


Classrooms in Peace Within Violent Contexts: Field Evaluation of *Aulas en Paz* in Colombia

Enrique Chau¹  · Madeleine Barrera¹ · Andrés Molano¹ · Ana María Velásquez¹ · Melisa Castellanos^{1,2} · Maria Paula Chaparro^{1,3} · Andrea Bustamante^{1,4}

© Society for Prevention Research 2017

Abstract Classrooms in Peace (*Aulas en Paz*) is an elementary school-based multicomponent program for prevention of aggression and promotion of peaceful relationships. Inspired by international programs and socio-emotional research, it includes (1) a classroom universal curriculum, (2) parent workshops and home visits to parents of the 10% most aggressive children, and (3) extracurricular peer groups of two aggressive and four prosocial children. Activities seek to promote socio-emotional competencies such as empathy, anger management, creative generation of alternatives, and assertiveness. A 2-year quasi-experimental evaluation was conducted with 1154 students from 55 classrooms of seven public schools located in neighborhoods with the presence of youth gangs, drug cartels, and high levels of community violence in two Colombian cities. Despite several implementation (e.g., about half of the activities were not implemented) and evaluation (e.g., randomization problems, large number of missing data, and changes between treatment and control groups) challenges, positive results were found in prosocial behavior and in reduction of aggressive behavior, according to teacher reports, and in assertiveness and reduction of verbal victimization, according to student reports. Furthermore, implementation cost (25 US

dollars per student per year) was very low compared to other programs in developed countries. This study shows that the Classrooms in Peace program has an important potential to generate positive results and highlights the challenges of implementing and evaluating prevention programs in highly violent environments.

Keywords Aggression · Socio-emotional competencies · Citizenship competencies · Bullying · Colombia

Many school-based programs seeking to promote peaceful relationships and to prevent aggression and violence have been developed and implemented around the world. Rigorous evaluations of some of these programs have found positive results (e.g., Beets et al. 2009; Durlak et al. 2011; Justicia-Arráez et al. 2015; Kärnä et al. 2011), but other evaluations have not (e.g., Farrington and Ttofi 2009; Social and Character Development Research Consortium 2010). It is still not clear which school-based programs, and especially in which contexts, are able to reduce aggression and violence. On the one hand, promoting peaceful relationships might be harder in contexts where community and domestic violence is common since exposure to violence among children is related to higher levels of aggression (e.g., Chau et al. 2009; Dodge et al. 1990; Gorman-Smith et al. 2004), emotion dysregulation (e.g., Schwartz and Proctor 2000), aggressive fantasies (Musher-Eizenman et al. 2004), and normative beliefs supporting aggression (e.g., Bradshaw et al. 2009; Chau et al. 2012; Guerra et al. 2003; Orue et al. 2011). On the other hand, some programs have shown that it is possible to reach positive results despite the difficult conditions of the context. For instance, evaluations in high-risk schools of programs such as The Incredible Years school program and Promoting Alternative THinking Strategies (PATHS) demonstrated

Electronic supplementary material The online version of this article (doi:10.1007/s11121-017-0754-8) contains supplementary material, which is available to authorized users.

✉ Enrique Chau
echaux@uniandes.edu.co

¹ Universidad de los Andes, Bogotá, Colombia

² Concordia University, Montreal, Canada

³ University of Toronto, Toronto, Canada

⁴ University of Missouri, St. Louis, MO, USA

positive results in students' behaviors and social competence (Bierman et al. 2010; Webster-Stratton et al. 2008). However, the effects are not always as large as they are expected. The current study evaluates the effect of Classrooms in Peace (*Aulas en Paz* in Spanish), a multicomponent program that seeks to prevent aggression and promote peaceful relationships in elementary schools located in urban contexts with high levels of community violence in Colombia.

Evaluations of other programs (e.g., Fast Track) have demonstrated the positive effects of including components focalized on children in higher risk in addition to universal intervention (Bierman et al. 2004; Conduct Problems Prevention Research Group 1999). Classrooms in Peace combines universal components, which aim to reach all students in the participating classrooms, with targeted (indicated) components, which specifically try to promote changes among those students with a higher risk of aggressive and violent behavior. The program consists of three components, all implemented from grade 2 to grade 5: (1) a classroom universal curriculum which includes modules based on children's literature, (2) workshops and home visits to parents, and (3) extracurricular activities with small heterogeneous peer groups. Although some of these components were inspired by existing programs (i.e., structure, types of activities), Classrooms in Peace specifically developed its original content for the sessions. In particular, home visits and extracurricular activities were inspired by the Montreal Prevention Program (Tremblay et al. 1995, 1996) and the children's literature modules of the classroom universal curriculum were inspired by the program Voices Reading (Walker et al. 2008). Classrooms in Peace seeks to promote the development of eight socio-emotional competencies (called citizenship competencies in Colombia, to be consistent with national educational policies; Chaux 2009; Patti and Cepeda 2007) which are key mediators in the program's theory of change (Chaux 2009, 2012): (1) empathy (Eisenberg 2000; Hoffman 2000), (2) anger management (Hanish et al. 2004), (3) assertiveness (Lange and Jakubowski 1980), (4) active listening (Gordon 1970), (5) perspective-taking (Selman 1980), (6) creative generation of options (de Bono 1970, 1985), (7) consideration of consequences (Crick and Dodge 1994; Slaby and Guerra 1988), and (8) critical thinking (Ennis 1987).

Previous evaluations of Classrooms in Peace have shown positive results, particularly in reducing aggression (e.g., Castellanos and Chaux 2010; Chaux 2012; Ramos et al. 2007). However, these evaluations had important limitations in terms of sample size (Ramos et al. 2007), missing data, and low levels of implementation (Chaux 2012), and they only lasted for one academic year (9 months). The current evaluation intended to address previous limitations by including a larger sample, using data imputation tools, ensuring closer and more frequent support to the implementation, and testing the effects over 2 years of implementation. In particular, the main

objective was to identify whether, how much, and at which costs children participating in Classrooms in Peace improve their socio-emotional competencies, such as empathy and assertiveness, and their prosocial behavior and reduce their levels of aggression, in comparison to non-participating children. More generally, it intended to address the dearth of rigorous evaluations conducted outside developed countries, especially in challenging contexts with high levels of exposure to violence.

Methods

Participants

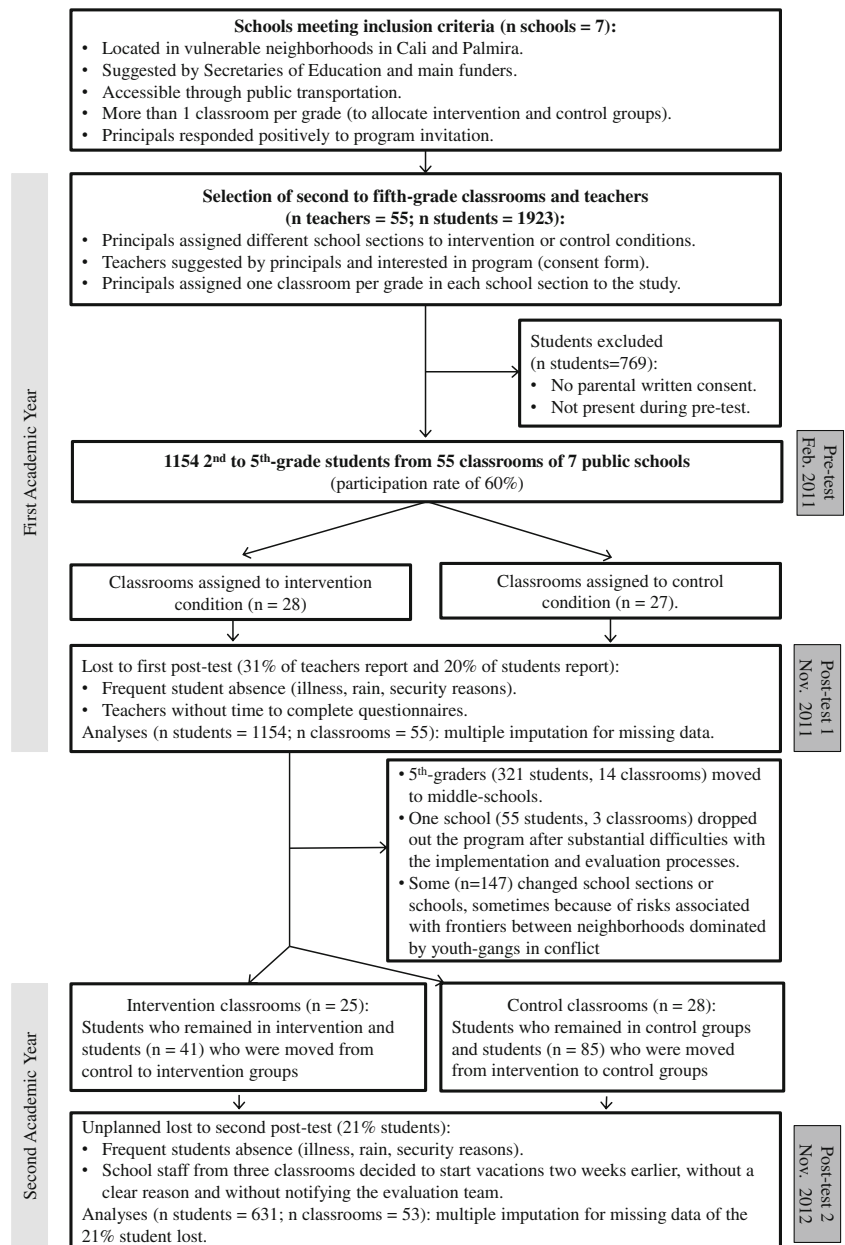
A total of 1154 second- to fifth-grade students from 55 classrooms of seven public schools participated in the pretest. At baseline, the average age was 8.7, and 48.9% were females. About 86% of second graders were between 7 and 8 years of age, 81% of third graders were between 8 and 10 years of age, and 77% of fourth graders were between 9 and 10 years of age.

Schools were located in two cities in the southwest of Colombia: Cali and Palmira. Cali is Colombia's third largest city with more than two million people and has a homicide rate of 64 per 100,000 habitants, making it the tenth most violent city in the world according to *Consejo Ciudadano para la Seguridad Pública y la Justicia Penal* (2016). The nearby city Palmira has about 300,000 habitants and is rated the eighth most violent city in the world with a homicide rate of 71 per 100,000 habitants (Consejo Ciudadano para la Seguridad Pública y la Justicia Penal 2016). Much of this violence is related to organized crime, fights between drug cartels, and youth gangs. Participating schools are located in some of the poorest neighborhoods with some of the highest levels of community violence in both cities. These seven schools were suggested by the offices of the Secretaries of Education and the main funders of the study. Figure 1 presents the CONSORT diagram which summarizes the evaluation procedure and the flow of participants in the different stages of the study.

Evaluation Procedure

Public schools in Colombian cities usually consist of two to four fairly independent sections (*Sedes* in Spanish), typically located in different buildings at walking distance from each other, and each with its own teachers and students. Principals selected, among the school sections, the one where the program would be implemented. Then, one classroom per grade in that section and one classroom per grade in another section were selected to be part of the evaluation. Selections were, in most cases, not completely random because principals had selected the intervening or comparison sections or classrooms

Fig. 1 CONSORT diagram



before the evaluation team was able to do it. It was not possible to identify the criteria with which selections were conducted. Prior to the second year, teachers were asked to sign a form if they wanted to continue with the program and other teachers were invited to participate.

Printed versions of the measures were administered during school time. Administration was supervised by research assistants who provided the instructions and assisted students who had questions. Teachers received their questionnaires on paper and returned them some days later to the research assistants. Pretests occurred at the beginning of the academic year (in February 2011) prior to implementation. The first posttest took place at the end of that academic year (November

2011), while the second posttest took place at the end of the following academic year (November 2012) (see Fig. 1).

Student Measures

Victimization (adapted from Velásquez and Chaux 2005; Cronbach’s $\hat{\alpha} = 0.83$) was measured with nine items inquiring for the frequency of physical (two items; e.g., “During the last month, has someone from your group hit you or hurt you on purpose?”), verbal (three items; e.g., “During the last week, has a classmate insulted you?”), and relational (four items; e.g., “During the last week, has a classmate invented bad things about you that were not true?”) aggression.

Aggression (adapted from Velásquez and Chaux 2005; Cronbach's $\hat{\alpha} = 0.88$) was measured with nine items about the frequency of physical (two items; e.g., "During the last month, have you on purpose hit or hurt someone from your school?"), verbal (three items; e.g., "During the last week, have you insulted someone from your school?"), and relational (four items; e.g., "During the last week, have you created false rumors about someone from your school?") aggression. Response options were "not once" (1), "once or a couple of times" (2), and "many times" (3).

Empathy (adapted from Schulz et al. 2011; Cronbach's $\hat{\alpha} = 0.87$) was measured with 14 items about how they feel when they see a child in a negative situation (e.g., "When I see a child sad because the others do not let him play, I feel..."; "when I see that others hit a girl that I don't like, I feel..."). Response options were "happy" (1), "nothing" (2), and "sad" (3) for fourth and fifth graders, and smiling (1), neutral (2), and (3) sad faces for second and third graders.

Assertiveness (adapted from Velásquez 2005; Cronbach's $\hat{\alpha} = 0.79$) was measured with four situations about being victim of aggression (e.g., "When a boy calls me by an ugly nickname to make me feel bad –like stupid, dumb, etc-...") and six options of responding: two aggressive responses ((1), e.g., "I call him an equally ugly nickname"), a passive response ((2), e.g., "I let him say what he wants and don't do anything"), an evasive response ((3), "I don't say anything and I leave"), asking support from an adult ((4), "I ask my teacher for help"), and an assertive response ((5), e.g., "Without shouting, I tell him that he made me feel bad and I ask him not to call me like that"). The rating of the response options reflects a progression from destructive responses that could escalate the problem to constructive responses that could deescalate it. For second and third graders, a cartoon illustration was included for each of the response options.

Demographic information, such as age, grade, and sex, and *classroom characteristics*, such as classroom size, were also collected.

Preliminary versions of all the instruments were piloted with students from similar ages and socio-economic backgrounds, and adaptations in language and complexity were made before the application of the pretest.

Teacher Measures

Teachers filled an instrument reporting seven items about each of their students (Chaparro and Chaux 2009). Four of these items were about student's physical, verbal, or relational aggression (e.g., "excludes other children from games or activities"; $\hat{\alpha} = 0.91$). Three items were about *prosocial behavior* (e.g., "If there is a physical fight, he/she tries to stop it"; $\hat{\alpha} = 0.75$). Response options for these two measures were "never" (1), "almost never" (2), "almost always" (3), and "always" (4).

Although family workshops and home visits intend to generate changes in parental practices and, in general, in the home climates, these changes were too difficult to measure, given the resources available.

The Program

Classroom Component The class curriculum consists of 40 sessions per grade, 24 of which are usually implemented in the ethics class and 16 are implemented in the language (Spanish) class. These sessions address topics of conflicts, bullying, and peer aggression. The curriculum has a particular focus on the role of bystanders by promoting assertive interventions to defend victims in bullying situations, and peer mediation during conflicts. This emphasis on the role of bystanders is crucial since it has been known since classical social psychology studies that human behavior depends greatly on what others around do (e.g., Milgram 1974; Zimbardo 2007).

Inspired by the Boston-based program Voices Reading (Walker et al. 2008), sessions in the language class are based on children's literature and theater (Classrooms in Peace developed its own reading list), promoting socio-emotional competencies (i.e., empathy, critical thinking) and language competencies such as reading, writing, and storytelling simultaneously. This curriculum is taught during school hours by their regular teachers.

Parenting Component It includes four workshops per grade offered to all parents of participating classes, as well as four home visits per year offered to the parents of children identified through peer and teacher reports as being in the 10% most aggressive children in their classes. These workshops and home visits seek to help parents promote family environments that are consistent with what children are learning at school. They seek to provide caregivers with parenting strategies, such as positive discipline (Nelsen 2006) and positive reinforcement schedules (Patterson 1976), and with similar socio-emotional competencies as those being learned by their children (e.g., anger management and conflict resolution skills). Home visits include similar activities as those in parent workshops but are focused on the particular needs of the specific families being visited. For instance, when children are learning conflict resolution in their classrooms, home visits seek to focus on each family's most common parent-child conflicts. Furthermore, home visits reach families which seem to be in great need of these strategies but which rarely participate in the workshops offered in the schools.

Heterogeneous Groups The third component of the program consists of 16 sessions of extracurricular activities conducted in small heterogeneous groups of six children: two identified by peer and teacher reports as being among the 10% most aggressive in their classes (the same who receive home visits)

and four identified as being among the 20% most prosocial students in their classes. This heterogeneous arrangement aims to promote peer-positive effects led by the most prosocial children. This configuration is designed to avoid deviancy training, which has been found to occur in interventions working exclusively with at-risk children or adolescents (Arnold and Hughes 1999; Dishion et al. 1999). In this component, students consolidate the learning and practice of socio-emotional competencies through different activities (e.g., role-plays, brainstorm, discussions).

The two targeted components of the program (i.e., parenting component and heterogeneous groups) were inspired by the Montreal Prevention Experiment, an intervention focused on high-risk children which demonstrated very positive long-term effects (Tremblay et al. 1995, 1996). The activities were original designs from Classrooms in Peace and were all piloted prior to intervention to ensure language appropriateness and cultural relevance. A previous evaluation of Classrooms in Peace showed larger changes among those groups where more extracurricular activities were implemented (Castellanos and Chaux 2010).

Program Implementation

Implementation of the program was led by the non-governmental organization *Convivencia Productiva*. The classroom component was implemented by the students' regular teachers. All those who implemented the program received, depending on the availability of time for training in each school, a short 2- to 4-h training or a longer 8- to 14-h training conducted by the implementing organization. Staff from this organization also provided on-site coaching through meetings at least every month with each teacher implementing the program. Meetings were more frequent with teachers who were identified as requiring additional support for having difficulties with classroom management or adapting to the pedagogical strategies of the program. Teachers were also observed (by staff from the implementing organization) implementing at least one session of the curriculum and provided feedback. The family and extracurricular components were implemented by undergraduate students from local pedagogical colleges doing their practicums, who also received training (8 to 14 h) and support from the non-governmental organization.

In total, 46% of the classroom sessions (on average, 17.3 sessions per group per year), 59% of the heterogeneous group sessions (9.4 sessions per group per year), and about 50% of the parent workshops and home visits (2 per group or family per year) were implemented.

The cost of the implementation of the program was 25 US dollars per student per year (57% in salaries of teacher trainers and coaches, 22% in materials, 8% in transport, and 12% in administration; teachers received no extra payments for

participating; *Convivencia Productiva*, personal communication, February 10, 2016).

Data Analysis

Initially, to compensate for a large amount of attrition which occurred at different moments of the process (see Fig. 1), multiple imputations of missing data were conducted (see online Appendix 1). Then, one-way ANOVAs were conducted to estimate the mean differences between treatment and control groups at pretest, first posttest, and second posttests (see online Appendix 2). To account for the multilevel structure of the data (repeated measures within students within classrooms) and to control for student and contextual characteristics, a set of *multilevel models for change* (Singer and Willett 2003) were estimated in which intraindividual change over time on selected outcomes is described as a function of individual- and classroom-level covariates and an indicator of treatment status. In these estimates, we focus on the intent-to-treat (ITT) effect, defined as the classroom-level condition (program or control) assigned at the beginning of the first year. To estimate these models, we fitted to our three-level data models to represent change over time, individual student differences, and differences between classrooms, according to the following collapsed equation:

$$Y_{ij} = \beta_0 + \beta_1(\text{Time}_i) + \beta_2(\text{Treatment}_j) + \beta_3(\text{Time}_i \times \text{Treatment}_j) + \alpha(\text{Demographic}_i) + \gamma(\text{Classroom characteristic}_j) + \varepsilon_{ij} \quad (1)$$

In this equation, Y_{ij} stands as the estimated value for each one of our outcome variables at time t , for child i , on classroom j . Of central note, the estimated value of β_3 represents the difference between the slopes of the treatment and control groups from time 0 to time 2 and allows us to test the intent-to-treat effect of the intervention on outcomes of children initially assigned to the treatment condition. These estimates could be regarded as obtained from a difference in difference model (Murnane and Willett 2011) where β_2 accounts for the mean differences between treatment and control groups at the beginning of the evaluation. Additionally, as noted in Eq. 1, to estimate our models, we include a vector of baseline covariates that describes individual (e.g., child gender and age) and classroom demographic (e.g., grade and city) characteristics. Finally, given the structure of our data (time nested in children and children nested in classrooms) along with Eq. 1, we estimated a set of variance components that account for the nesting and process of error autocorrelation. For ease of interpretation, we focus our presentation of results and discussion on the estimates obtained for β_3 , our intent-to-treat estimates.

Results

Differences Between Treatment and Control Before and After the Intervention

At the pretest, one-way ANOVAs showed that teacher-reported aggression, student-reported verbal victimization, verbal aggression, and relational aggression were significantly higher in the treatment group compared to the control group, while teacher-reported prosocial behavior and student-reported assertiveness were significantly higher in the control group in comparison to the treatment group (see online Appendix 2). In contrast, indicating positive effects of the program, there were no significant differences in any of the outcomes at the second posttest, except that teacher-reported aggression was significantly higher in the control group than in the treatment group (see online Appendix 2).

Growth Over Time Controlled for Covariates

Following our formulation of Eq. 1, we estimated our difference in difference model by employing a multilevel model for change (Singer and Willett 2003). Intraclass correlation analysis from each outcome's unconditional growth model is available online (see online Appendix 3). Table 1 presents the results of our estimations. According to our estimated results, treatment and control groups did differ statistically at baseline on some teacher-reported and self-reported variables: program classrooms were lower on prosocial behavior as reported by teachers ($\hat{\beta}_2 = -0.18$) and higher on self-reported verbal victimization ($\hat{\beta}_2 = 0.12$) and self-reported verbal and relational aggression ($\hat{\beta}_2 = 0.08$ and $\hat{\beta}_2 = 0.07$, respectively).

Of the control variables, only gender had consistent effects, except for prosocial behavior reported by teachers. At baseline, girls were lower than boys on aggression reported by teachers ($\hat{\alpha}_{\text{female}} = -0.33$) and self-reported verbal and physical victimization ($\hat{\alpha}_{\text{female}} = -0.06$ and $\hat{\alpha}_{\text{female}} = -0.10$, respectively), as well as on self-reported aggression indicators (verbal, $\hat{\alpha}_{\text{female}} = -0.14$; relational, $\hat{\alpha}_{\text{female}} = -0.04$; and physical, $\hat{\alpha}_{\text{female}} = -0.16$). However, girls were higher than boys on relational victimization ($\hat{\alpha}_{\text{female}} = 0.06$). Concerning self-reported socio-emotional competencies, girls were more assertive ($\hat{\alpha}_{\text{female}} = 0.27$) and empathic ($\hat{\alpha}_{\text{female}} = 0.11$).

Program effects were examined for the 2 years of implementation. Aggressive behavior decreased significantly more in treatment classrooms in comparison to changes in control classrooms ($\hat{\beta}_3 = -0.13$), and prosocial behavior increased significantly more in treatment than in control classrooms ($\hat{\beta}_3 = 0.13$) according to teacher reports. According to student reports, verbal victimization decreased significantly more in treatment classrooms in comparison with changes in control groups ($\hat{\beta}_3 = -0.05$). All other effects did not reach statistical significance.

Effect sizes in the metric of Cohen's d were from 0 to 0.27 in absolute values (see Table 1), all in the expected direction.

Also, to estimate differential effects in the function of time, we conducted the original analysis including two time dummy variables instead of one continuous time variable. Table 2 presents the results for these estimations.

According to these estimated results, program effects on aggressive and prosocial behavior, according to teacher report, were significant on both years. Regarding student report, significant effects for verbal victimization and on assertiveness were found for the second year of implementation. Effect sizes in the metric of Cohen's d were from 0 to 0.44 in absolute values (see Table 2), all in the expected direction in year 2.

Finally, to estimate the differential effects of the program based on students' initial level of aggression, multilevel models for change were conducted with an interaction between treatment and a dichotomous variable coding whether or not initial levels of aggression were above 1 standard deviation from the mean. No significant results were found. For sample size reasons, interactions based on smaller groups were not tested.

Discussion

This study examined the effects of the Classrooms in Peace (*Aulas en Paz*) multicomponent program on behavioral outcomes (i.e., aggression, prosociality, and victimization) as well as on some socio-emotional competencies (i.e., assertiveness and empathy) of second to fifth graders in Colombian schools, located in a high-risk community context. Overall, a positive impact of the program was found on teacher-reported aggression and prosocial behavior and on student-reported verbal victimization and assertiveness. This adds to the current evidence that it is possible to have positive results with school-based programs based on socio-emotional development and implemented in contexts with high levels of community violence as some other evaluations have shown (e.g., Bierman et al. 2010; Webster-Stratton et al. 2008).

The study also highlights several challenges associated with implementing and evaluating school-based programs in vulnerable and under-resourced contexts (e.g., about 50% of the program activities were not implemented). Also, recruitment of teachers was mostly based on their expressed desire to participate. However, teacher's initial motivation towards the program might not be enough to guarantee full implementation since the risk of frustration is large, given that most participating teachers work in very difficult contexts, surrounded by high levels of exposure to violence which legitimize and reinforce their students' aggression. In contrast to other international contexts where similar programs are implemented, they have to face a larger distance between what is expected of their students and what their students are exposed to outside

Table 1 Multilevel modeling results for teacher- and self-reported outcomes

	Teacher report			Student self-report								
	Aggression	Prosocial behavior		Victimization			Aggression			Socio-emotional competencies		
			Verbal	Relational	Verbal	Relational	Physical	Verbal	Relational	Physical	Assertiveness	Empathy
Baseline												
Intercept	0.44***	3.32***		1.33***	1.23***	1.53***	0.85***	1.02***	1.02***	4.25***	3.04***	
Student level												
Girl	-0.33***	0.01		-0.06**	0.06**	-0.10***	-0.14***	-0.04**	-0.16***	0.27***	0.11***	
Age	0.13***	-0.04***		0.01	0.02*	0.00	0.05***	0.03***	0.03***	-0.05***	-0.03***	
Classroom level												
Treatment	0.07	-0.18*		0.12**	0.03	0.05	0.08**	0.07*	0.04	-0.10#	-0.02	
Grade 3	-0.23**	0.09		0.02	-0.01	-0.02	-0.07#	-0.01	-0.05	0.01	-0.04	
Grade 4	-0.25***	0.03		0.02	-0.02	-0.04	-0.06	-0.08*	-0.08*	-0.19**	-0.03	
Grade 5	-0.48***	0.07		0.04	-0.05	-0.05	-0.08#	-0.08*	-0.09*	-0.22**	-0.05	
Change over time												
Time	0.05**	0.04#		0.01	0.00	-0.06**	0.03*	0.00	0.00	-0.05#	-0.03**	
Treatment × time	-0.13***	0.13***		-0.05*	-0.01	-0.01	-0.03	-0.02	-0.02	0.07#	0.00	
Random effects												
Level 1												
Intercept	0.28***			0.11***	0.07***	0.13	0.06***	0.06***	0.08***	0.19***	0.06***	
Time	0.04***	0.03***		0.02***	0.01***	0.02	0.01***	0.01***	0.01***	0.03***	0.01***	
Covariate (intercept, time)	-0.09***			-0.03**	-0.02***	-0.05	-0.02***	-0.03**	-0.03***	-0.07*	-0.02***	
Level 2												
Intercept	0.02***	0.06***		0.01***	0.01***	0.01	0.005***	0.003***	0.002***	0.008***	0.003***	
Residual	0.17**	0.30***		0.194**	0.16***	0.22	0.14***	0.10***	0.13***	0.45***	0.08***	
Model fit (-2ll)	4557.26	5155.78		4383.53	3621.78	4627.20	3191.93	2353.65	3023.90	6566.25	1693.20	
Effect size (Cohen's <i>d</i>)	-0.27	0.26		-0.11	-0.03	-0.02	-0.08	-0.06	-0.06	0.11	0.00	

Cohen's *d* were calculated as the adjusted group mean difference (from β_3 ; intervention × time) divided by the unadjusted pooled within-group standard deviation, as proposed by Kärrä et al. (2011). All models were controlled by city (Cali or Palmira)

Most relevant results are highlighted in bold

$p < 0.10$; * $p < 0.05$; ** $p < 0.01$; *** $p < 0.001$

Table 2 Multilevel modeling results for teacher- and self-reported outcomes for both years, separately

	Teacher report		Student self-report							
	Aggression	Prosocial behavior	Victimization			Aggression			Socio-emotional competencies	
			Verbal	Relational	Physical	Verbal	Relational	Physical	Assertiveness	Empathy
Baseline										
Intercept	0.68***	3.10***	1.39***	1.28***	1.61***	0.89***	1.04***	1.06***	4.24***	2.99***
Student level										
Girl	-0.36***	0.01	-0.06*	0.06**	-0.10***	-0.13***	-0.04**	-0.15***	0.27***	0.10***
Age	0.14***	-0.04***	0.01	0.02*	0.00	0.05***	0.03***	0.03***	-0.06***	-0.03***
Classroom level										
Treatment	0.10	-0.21**	0.12**	0.04	0.03	0.08*	0.06*	0.02	-0.10	-0.02
Grade 3	-0.23**	0.09	0.02	-0.01	-0.02	-0.07	-0.01	-0.04	0.04	-0.04
Grade 4	-0.26***	0.03	0.02	-0.01	-0.03	-0.06	-0.08*	-0.07	-0.17**	-0.01
Grade 5	-0.48***	0.07	0.04	-0.05	-0.05	-0.09	-0.09*	-0.09*	-0.16*	-0.04
Change over time										
Year 1	0.06*	0.00	0.06*	0.03	-0.07*	0.05	0.01	0.00	-0.05	-0.04*
Year 2	0.09*	0.08*	0.01	0.00	-0.11**	0.06	0.01	0.00	-0.09	-0.06**
Treatment × year 1	-0.25***	0.23***	-0.07	-0.04	0.03	-0.01	-0.01	0.02	0.06	-0.01
Treatment × year 2	-0.19***	0.22***	-0.11*	-0.03	-0.05	-0.07	-0.07	-0.06	0.14*	0.00
Random effects										
Intercept	0.12***	0.03***	0.06***	0.04***	0.05***	0.03***	0.02***	0.03***	0.09***	0.03***
Time	0.03***	0.07***	0.01***	0.01***	0.01***	0.01***	0.01***	0.01***	0.01***	0.01***
Residual	0.20***	0.30***	0.21***	0.17***	0.24***	0.15***	0.12***	0.14***	0.47***	0.08***
Effect size (Cohen's <i>d</i>), year 1	-0.38	0.34	-0.13	-0.09	0.06	-0.02	-0.03	0.05	0.07	-0.03
Effect size (Cohen's <i>d</i>), year 2	-0.40	0.44	-0.23	-0.08	-0.11	-0.19	-0.22	-0.18	0.22	0.00

Cohen's *d* were calculated as the adjusted group mean difference (from β_3 : intervention \times time) divided by the unadjusted pooled within-group standard deviation, as proposed by Kärnä et al. (2011). All models were controlled by city (Cali or Palmira)

Most relevant results are highlighted in bold

* $p < 0.05$; ** $p < 0.01$; *** $p < 0.001$

of the school. Also, they have to deal with large classroom sizes (56% of classrooms had between 30 and 39 students, while 26% had between 40 and 49 students, all with only one adult per classroom) and precarious room conditions (e.g., no fan or air conditioner, in extreme heat). This requires a greater effort, persistence, self-efficacy, stress coping strategies, and higher tolerance to frustration, in addition to having developed the same socio-emotional competencies that they have to teach. In Colombia, teachers receive little to no training at all in their undergraduate programs regarding socio-emotional development or classroom management skills (García et al. 2014). Within this context, teacher training for the implementation of programs such as Classrooms in Peace may require additional efforts to compensate for this lack of previous training.

There are also several challenges related to evaluation. Some of the current measures in this study might have ceiling

effects and social desirability problems or might not be sensitive enough to detect small changes in students. Second, despite the efforts to use rigorous experimental designs, it was not possible to conduct a rigorous randomization of experimental and control groups. At the pretest, we found higher levels of aggression and lower levels of prosocial and assertive behaviors among the students assigned to the treatment group as compared to those in the control groups. This was probably due to schools' eagerness to assign intervention to the most at-risk sections, classrooms, or students. Also, there were large levels of attrition, and some students were changed, during the evaluation, from treatment to control groups or from control to treatment groups. All of these methodological limitations speak to the challenges of aiming for a randomized control trial design in diverse cultural and vulnerable settings. Elaborated statistical analyses, such as some of those conducted here, may be needed to address these difficulties.

Furthermore, we are currently attempting to correct these sources of bias by implementing a nested propensity score-matching procedure (Zhai et al. 2012), while accounting for the time variant movement of children across treatment groups (Hirano et al. 2000).

Future evaluation studies could examine in more detail which components, in which contexts, and for which students may be more effective, even though testing the added benefits of each component requires more complex designs, with differentiated experimental conditions, and a much larger sample size than that of the current study. Also, a closer look at the conditions of implementation entails having measures of the teachers who lead the implementation, as well as improving the training they should get to develop their own pedagogical and socio-emotional competencies. Accounting for the variability of conditions also calls for a very strict follow-up of each student's and teacher's experimental condition, especially when both students and teachers may drop out of schools or change classrooms (Jones and Molano 2016). Lastly, it is possible that the effects of the program might be more evident by focusing the analyses in those students who, at baseline, lack these competencies or the desired behavioral outcomes. Again, a larger sample size would be needed to check this possibility.

At the methodological level, further studies could try to qualify their assessment and analytical approaches. We relied in this study on teachers' and students' reports. However, teachers were not blinded to the experimental condition of the students, and students might have been biased when reporting on their own behavior. These limitations could be overcome by using alternative measurement tools, such as those based on virtual reality (e.g., Han et al. 2009), mixed methods, peer nominations, experimental manipulations (e.g., Yeager et al. 2013), or observational data with planned missing data designs (Little et al. 2014).

The cost of implementation of Classrooms in Peace was 25 US dollars per student per year. This is very low compared to other multicomponent programs. Fast Track (Conduct Problems Prevention Research Group 2002), for instance, has been estimated in a cost of 5800 US dollars per child per year (Foster 2010), that is, 232 times more. One of the main contrasts between implementing school-based multicomponent programs in developing countries, compared to developed ones, might be their associated costs. This should be taken into account when conducting cost-benefit analyses.

Despite the discussed limitations, the evaluation of the Classrooms in Peace program shows that it is not only possible to implement a multicomponent program through alliances and support from local universities and non-governmental organizations but also possible to reach, at very low costs, positive results in aggression, prosocial behavior, and assertiveness. Furthermore, understanding challenges as those outlined by this study may represent an important contribution to our knowledge on prevention science, as they unveil the

complexity of intervention and evaluation of programs in the vulnerable contexts where they are needed the most.

Acknowledgements We are very grateful to the Manuelita S.A. and its affiliated Harold Eder Foundation for financing the current study. We also thank the non-governmental organization Convivencia Productiva for leading the implementation of the program and for providing support during the data collection process. A substantial part of the writing of the article was made possible by a Georg Forster Fellowship granted to the first author by the Alexander von Humboldt Foundation, Germany. The second author was financed by a full scholarship by Colciencias. We are very thankful to Diana Andrade, Martha Isabel Cerón, Ana Lucía Jaramillo, Manuela Jiménez, José Fernando Mejía, Juan Jacobo Ospina, Mariajosé Otálora, and Álvaro Valencia for their crucial participation in the design and implementation of the evaluation. We really appreciate the careful reading and insightful suggestions by Guest Editor Dr. Nancy Guerra, Editor-in-Chief Dr. Catherine Bradshaw, and the two anonymous reviewers. We would like to specially thank all the students, teachers, principals, school coordinators, school psychologists, college students, volunteers, and research assistants who participated in this study. We would like to dedicate this paper to the memory of Anna King, whose strong convictions about the need for socio-emotional development helped bring Aulas en Paz to thousands of children, their teachers, and their families.

Author's Contributions Chaux, Barrera, Castellanos, and Chaparro designed the evaluation. Analyses were conducted by Chaux, Barrera, Molano, and Velásquez. Drafts of the manuscript were written by Chaux, Barrera, Molano, and Velásquez. The other authors were part of the evaluation team and helped in the writing of the manuscript.

Compliance with Ethical Standards

Funding The study was funded by the private group Manuelita S.A. and its affiliated Harold Eder Foundation. A substantial part of the writing of the article was made possible by a Georg Forster Fellowship granted to the first author by the Alexander von Humboldt Foundation, Germany. The second author was financed by a full scholarship by Colciencias.

Conflict of Interest As potential conflict of interests, we declare that some of the authors of this manuscript (Chaux, Velásquez, Castellanos, Chaparro, and Bustamante) were part of the team that designed the Classrooms in Peace (*Aulas en Paz*) program and that the first author (Chaux) was the leader of that team.

Ethical Approval All procedures were in accordance with the ethical standards of the institutional and/or national research committee and with the 1964 Helsinki Declaration and its later amendments or comparable ethical standards.

The study was approved by the Ethics Committee of Universidad de los Andes, Bogotá, Colombia.

Informed Consent Written consents were obtained from parents of all students who participated in the study as well as from all teachers who participated.

References

- Arnold, M. E., & Hughes, J. N. (1999). First do no harm: Adverse effects of grouping deviant youth for skills training. *Journal of School Psychology, 37*, 99–115.

- Beets, M. W., Flay, B. R., Vuchinich, S., Snyder, F. J., Acock, A., Li, K., Burns, K., Washburn, I. J., & Durlak, J. (2009). Use of a social and character development program to prevent substance use, violent behaviors, and sexual activity among elementary-school students in Hawaii. *American Journal of Public Health, 99*, 1438–1445.
- Bierman, K. L., Coie, J. D., Dodge, K. A., Foster, E. M., Greenberg, M. T., Lochman, J. E.,... & Pinderhughes, E. E. (2004). The effects of the Fast Track program on serious problem outcomes at the end of elementary school. *Journal of Clinical Child and Adolescent Psychology, 33*(4), 650–661.
- Bierman, K. L., Coie, J. D., Dodge, K. A., Greenberg, M. T., Lochman, J. E., McMahon, R. J., & Pinderhughes, E. (2010). The effects of a multiyear universal social-emotional learning program: The role of student and school characteristics. *Journal of Consulting and Clinical Psychology, 78*, 156.
- Bradshaw, C. P., Rodgers, C. R. R., Ghandour, L. A., & Garbarino, J. (2009). Social-cognitive mediators of the association between community violence exposure and aggressive behavior. *School Psychology Quarterly, 24*, 199–210.
- Castellanos, M., & Chaux, E. (2010). *Informe de evaluación de impacto del Programa Aulas en Paz*. Bogotá: Universidad de los Andes. Ministerio de Educación Nacional. Unpublished manuscript.
- Chaparro, M. P., & Chaux, E. (2009). *Instrumento multiusos para conformación de grupos heterogéneos*. Bogotá: OIM. Universidad de los Andes. Unpublished document.
- Chaux, E. (2009). Citizenship competencies in the midst of a violent political conflict: The Colombian educational response. *Harvard Educational Review, 79*, 84–93.
- Chaux, E. (2012). *Educación, convivencia y agresión escolar*. Bogotá: Ediciones Uniandes. Taurus, Santillana.
- Chaux, E., Molano, A., & Podlesky, P. (2009). Socio-economic, socio-political and socio-emotional variables explaining school bullying: A country-wide multilevel analysis. *Aggressive Behavior, 35*, 520–529.
- Chaux, E., Arboleda, J., & Rincón, C. (2012). Community violence and reactive and proactive aggression: The mediating role of cognitive and emotional variables. *Revista Colombiana de Psicología, 21*, 233–251.
- Conduct Problems Prevention Research Group. (1999). Initial impact of the Fast Track prevention trial for conduct problems: I. The high-risk sample. *Journal of Consulting and Clinical Psychology, 67*(5), 631.
- Conduct Problems Prevention Research Group. (2002). Evaluation of the first 3 years of the Fast Track prevention trial with children at high risk for adolescent conduct problems. *Journal of Abnormal Child Psychology, 30*, 19–35.
- Consejo Ciudadano para la Seguridad Pública y la Justicia Penal (2016). *Caracas, Venezuela, the most violent city in the world*. Retrieved from: <http://www.seguridadjusticiaypaz.org.mx>.
- Crick, N. R., & Dodge, K. A. (1994). A review and reformulation of social information-processing mechanisms in children's social adjustment. *Psychological Bulletin, 115*, 74–101.
- de Bono, E. (1970). *Lateral thinking: A textbook of creativity*. London: Ward Lock Educational.
- de Bono, E. (1985). *Six thinking hats*. Boston: Little, Brown.
- Dishion, T. J., McCord, J., & Poulin, F. (1999). When interventions harm: Peer groups and problem behavior. *American Psychologist, 54*, 755–761.
- Dodge, K. A., Bates, J. E., & Pettit, G. S. (1990). Mechanisms in the cycle of violence. *Science, 250*, 1678–1683.
- 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.
- Eisenberg, N. (2000). Emotion, regulation, and moral development. *Annual Review of Psychology, 51*, 665–697.
- Ennis, R. (1987). A taxonomy of critical thinking dispositions and abilities. In J. B. Baron & R. J. Sternberg (Eds.), *Teaching thinking skills: Theory and practice* (pp. 9–26). New York: W. H. Freeman.
- Farrington, D. P. & Ttofi, M. M. (2009). School-based programs to reduce bullying and victimization. *Campbell Systematic Reviews, 6*. doi:10.4073/csr.2009.6.
- Foster, E. M. (2010). Costs and effectiveness of the Fast Track intervention for antisocial behavior. *The Journal of Mental Health Policy and Economics, 13*, 101–119.
- García, S., Maldonado, D., Perry, G., Rodríguez, C., & Saavedra, J. E. (2014). *Tras la excelencia docente. Cómo mejorar la calidad de la educación para todos los colombianos*. Bogotá: Fundación Compartir.
- Gordon, T. (1970). *Parent effectiveness training: The no-lose program for raising responsible children*. New York: P.H. Wyden.
- Gorman-Smith, D., Henry, D. B., & Tolan, P. H. (2004). Exposure to community violence and violence perpetration: The protective effects of family functioning. *Journal of Clinical Child and Adolescent Psychology, 33*, 439–449.
- Guerra, N., Huesmann, R., & Spindler, A. (2003). Community violence exposure, social cognition and aggression among urban elementary school children. *Child Development, 74*, 1561–1576.
- Han, K., Ku, J., Kim, K., Jang, H. J., Park, J., Kim, J. J., Kim, C. H., Choi, M. H., Kim, I. Y., & Kim, S. I. (2009). Virtual reality prototype for measurement of expression characteristics in emotional situations. *Computers in Biology and Medicine, 39*, 173–179. doi:10.1016/j.compbimed.2008.12.002.
- Hanish, L. D., Eisenberg, N., & Fabes, R. A. (2004). The expression and regulation of negative emotions: Risk factors for young children's peer victimization. *Development and Psychopathology, 16*, 335–353.
- Hirano, K., Imbens, G. W., Rubin, D. B., & Zhou, X. H. (2000). Assessing the effect of an influenza vaccine in an encouragement design with covariates. *Biostatistics, 1*, 69–88.
- Hoffman, M. L. (2000). *Empathy and moral development: Implications for caring and justice*. New York: Cambridge University Press.
- Jones, S. M., & Molano, A. (2016). Seasonal and compositional effects of classroom aggression: A test of developmental-contextual models. *Journal of Cognitive Education and Psychology, 15*, 225–247.
- Justicia-Arráez, A., Pichardo, C., & Justicia, F. (2015). Efecto del programa Aprender a Convivir en la competencia social y en los problemas de conducta del alumnado de 3 años. *Anales de Psicología, 31*, 825–836. doi:10.6018/analesps.31.3.185621.
- Kärnä, A., Voeten, M., Little, T. D., Poskiparta, E., Kaljonen, A., & Salmivalli, C. (2011). A large-scale evaluation of the KiVa anti-bullying program: Grades 4–6. *Child Development, 82*, 311–330. doi:10.1111/j.1467-8624.2010.01557.x.
- Lange, A., & Jakubowski, P. (1980). *Responsible assertive behavior: Cognitive, behavioural procedures for trainers*. Champaign: Research.
- Little, T. D., Jorgensen, T. D., Lang, K. M., & Moore, E. W. (2014). On the joys of missing data. *Journal of Pediatric Psychology, 39*, 151–162. doi:10.1093/jpepsy/jst048.
- Milgram, S. (1974). *Obedience to authority: An experimental view*. New York: Harper & Row.
- Murnane, R. J., & Willett, J. B. (2011). *Methods matter: Improving causal inference in educational and social science research*. New York: Oxford University Press.
- Musher-Eizenman, D. R., Boxer, P., Danner, S., Dubow, E. F., Goldstein, S. E., & Heretick, D. M. L. (2004). Social-cognitive mediators of the relation of environmental and emotion regulation factors to children's aggression. *Aggressive Behavior, 30*, 389–408.
- Nelsen, J. (2006). *Positive discipline*. New York: Ballantine Books.
- Orue, I., Bushman, B. J., Calvete, E., Thomaes, S., Orobio de Castro, B., & Hutteman, R. (2011). Monkey see, monkey do, monkey hurt: Longitudinal effects of exposure to violence on children's

- aggressive behavior. *Social Psychological and Personality Science*, 2, 432–437.
- Patterson, G. (1976). *Families: Applications of social learning to family life*. Champaign: Research.
- Patti, J., & Cepeda, A. (2007). Citizenship competencies in Colombia: Learning from policy and practice. *Conflict Resolution Quarterly*, 25, 109–125.
- Ramos, C., Nieto, A. M., & Chaux, E. (2007). *Aulas en Paz*: Preliminary results of a multi-component program. *Interamerican Journal of Education for Democracy*, 1, 36–56.
- Schulz, W., Ainley, J., Friedman, T., & Lietz, P. (2011). *ICCS 2009 Latin American Report. Civic knowledge and attitudes among lower-secondary students in six Latin American countries*. Amsterdam: International Association for the Evaluation of Educational Achievement (IEA). ACER. Nfer. Università degli Studi Roma Tre.
- Schwartz, D., & Proctor, L. J. (2000). Community violence exposure and children's social adjustment in the school peer group: The mediating roles of emotion regulation and social cognition. *Journal of Consulting and Clinical Psychology*, 68, 670–683.
- Selman, R. L. (1980). *The growth of interpersonal understanding: Developmental and clinical analyses*. New York: Academic.
- Singer, J., & Willett, J. (2003). *Applied longitudinal data analysis: Modeling change and event occurrence*. New York: Oxford University Press.
- Slaby, R. G., & Guerra, N. G. (1988). Cognitive mediators of aggression in adolescent offenders: 1. Assessment. *Developmental Psychology*, 24, 580–588.
- Social and Character Development Research Consortium (2010). *Efficacy of schoolwide programs to promote social and character development and reduce problem behavior in elementary school children (NCER 2011-2001)*. Washington, DC: National Center for Education Research, Institute of Education Sciences, U.S. Department of Education.
- Tremblay, R. E., Pagani, L., Mâsse, L. C., Pagani, F., & Pihl, R. O. (1995). A bimodal preventive intervention for disruptive kindergarten boys: Its impact through mid-adolescence. *Journal of Consulting and Clinical Psychology*, 63, 560–568.
- Tremblay, R. E., Mâsse, L. C., Pagani, L., & Vitaro, F. (1996). From childhood physical aggression to adolescent maladjustment: The Montreal prevention experiment. In R. D. Peters & R. J. McMahon (Eds.), *Preventing childhood disorders, substance abuse, and delinquency* (Vol. 3). Thousand Oaks: Banff International Behavioral Science Series, Sage.
- Velásquez, A.M. (2005). *Desarrollo de la asertividad: Comparación entre dos intervenciones pedagógicas*. Master's Thesis in Education. Bogotá: Universidad de los Andes.
- Velásquez, A. M., & Chaux, E. (2005). *AVC-agresión, violencia & competencias*. Bogotá: Universidad de los Andes. Unpublished document.
- Walker, P., Selman, R., & Snow, C. (2008). *Voices reading*. Columbus: Zaner-Bloser.
- Webster-Stratton, C., Reid, M. J., & Stoolmiller, M. (2008). Preventing conduct problems and improving school readiness: Evaluation of the incredible years teacher and child training programs in high-risk schools. *Journal of Child Psychology and Psychiatry*, 49, 471–488.
- Yeager, D. S., Trzesniewski, K. H., & Dweck, C. S. (2013). An implicit theories intervention reduces adolescent aggression in response to victimization and exclusion. *Child Development*, 84, 970–988. doi:10.1111/cdev.12003.
- Zhai, F., Raver, C. C., & Jones, S. (2012). Academic performance of subsequent schools and impacts of early interventions: Evidence from a randomized controlled trial in Head Start settings. *Children and Youth Services Review*, 34, 946–954.
- Zimbardo, P. G. (2007). *The Lucifer effect: Understanding how good people turn evil*. New York: Random House.