worksheets. The lessons are grouped in three major units: readiness and self-control, feelings and relationships, and interpersonal cognitive problem-solving (Greenberg and Kusché 1993).

1.1. PATHS in the Netherlands

For a couple of decades now, the efficacious school-based prevention program PATHS (Providing Alternative THinking Strategies) (Domitrovich 1999; Kusché and Greenberg 1994) has been implemented in the Netherlands by a small group of dedicated professionals employed at the Dutch national licensee, which has directly trained staff at three hundred schools over the last twenty-five years. However, with approximately seven thousand elementary schools in the Netherlands this is no more than 5 percent of the potential, and its impact on society is therefore probably still limited. The present study examines the effectiveness of PATHS when implemented through health promotion professionals from Dutch municipal health services (MHSs). In this implementation strategy, the national licensee trained MHS health promotion professionals who in turn trained teachers. As it is the daily job of such professionals to implement school-based prevention programs on topics like bullying, alcohol, obesity, sexually transmitted diseases,

and social emotional competence, it was assumed that they could also be a useful link in implementing PATHS.

1.2. PATHS Efficacy and Effectiveness

PATHS has been studied in a number of trials within a variety of populations, including children in regular education, children at risk of behavior problems, children in special education, and children with hearing impairments. As the current study concerns the implementation of PATHS for children in regular education, we concentrate on studies focusing on that target group, as well as studies that examine the implementation process.

The results of the first PATHS study in regular education indicated that the intervention was effective in improving grade 2 and 3 children's range of vocabulary and fluency in discussing emotional experiences, their efficacy beliefs regarding the management of emotions, and their developmental understanding of some aspects of emotions (Greenberg et al. 1995). A second, larger study of 329 second and third graders showed that the intervention promoted inhibitory control, verbal fluency, and diminished internalizing and externalizing problem behaviors (Riggs et al. 2006). A third study, with 246 pre-school children, showed that children exposed to PATHS intervention had higher emotion knowledge skills and were more socially competent than peers (Domitrovich, Cortes, and Greenberg 2007). In all these studies, the program developers were involved in the research, and the level of support was high (i.e. teachers received monthly or even weekly consultation from the project staff to enhance the quality of implementation).

These intensive procedures may have contributed significantly to the implementation quality and the effects found. This assumption is supported by the results of an effectiveness trial involving 350 first graders in six inner-city public schools in a high-risk urban community in the United States (Kam, Greenberg, and Walls 2003). In this study, significant intervention effects were found only in schools where both implementation quality and support from the principal were high. Another study showed that it was not the number of PATHS sessions received, but the quality of these sessions that positively influenced the outcomes (CPPRG 1999).

Interestingly, PATHS with less intense support was also studied. A Dutch study concerning the effectiveness of PATHS for boys with severe aggressive behavior problems showed positive effects on proactive and reactive aggression (Louwe et al. 2007), in a context where training and support were provided directly by the national licensee. However, a negative trend was found for children in special needs schools, where implementation quality was substantially lower than in other schools (Louwe et al. 2007b). A recent study in Zurich showed intervention effects on teacher- and parent-rated aggressive behavior, and teacher-rated ADHD, but no significant positive effects for nine other externalizing and internalizing outcomes. The teachers who implemented PATHS received two days' training, and local coaches were trained to visit the classes and provide feedback to the teachers (Malti, Ribeaud, and Eisner 2011). Finally, the intervention was studied in ten US public elementary schools (SCDRC 2010). Of the twenty child-level outcomes, none was significant. The authors suggest that the lack of positive findings was probably caused by the control condition not being a non-treatment condition, but a standard practice condition including schools that use other social and character development activities. In sum, PATHS has been shown to be efficacious and potentially effective, but effectiveness depends to a large extent on the implementation conditions.

1.3. Aims and Hypotheses

The present study explored the effectiveness of PATHS when implemented through a regular health service system, i.e. Dutch municipal health services. We hypothesized finding an intervention effect of PATHS in terms of a significant reduction in teacher- and student-rated externalizing and internalizing problem behaviors, and a significant improvement in teacher-, student-, and peer-rated social and emotional skills. We further hypothesized finding a larger effect in classes with higher implementation quality, and in classes with higher implementation quantity.

2. Method

The effectiveness of PATHS was assessed using a quasi-experimental design with an intervention ($n = 674$) and a waiting list comparison ($n = 620$) condition, each containing nine regular elementary schools. The intervention

effects were measured during the first two years of PATHS implementation.

2.1. Implementation Strategy

First, the national project team recruited MHSs to participate in the study. All thirty Dutch MHSs were approached by e-mail, and three were willing and able to participate in the study. At each participating MHS, three health promotion professionals were recruited. These nine professionals were facilitated in their task of training and supporting teachers by means of 1) a two-day PATHS course run by the Dutch PATHS licensee, 2) a manual, a model recruitment letter, and a PATHS brochure for recruiting schools, 3) a pre-implementation teacher training protocol, and 4) follow-up support by the Dutch PATHS licensee.

Second, the health promotion professionals sent all principals of elementary schools in their region the PATHS model recruitment letter and the PATHS brochure. If principals and teachers expressed the intention to participate in the study, an implementation agreement was signed between the school and the MHS.

Third, schools were allocated randomly to either the intervention condition or the waiting list comparison condition, with four exceptions. In order to maximize ecological validity, we refrained from randomizing four of the schools. All the schools were very willing to implement PATHS, but two preferred to start two years later for organizational reasons and two were willing to participate only if they could start directly. These requests were complied with.

Fourth, under supervision of the national licensee, the health promotion professionals provided a two-day pre-implementation training course for the teachers in the intervention condition. Directly after this course, all teachers in the intervention condition started giving their PATHS lessons.

Fifth, in the course of each school year the teachers received a half-day booster session. Additionally, the PATHS coordinators received a half-day support session. Each school designated a staff member as PATHS coordinator with the task of supporting the implementation in their school by delivering supervision and feedback to the teachers.

Sixth, each school in the experimental condition organized, in co-operation with the MHS, an information meeting for parents. Parents also received frequent written information about the content of the program.

2.2. The Health Promotion Context

The PATHS strategy had previously been applied successfully to the implementation of a school-based sex education program by MHSs in the Netherlands (Wiefferink et al. 2005). The strategy had a positive impact on teachers' extent of use, as well as their curriculum-related beliefs. Moreover, implementing school-based prevention programs through MHS health promotion professionals is a common strategy for the implementation of school-based prevention programs in the Netherlands. So although this strategy clearly differs from the more intensive procedures advised by the developers – and as generally used in PATHS trials – it is a good example of how prevention programs are implemented in the Dutch context of school-based health promotion and thus complies with our aim of exploring the effectiveness of PATHS when implemented through a regular health service system.

2.3. The Intervention

The version of PATHS used in this study (SvO 2005) consisted of 161 lessons of 20 to 30 minutes, spread over the eight years of elementary school. It was an update of a Dutch translation of the US curriculum for regular schools that had already been in use for several years in the Netherlands, supplemented by a translation of the pre-school PATHS program (Domitrovich et al. 1999). During the study, all children in the intervention condition received the PATHS program for two years, consisting of approximately forty PATHS lessons. As children in the higher grades did not start the PATHS lessons from kindergarten, they received extra lessons to inform them about the story line and basic principles of the PATHS program. Nevertheless, as PATHS is a cyclical program, all major units are discussed each school year. The study proposal was approved by the medical ethical committee, which stated that passive informed consent was adequate.

2.4. Data Collection

Teacher assessments and student assessments were conducted in both conditions: at the start (T0) and the end

(T1) of the first year, and at the start (T2) and the end (T3) of the second year. Each teacher received his/her questionnaires by post and was asked to fill out ten to thirty questionnaires each wave. Child questionnaires were completed in face-to-face interviews with the three youngest cohorts that were followed in this study (kindergarten, grade 1 and grade 3) and by means of a self-report questionnaire for the oldest cohort (grade 5). The student assessments lasted 20 to 30 minutes per child per wave and were conducted by approximately seventy-five trained graduate psychology students who did not know the intervention condition of the school. Questionnaires concerning implementation quantity were sent to all teachers in the intervention condition four times a year. Implementation quality was measured twice, at the end of the first year and at the end of the second year. Children received a gift worth +/- €0.75 at each measurement point. To incentivize school participation, training and materials were provided free of charge (+/- €4,000).

2.5. Measures

2.5.1. Problem Behavior

The Problem Behavior at School Interview (PBSI) (ErasmusMC 2000) is a forty-two-item questionnaire enquiring about externalizing problems: attention deficit and hyperactivity (ADH), oppositional defiant disorder (ODD), conduct problems, and relational aggression, and internalizing problems: anxiety and depression. Teachers rated the child's behavior on a five-point Likert scale ranging from 0 (never applicable) to 4 (often applicable). Cronbach's α in this study varied between .78 and .92. The convergent validity of the PBSI was found to be good as indicated by the correlations between the PBSI and the Teacher's Report Form (Achenbach 1991), which were .75 for externalizing behavior and .55 for internalizing behavior (Witvliet et al. 2010). This measure was used for all cohorts.

The Social Experience Questionnaire – Teacher Report (SEQ-T) (Crick and Grotpeter 1996) was used to measure relational victimization, physical victimization, and prosocial behavior on a 5-point Likert scale ranging from 0 (never applicable) to 4 (often applicable). The relational victimization and physical victimization scales included three items and the prosocial behavior scale four items. Cronbach's α was .87 (relational victimization), .85 (physi-

cal victimization), and .75 (prosocial behavior). This measure was used for all four cohorts.

The twenty-four-item Dimensions of Depression Profile for Children (DDPC) (four-point Likert format) was used to measure the level of depression in the oldest cohort (Harter and Nowakowski 1987). It contained four subscales: depressed mood ($\alpha = .69$), self-blame ($\alpha = .59$), low energy/interest ($\alpha = .75$), low global self-worth ($\alpha = .77$), and one total score ($\alpha = .85$).

2.5.2. Social and Emotional Skills

Peer social preference was obtained using peer nominations of like most and like least, as described by Coie, Dodge, and Coppotelli (1982). Children (cohorts 1 and 4) were asked to nominate an unlimited number of classmates that they liked most and that they liked least. Each child could therefore be nominated by each classmate as “liked most,” “liked least,” or gain no score. For each child, the liked most as well as the liked least nominations were summed and divided by the number of children in the class minus one (self-nomination was not allowed). The standardized liked least score was then subtracted from the standardized liked most score to generate a social preference score. Social preference is generally regarded as a reliable and valid measure of sociometric status (Cillessen and Mayeux 2004; Rubin et al. 2006).

The teacher-based Preschool and Kindergarten Behavior Scale (PKBS) (Merrel 1996) used for cohorts 1 and 2 is designed to assess social skills and problem behaviors in children aged three to six years. In this study, the social skills scale, which has three subscales, was used. The social cooperation scale included twelve items reflecting behaviors and characteristics deemed important in following instructions from adults, cooperating and compromising with peers. The social interaction scale included 11 items reflecting behaviors and characteristics deemed important in gaining and maintaining the acceptance and friendship of others. The social independence scale included eleven items reflecting behaviors and characteristics deemed important in achieving social independence within the domain of the peer group. For all items, responses were based on a four-point scale. The internal consistency of all three subscales was high (respectively $\alpha = .89$, $\alpha = .87$, $\alpha = .86$).

The Head Start Competence Scale (HSCS) (Domitrovich, Cortes, and Greenberg 2001) used for cohorts 1 and 2 is a twelve-item measure of children’s social and emotional skills reflecting interpersonal relationships and emotion regulation. Teachers were asked to indicate on a four-point scale how well each item on the scale described the child. Internal consistency was high ($\alpha = .95$).

Children’s emotional awareness scores were obtained for cohort 1 and 4 using the Levels of Emotional Awareness Scale for Children (LEAS-C), which assesses the complexity of children’s emotional awareness (Bajgar et al. 2005). It contains twelve interpersonal scenarios featuring the child and another person. After each scenario had been presented, the children were asked to describe their own feelings and those of the other person. For each scenario, the self and other response was rated on a scale ranging from 0 to 4. The higher of the scores for the self and the other was taken as the total score for each scenario. In cases where both the self and the other score were 4, a total score of 5 was awarded. The total scores were summed across the scenarios. Cronbach’s α ranged from .89 to .92 over the assessments.

The child-based Difficulties in Emotion Regulation Scale (DERS) (Gratz and Roemer 2004) was used to measure deficits of emotion regulation. It contains thirty-six questions (five-point Likert scale) in six scales: non-acceptance of emotional responses ($\alpha = .73$), difficulties engaging in goal-directed behavior ($\alpha = .82$), impulse control difficulties ($\alpha = .80$), lack of emotional awareness ($\alpha = .78$), limited access to emotion regulation strategies ($\alpha = .74$), and lack of emotional clarity ($\alpha = .61$). This measure was used for the oldest cohort only.

To measure the children’s affective sharing of others’ emotions, a short ten-item version of Bryant’s Empathy Index was used (Bryant 1982; De Wied et al. 2007). This child-based measure was used for the youngest and oldest cohort only. Cronbach’s α was .68.

2.5.3. Implementation Quality

Implementation quality was operationalized as “conceptual use of the program” i.e. to what extent do teachers act according to the PATHS basic principles. Teachers received a

list of ten questions describing daily classroom situations. For each situation, teachers could choose one of four answers that described how they would react in this specific situation. The answers varied from most desired reaction according to the PATHS basic principles (score = 4) to least desired reaction (score = 1). All scores were averaged, resulting in a mean score for the first year and a mean score for the second year. For each class, we calculated one mean score (range 1 – 4) from these two scores.

2.5.4. Implementation Quantity

To measure implementation quantity (i.e. completeness), teachers completed a monthly log describing all the required lessons and elements thereof, recording whether they completed each specific element of each lesson. For both the first year and the second year, intervention completeness was assessed as the completed proportion of all prescribed activities for that year. We summed these two proportions to obtain one total score (range 0 – 2).

2.5.5. Covariates

The Peabody Picture Vocabulary Test – Third Edition (PPVT-III) (Dunn and Dunn 1997) was used to measure verbal ability. This assessment is a well-known and widely used measure of children's receptive vocabulary. A Word Comprehension Quotient (WCQ) score, using age-appropriate norms, was calculated from the raw total number of correct answers. The internal consistency of the PPVT-III standard scores ranged from .92 to .98. Child verbal ability was included as a covariate because of its potential to affect children's performance in the testing situation.

2.6. Statistical Analyses

The data were analyzed using Stata version 11.1 (StataCorp 2009) over all 1,294 students in accordance with the intention-to-treat principle. Missing data were handled through Full Information Maximum Likelihood estimation (FIML). ANOVAs and Chi-square analyses were conducted to check whether there was a balanced distribution of student characteristics and outcome variable values across the two conditions at baseline ($p < .05$).

The data in this study is hierarchically structured, i.e. measurement waves (T0, T1, T2, and T3) are nested within

students, students are nested within classes, and classes are nested within schools. In such a hierarchical structure, standard statistical formulas will underestimate the sampling variance, and consequently lead to significance tests with an inflated alpha level (type 1 error rate). Multilevel models are specifically geared toward the statistical analyses of nested or clustered data (Hox 2010). In the present study we use multilevel mixed-effects linear regression analyses (Twisk 2006) to test, for each of the outcome variables, whether changes over time varied as a function of the intervention.

For each outcome variable we analyzed the whole development of each outcome variable over time in one analysis by calculating the three change scores (T0 – T1, T1 – T2, and T2 – T3) for each variable. However, as there were baseline differences between the conditions we had to correct the change scores and used the analysis of covariance combination approach as described by Twisk and De Vente (2008) for that purpose. In this analytical approach, the change between baseline and the first reassessment is corrected for the baseline value by subtracting the individual baseline value from the first individual change score. For computational reasons only, a correction was also made for the remaining two change scores. These three adjusted change scores per variable were included in a longitudinal multilevel mixed-effects linear regression model. The resulting β coefficient represents the intervention effect over the whole period, i.e. from T0 to T3.

In order to assess short-term effects, the analyses were also performed for the T0 – T1 period alone (i.e. the models included only the T0 – T1 adjusted change score). In addition, we tested for possible moderating effects of group, gender, program completeness, and conceptual use by including interaction terms between these variables and condition in the models (each interaction term was tested in separate models). Because of multiple testing (27 outcomes), the level of statistical significance was set at $p < .01$ in all tests.

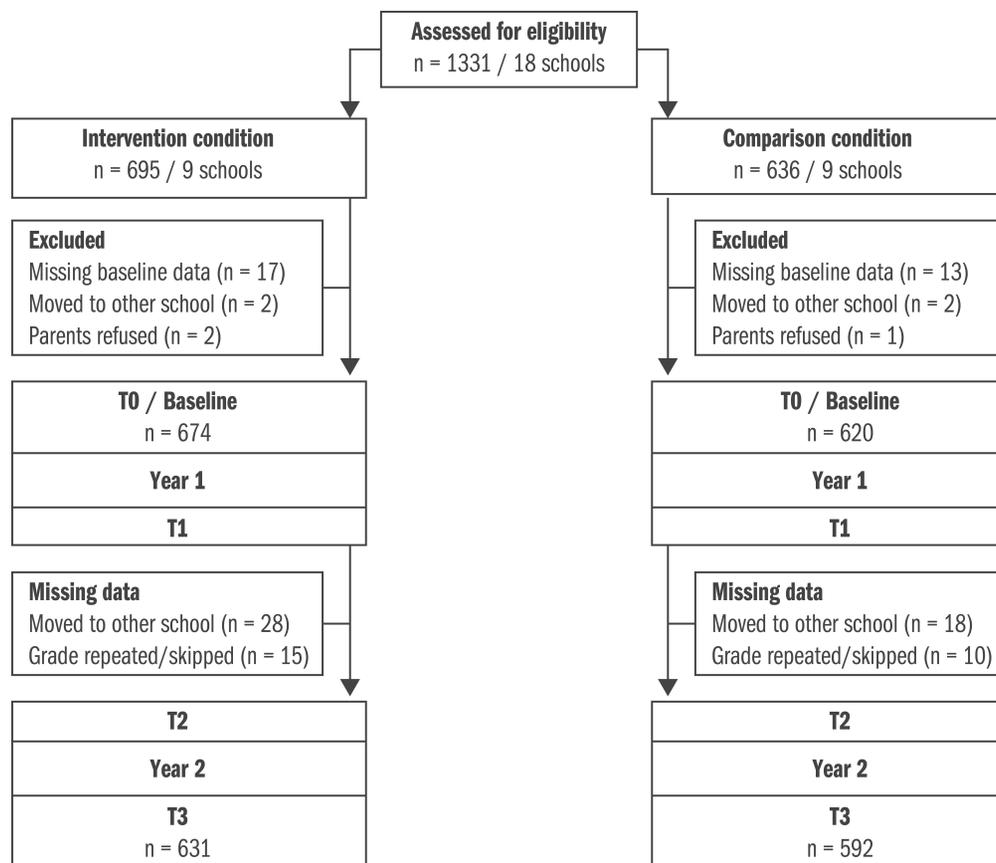
2.7. The Sample

In total, 1,331 children (five to 11 years old) from kindergarten and elementary school grades 1, 3, and 5 were eligible for inclusion (Figure 1). The 18 participating schools were located in rural areas and provincial towns in the

western (Noord-Holland) and eastern part (Gelderland and Overijssel) of the Netherlands. The baseline data for 30 children were missing or incomplete. Four children were excluded because they moved to another school soon after the baseline assessment, and three children's parents refused to allow their children to participate in the study. Therefore the baseline sample included 1,294 children. Of these, 65 changed school during the study. In accordance

with the intention-to-treat principle, we sought to collect data from these children by sending a questionnaire to their home address, asking the parents to fill out the forms. We were able to collect data from nineteen of them. We were also able to collect reassessment data for ten of the thirty-five children who repeated or skipped a grade. Participation in our study was high (97 percent), and the attrition rate low (5 percent at last assessment).

Figure 1: Participation flowchart



Age, gender, ethnicity, and verbal ability did not differ between the intervention and the comparison group (Table 1). Significant baseline differences ($p < .05$) were present with respect to the level of ADHD ($F(1,1292) = 10.443, p < .001$), ODD ($F(1,1292) = 20.896, p < .001$), conduct problems ($F(1,1292) = 30.338, p < .001$), relational aggression ($F(1,1292) = 60.891, p < .001$), anxiety ($F(1,1293) = 11.400, p < .001$), depression ($F(1,1292) = 10.161, p < .001$), relational victimization ($F(1,1292) = 45.082, p < .001$),

physical victimization ($F(1,1292) = 45.594, p < .001$), prosocial behavior ($F(1,1292) = 5.098, p < .05$), low energy ($F(1,1292) = 4.924, p < .05$), social interaction ($F(1,1292) = 4.722, p < .05$), social independence ($F(1,1292) = 4.725, p < .05$), and social and emotional skills ($F(1,1292) = 10.691, p < .01$). In general, the levels of problem behavior were higher in the intervention condition, and the levels of social and emotional skills were lower (Table 2). These differences were statistically corrected for in the analyses.

Table 1: Socio-demographic variables at baseline by cohort

		Intervention condition	Comparison condition
Cohort 1	n	158	166
(Kindergarten)	Male	56.3 percent	52.4 percent
	Age	5.4 years (SD = .36)	5.5 years (SD = .31)
	Dutch	98.6 percent	92.4 percent
Cohort 2	n	159	151
(Grade 1)	Male	52.2 percent	46.4 percent
	Age	6.5 years (SD = .38)	6.5 years (SD = .42)
	Dutch	96.7 percent	95.1 percent
Cohort 3	n	173	152
(Grade 3)	Male	49.1 percent	56.6 percent
	Age	8.6 years (SD = .53)	8.5 years (SD = .50)
	Dutch	92.1 percent	93.2 percent
Cohort 4	n	184	151
(Grade 5)	Male	44 percent	47 percent
	Age	10.5 years (SD = .42)	10.6 years (SD = .45)
	Dutch	94.4 percent	94.5 percent

3. Results

Table 2 shows the outcome variables over time for both conditions. A decline in the PBSI, SEQ-T (except prosocial behavior scale), DDPC, and DERS scores represents a decrease in these problem behaviors/skills. An increase in the PEER, PKBS, HSCS, LEAS, EMPATHY, and prosocial behavior scale (SEQ-T) scores represents an improvement in these skills.

Table 2: Mean scores (M) and standard deviations (SD) at baseline and reassessments for all four cohorts combined

Measure	Scale	Intervention condition					Comparison condition				
		T0		T1	T2	T3	T0		T1	T2	T3
		M	SD	M	M	M	M	SD	M	M	M
PBSI	ADH*	1.26	(.81)	1.00	1.07	1.11	1.11	(.84)	.97	1.10	1.07
	ODD*	1.11	(.75)	.96	.98	.98	.91	(.79)	.88	.88	.89
	Conduct problems*	.62	(.55)	.51	.54	.56	.45	(.52)	.45	.41	.42
	Relational aggression*	1.12	(.73)	.94	1.05	1.04	.80	(.76)	.82	.89	.86
	Anxiety*	1.29	(.71)	1.10	1.20	1.15	1.16	(.72)	1.09	1.13	1.01
	Depression*	.92	(.64)	.86	.95	.92	.80	(.69)	.81	.84	.82
SEQ-T ¹	Relational victimization*	.87	(.69)	.72	.	.79	.62	(.63)	.61	.	.62
	Physical victimization*	.57	(.55)	.	.45	.50	.37	(.50)	.	.31	.29
	Prosocial behavior*	2.68	(.66)	.	2.78	2.78	2.76	(.66)	.	2.76	2.83
DDPC	Depressed mood	10.63	(2.99)	10.37	10.06	9.98	10.35	(3.00)	9.93	9.82	9.88
	Self blame	15.31	(2.74)	14.43	14.17	14.44	15.01	(3.09)	14.63	14.28	14.27
	Low energy*	11.08	(3.46)	10.68	10.45	10.53	10.25	(3.25)	10.40	9.78	10.19
	Low self-worth	10.23	(3.30)	9.71	9.71	9.79	9.70	(3.25)	9.44	9.43	9.50
	Total depression	47.25	(9.14)	45.24	44.39	44.73	45.31	(8.89)	44.39	43.31	43.90
PEER	Peer nominations	.11	(.27)	.15	.18	.15	.13	(.29)	.17	.18	.17
PKBS	Social cooperation	31.65	(4.35)	32.59	32.35	32.24	32.25	(4.00)	32.69	32.02	32.32
	Social interaction*	24.45	(5.02)	25.76	25.44	26.07	24.54	(5.22)	25.46	24.79	25.55
	Social independence*	27.95	(4.11)	29.07	28.58	28.57	28.35	(4.16)	28.91	28.15	28.87
HSCS	Social and emotional skills*	21.61	(7.39)	23.63	23.55	23.98	24.47	(7.06)	25.08	24.45	25.48
LEAS	Emotional awareness	25.04	(11.27)	31.47	34.03	35.90	23.84	(10.69)	28.10	30.15	30.98
DERS	Lack emotional awareness	21.31	(5.67)	22.50	23.60	24.51	20.93	(5.50)	22.49	22.20	22.43
	Non-acceptance of emotional responses	11.91	(4.46)	10.86	10.14	9.91	11.61	(4.63)	10.50	10.04	10.7
	Difficulty in goal-directed behavior	13.19	(5.14)	12.73	11.87	11.30	13.19	(5.33)	12.13	11.54	11.51
	Impulse control difficulties	13.16	(5.09)	12.69	12.49	11.94	13.60	(5.67)	12.17	11.57	11.90
	Limited access strategies	17.42	(5.59)	17.20	15.99	15.83	17.60	(5.97)	15.97	15.38	15.76
	Lack of emotional clarity	15.52	(4.32)	14.54	14.64	14.60	15.35	(4.16)	14.19	14.06	14.44
EMPATHY	Empathy	5.20	(2.36)	5.21	4.91	4.94	4.99	(2.32)	5.14	5.19	5.24

1 To limit the number of questions teachers had to answer each wave, this questionnaire was sent at two instead of three reassessments.

* Significant difference between intervention condition and comparison condition at baseline ($p < .05$).

First, intervention effects were examined by longitudinal multilevel mixed-effects linear regression analyses (Table 3). We combined all four cohorts and tested whether changes over time varied as a function of the intervention.

This first set of analyses resulted in a positive intervention effect found on the emotional awareness scale. Note that this is the only significant intervention effect out of twenty-seven tests.

Table 3: Intervention effect (β coefficient) over all four assessments for all four cohorts combined

Measure	Scale	β	SE	z	ρ	95 percent CI	
PBSI	ADH	-.022	.029	-0.74	0.458	-.079	.035
	ODD	-.012	.028	-0.42	0.675	-.067	.043
	Conduct problems	.008	.020	0.41	0.681	-.032	.048
	Relational aggression	.005	.030	0.18	0.857	-.054	.065
	Anxiety	.017	.032	0.53	0.599	-.045	.078
	Depression	.005	.031	0.17	0.868	-.055	.065
SEQ-T	Relational victimization	.025	.039	0.64	0.522	-.051	.101
	Physical victimization	.102	.071	1.45	0.148	-.036	.241
	Prosocial behavior	.074	.075	0.99	0.323	-.073	.222
DDPC	Depressed mood	.016	.168	0.09	0.925	-.314	.345
	Self blame	.114	.173	0.66	0.509	-.225	.454
	Low energy	-.032	.182	-0.18	0.860	-.388	.324
	Low self-worth	.090	.188	0.48	0.632	-.279	.460
	Total depression	.104	.473	0.22	0.825	-.823	1.032
PEER	Peer nominations	.004	.012	0.33	0.738	-.019	.027
PKBS	Social cooperation	-.065	.174	-0.37	0.710	-.280	.411
	Social interaction	.466	.295	1.58	0.115	-.113	1.046
	Social independence	.187	.276	0.68	0.498	-.353	.727
HSCS	Social and emotional skills	.111	.328	0.34	0.736	-.532	.753
LEAS	Emotional awareness	1.186	.393	3.02	0.003 *	.416	1.957
DERS	Lack emotional awareness	.694	.305	2.27	0.023	.095	1.292
	Non-acceptance of emotional responses	-.059	.245	-0.24	0.809	-.539	.421
	Difficulty in goal-directed behavior	-.139	.277	-0.50	0.616	-.683	.404
	Impulse control difficulties	.135	.297	0.46	0.648	-.446	.717
	Limited access strategies	.123	.332	0.37	0.712	-.528	.774
	Lack of emotional clarity	.141	.298	0.47	0.635	-.442	.725
EMPATHY	Empathy	-.200	.114	-1.76	0.079	-.422	.023

* Significant difference between intervention condition and comparison condition ($p < .01$).

Note: A minus sign (-) before the ρ of the PBSI, SEQ-T (except prosocial behavior scale), DDPC, and DERS represents a decrease in these problem behaviors/skills in the intervention condition over time, relative to the control condition.

Second, we checked whether there was a positive effect in the first year of implementation. We performed the same analyses as above, but limited them to the T0 – T1 adjusted change score. We found no intervention effects on any of the variables in the first year.

Third, we tested whether intervention effects were moderated by grade. We found three significant moderations on the prosocial scale. Effects were stronger for children in kindergarten in the intervention condition and weaker for children in grades 1 and 5 in the intervention condition. The analyses also showed that the effects for the LEAS were stronger for children in kindergarten. Otherwise, grade did not moderate outcomes.

Fourth, we checked whether there was a difference in effect for boys and girls by adding an interaction term (sex x condition). We did not find any gender-related difference.

Fifth, we tested whether intervention effects depended on the level of program completeness. Mean completeness was 50 percent in the first year (SD = 23 percent) and 49 percent in the second year (SD = 24 percent). We summed these two proportions to obtain one score (mean = .99, SD = .42, range .16 – 1.70), and tested whether this score was related to intervention effects. The analyses showed that program completeness did not moderate intervention effects.

Sixth, we tested whether intervention effects depended on implementation quality. The mean level of “conceptual use” was around 3.05 in the first year (SD = .27, range 2.30 – 3.75) and 3.07 in the second year (SD = .37, range 2.20 – 3.70). We calculated a mean of these two scores for each class and tested whether this score was related to intervention effects. Conceptual use did not moderate intervention effects.

4. Discussion

The purpose of this study was to examine the effectiveness of the efficacious school-based prevention program PATHS when implemented through Dutch municipal health services. In contrast to other studies, the intervention was implemented by health promotion professionals from a regular health service system rather than dedicated PATHS

trainers. As health promotion professionals implement school-based prevention programs on a regular basis, this linkage group could in the long run potentially support the national licensee in the dissemination of PATHS. However, virtually no intervention effects were found in this study.

It seems unlikely that the underlying theoretical model of the PATHS intervention can explain the lack of intervention effects. Previous research has demonstrated that PATHS is effective when delivered adequately. The implementation process therefore seems a more likely reason. In the present study, there appears to be a large gap between the intended intervention and the intervention participants actually received. The implementation strategy most probably affected the teacher support negatively. This resulted in low program completeness and probably affected some other implementation variables negatively as well. This finally resulted in poor intervention outcomes.

Although we cannot exactly identify the (implementation) variables that moderated the outcomes, we can conclude that the tested PATHS implementation strategy is not a recipe for effective prevention of problem behavior on a large scale.

4.1. Strengths and Limitations

This is the first study to test the effectiveness of PATHS when implemented through a regular health service, and sought to explore how problem behavior could be prevented on a large scale. Furthermore, this study had very high participation (97 percent) and low attrition (5 percent). We also had a fairly large sample that gave us enough power to detect small effects, even for the four grades/cohorts separately. It must further be recognized that the results are fairly comprehensive, and the measures included constructs of social and emotional skills and problem behavior of both teachers and children, on four occasions.

However, the study also suffers from important limitations. First, we did not randomize all schools. This may have caused baseline differences between the two conditions. However, we corrected for these differences with a thorough statistical technique specially developed for this kind

of problem (Twisk and De Vente 2008). A second limitation is the limited duration of implementation. As PATHS is an eight-year curriculum, two years of implementation may have been insufficient to achieve the desired change. However, previous studies were able to report positive outcomes within a similar timespan. Third, instead of using independent observations, self-reporting was used to assess teachers' quality and quantity of program implementation. Self-reports may have led to an over-estimation of the quantity and quality of implementation. Besides, our number of implementation measures was limited. Fourth, although we did not monitor this, the lack of effects could have been caused by the level of standard social and character development activities that were given in the control group, for example by other preventive programs. Nonetheless, the practical value of PATHS would have been reflected in additional effects on top of care as usual, and this is exactly what we studied.

4.2. Conclusions

It seems unfortunate that the combination of an efficacious prevention program and a health service system specifically designed to be a linking agent for the implementation of school-based health promotion interventions does not provide better results – especially as universal prevention programs can only have population impact on the prevention of problem behavior when they are implemented on a large scale.

One could question whether this implementation strategy was the best possible. Obviously, the implementation strategy could have been enhanced with greater levels of support for both health promotion professionals and teachers. However, our aim was not to study the effectiveness of an intervention using a theoretically optimal implementation strategy, but rather to study its effectiveness when employing an implementation strategy that has been used successfully before, but, even more importantly, is commonly used and achievable within the Dutch public health sector. Along with efficacy trials, this kind of study helps us to identify interventions and implementation strategies that could be helpful (or not) in preventing problem behavior.

In addition, we believe that the present findings are probably not just relevant for PATHS or limited to prevention in the realm of social behavior. Implementation seems likely to play a key role in establishing societal impact. Our study underlines the importance of studying the process of transferring outcomes of efficacy studies to the more naturalistic settings for program implementation, and of monitoring program application in different countries and settings. More, well-designed large-scale field trials are urgently needed to provide policymakers with realistic estimates of the investments required to obtain intervention effects that are replicable at population level.

References

- Achenbach, Thomas M. 1991. *Manual for the Teacher's Report Form and 1991 Profiles*. Burlington: University of Vermont, Department of Psychiatry.
- Bajgar, Jane, Joseph Ciarrochi, Richard Lane, and Frank P. Deane. 2005. Development of the Levels of Emotional Awareness Scale for Children (LEAS-C). *British Journal of Developmental Psychology* 23:569–86. doi:10.1348/026151005X35417
- Beelmann, Andreas, and Tobias Raabe. 2009. The Effects of Preventing Antisocial Behavior and Crime in Childhood and Adolescence: Results and Implications of Research Reviews and Meta-Analyses. *European Journal of Developmental Science* 3:260–81. doi:10.3233/DEV-2009-3305
- Bryant, Brenda K. 1982. An Index of Empathy for Children and Adolescents. *Child Development* 53 (2): 413–25. doi:10.2307/1128984
- Cillissen, Antonius H. N., and Lara Mayeux. 2004. Sociometric Status and Peer Group Behavior: Previous Findings and Current Directions. In *Children's Peer Relations: From Development to Intervention*, ed. Janis B. Kupersmidt and Kenneth A. Dodge, 3–20. Washington, DC: American Psychological Association.
- Coie, John D., Kenneth A. Dodge, and Heide Coppotelli. 1982. Dimensions and Types of Social Status: A Cross-Age Perspective. *Developmental Psychology* 18:557–70.
- CPPRG (Conduct Problems Prevention Research Group). 1999. Initial Impact of the Fast Track Prevention Trial for Conduct Problems: II. Classroom Effects. *Journal of Consulting and Clinical Psychology* 67:648–57. doi:10.1037/0022-006X.67.5.648
- Crick, Nicki R., and Jennifer K. Grotpeter. 1996. Children's Treatment by Peers: Victims of Relational and Overt Aggression. *Development and Psychopathology* 8:367–80. doi:10.1017/S0954579400007148
- De Wied, Minet, Cora Maas, Stephanie Van Goozen, Marjolijn Vermande, Rutger Engels, Wim Meeus, Walter Matthys, and Paul Goudena. 2007. Bryant's Empathy Index: A Closer Examination of its Internal Structure. *European Journal of Psychological Assessment* 23:99–104. doi:10.1027/1015-5759.23.2.99
- Dodge, Kenneth A. 2009. Community Intervention and Public Policy in the Prevention of Antisocial Behaviour. *Journal of Child Psychology and Psychiatry* 50:194–200. doi:10.1111/j.1469-7610.2008.01985.x
- Domitrovich, Celene E., Rebecca C. Cortes, and Mark T. Greenberg. 2001. Head Start Competence Scale Technical Report. Unpublished manuscript, Pennsylvania State University.
- Domitrovich, Celine E., Rebecca C. Cortes, and Mark T. Greenberg. 2007. Improving Young Children's Social and Emotional Competence: A Randomized Trial of the Preschool "PATHS" Curriculum. *Journal of Primary Prevention* 28 (2): 67–91. doi:10.1007/s10935-007-0081-0
- Domitrovich, Celine E., Mark T. Greenberg, Rebecca C. Cortes, and Carol Kusché. 1999. *Manual for the Preschool PATHS Curriculum*. South Deerfield, MA: Channing-Bete.
- Dunn, Lloyd M., and Leota M. Dunn. 1997. *Peabody Picture Vocabulary Test – Revised*. Circle Pines, MN: American Guidance Service.
- Durlak, Joseph A., Roger P. Weissberg, Allison B. Dymnicki, Rebecca D. Taylor, and Kriston B. Schellinger. 2011. The Impact of Enhancing Students' Social and Emotional Learning: A Meta-Analysis of School-Based Universal Interventions. *Child Development* 82:405–32. doi:10.1111/j.1467-8624.2010.01564.x
- Eisner, Manuel. 2009. No Effects in Independent Prevention Trials: Can We Reject The Cynical View? *Journal of Experimental Criminology* 5:163–83. doi:10.1007/s11292-009-9071-y
- Elliott, Delbert S., and Sharon Mihalic. 2004. Issues in Disseminating and Replicating Effective Prevention Programs. *Prevention Science* 5:47–53. doi:10.1023/B:PREV.0000013981.28071.52
- ErasmusMC. 2000. *Problem Behaviour at School Interview*. Rotterdam: Department of Child and Adolescent Psychiatry.
- Flay, Brian R., Anthony Biglan, Robert F. Boruch, Felipe González Castro, Denise Gottfredson, Sheppard Kellam, Eve K. Mo cicki, Steven Schinke, Jeffrey C. Valentine, and Peter Ji. 2005. Standards of Evidence: Criteria for Efficacy, Effectiveness and Dissemination. *Prevention Science* 6:151–75. doi:10.1007/s11121-005-5553-y
- Gratz, Kim L., Lizabeth Roemer. 2004. Multidimensional Assessment of Emotion Regulation and Dysregulation: Development, Factor Structure and Initial Validation of the Difficulties in Emotion Regulation Scale. *Journal of Psychopathology and Behavioural Assessment* 26:41–54. doi:10.1023/B:JOBA.0000007455.08539.94
- Greenberg, Mark T., and Carol A. Kusché. 1993. *Promoting Social and Emotional Development in Deaf Children: The PATHS Project*. Seattle, WA: University of Washington Press.
- Greenberg, Mark T., Carol A. Kusché, Elizabeth T. Cook, and Julie P. Quamm. 1995. Promoting Emotional Competence in School-Aged Children: The Effects of the PATHS Curriculum. *Development and Psychopathology* 7:117–136. doi:10.1017/S0954579400006374
- Harter, Susan, and M. Nowakowski. 1987. *Manual for the Dimensions of Depression Profile for Children and Adolescents*. Denver, CO: University of Denver, Department of Psychology.
- Hox, Joop J. 2010. *Multilevel Analysis. Techniques and Applications*, 2d ed. New York: Routledge.
- Kam, Chi-Ming, Mark T. Greenberg, and Carla T. Walls. 2003. Examining the Role of Implementation Quality in School-Based Prevention using the PATHS Curriculum. *Prevention Science* 4:55–63. doi:10.1023/A:1021786811186
- Kusché, Carol A., and Mark T. Greenberg. 1994. *The PATHS Curriculum*. South Deerfield, MA: Channing-Bete.
- Louwe, Jos J., Cornelis W. van Overveld, Welmoet Merk, Bram Orobio de Castro, and Willem Koops. 2007a. De invloed van het Programma Alternatieve Denkstrategieën op reactieve en proactieve agressie bij jongens in het primair onderwijs: Effecten na één jaar. *Pedagogische Studiën* 84:277–92.
- Louwe, Jos J., Cornelis W. van Overveld, Welmoet Merk, Bram Orobio de Castro, and Willem Koops. 2007b. Het PAD van de school en de leraar. In *Jaarboek Lectoraten Faculteit Educatie 2007*. Utrecht: Hogeschool Utrecht.
- Malti, Tina, Denis Ribeaud, and Manuel Eisner. 2011. The Effectiveness of Two Universal Preventive Interventions in Reducing Children's Externalizing Behaviour: A Cluster Randomized Controlled Trial. *Journal of Clinical Child and Adolescent Psychology* 40 (5): 677–92. doi:10.1080/15374416.2011.597084
- Merrill, Kenneth W. 1996. Socio-Emotional Assessment in Early Childhood: The Preschool and Kindergarten Behaviour Scales. *Journal of Early Intervention* 20:132–45. doi:10.1177/105381519602000205
- Riggs, Nathaniel R., Mark T. Greenberg, Carol A. Kusché, and Mary Ann Pentz. 2006. The Medial Role of Neurocognition in the Behavioural Outcomes of a Social-Emotional Prevention Program in Elementary School Students: Effects of the PATHS Curriculum. *Prevention Science* 7:91–102. doi:10.1007/s11121-005-0022-1
- Rubin, Kenneth H., Julie C. Wojslawowicz, Linda Rose-Krasnor, Cathryn Booth-LaForce, and Kim B Burgess. 2006. The Best Friendships of Shy/Withdrawn Children: Prevalence, Stability, and Relationship Quality. *Journal of Abnormal Child Psychology* 34 (2): 143–57. doi:10.1007/s10802-005-9017-4
- Salmivalli, Christina, Ari Kaukiainen, and Marinus Voeten. 2005. Anti-Bullying Intervention: Implementation and Outcome. *British Journal of Educational Psychology* 75:465–87. doi:10.1348/000709905X26011
- SCDR (Social and Character Development Research Consortium). 2010. *Efficacy of School Wide Programs to Promote Social and Character Development and Reduce Problem Behaviour in Elementary School Children (NCER 2011–2001)*. Washington, DC: National Center for Education Research, Institute of Education Sciences, U.S. Department of Education.
- StataCorp. 2009. *Stata Statistical Software: Release 11*. College Station, TX: Stata-Corp LP.
- SvO (Seminarium voor Orthopedagogiek). 2005. *PAD, Programma Alternatieve Denkstrategieën, Lessen voor sociale en emotionele vaardigheden in de basisschool*. Utrecht: Trimbos Institute.
- Twisk, Jos W. R. 2006. *Applied Multilevel Analysis*. Cambridge: Cambridge University Press.

- Twisk, Jos W. R., and Wieke De Vente. 2008. The Analysis of Randomised Controlled Trial Data with More than One Follow-Up Measurement. A Comparison Between Different Approaches. *European Journal of Epidemiology* 23:655–60. doi:10.1007/s10654-008-9279-6
- Weisz, John R., Geri R. Donenberg, Susan S. Han, and Danika Kauneckis. 1995. Child and Adolescent Psychotherapy Outcomes in Experiments versus Clinics: Why the Disparity? *Journal of Abnormal Child Psychology* 23:83–106. doi:10.1007/BF01447046
- Wiefferink, Karin. H., Jos Poelman, M. Linthorst, Ine. Vanwesenbeeck, J. C. M. van Wijngaarden, and Theo G. W. Paulussen. 2005. Outcomes of a Systematically Designed Strategy for the Implementation of Sex Education in Dutch Secondary Schools. *Health Education Research* 20 (3): 323–33. doi:10.1093/her/cyg120
- Wilson, Sandra J., and Mark W. Lipsey. 2007. School-Based Interventions for Aggressive and Disruptive Behavior. Update of a Meta-Analysis. *American Journal of Preventive Medicine* 33 (2S): S130–43. doi:10.1016/j.amepre.2007.04.011
- Witvliet, Miranda, Pol A. C. van Lier, Pim Cuijpers, and Hans M. Koot. 2010. Change and Stability in Childhood Clique Membership, Isolation from Cliques, and Associated Child Characteristics. *Journal of Clinical Child and Adolescent Psychology* 39:12–24. doi:10.1080/15374410903401161

Ferry X. Goossens

fgoossens@trimbos.nl

Evelien M. J. C. Gooren

emjc.gooren@psy.vu.nl

Bram Orobio de Castro

b.orobiodecastro@uu.nl

Kees W. van Overveld

kees.vanoverveld@hu.nl

Goof J. Buijs

gbuijs@nigz.nl

Karin Monshouwer

kmonshouwer@trimbos.nl

Simone A. Onrust

sonrust@trimbos.nl

Theo G. W. M. Paulussen

theo.paulussen@tno.nl