Preventing Violence in the Most Violent Contexts

Behavioral and Neurophysiological Evidence

Lelys Dinarte
Pablo Egana-delSol

WORLD BANK GROUP

Abstract

This paper provides experimental evidence of the impact of an after-school program on vulnerable public-school students in El Salvador. The program combined a behavioral intervention with ludic activities for students aged 10–16 years old. The authors hypothesize that it affects violence, misbehaviors, and academic outcomes by modulating emotional regulation or automatic reactions to external stimuli. Results indicate the program reduced reports of bad behavior and school absenteeism while increasing students' grades. Neurophysiological results suggest that the impacts on behavior and academic performance are driven by the positive effects of the program on emotional regulation. Finally, the study finds positive spillover effects for untreated children.

Preventing Violence in the Most Violent Contexts: Behavioral and Neurophysiological Evidence[*]

Lelys Dinarte[†]   Pablo Egana-delSol[‡]

Keywords: After-School Programs, Emotional Regulation, Violence, Behavioral Economics.

JEL Classification: I29, K42, Z13, I25, D87

1. Introduction

Violence and crime cause critical welfare losses in the developing world by forcing countries to spend substantial amounts of public and private resources to reduce their adverse effects (Krug et al., 2002; Soares and Naritomi, 2010).[1] In addition to direct economic and social costs, another concern is the "snowball effect" of this high exposure to violence during the early stages of an individual's life. Evidence shows children and adolescents with early exposure to violence tend to be involved in crimes later in life (Sousa et al., 2011; Damm and Dustmann, 2014).[2] Considering this, how vulnerable to violence and crime are young people in developing countries? Recent statistics reveal high exposure across all stages of childhood and adolescence. For example, 43% of total worldwide homicides occur among youth between 10 and 29 years old, and nearly all of these deaths occur in developing countries (WHO, 2016). Moreover, violent deaths are more common during adolescence. In 2015, there were roughly 119,000 violent deaths among children and adolescents below age 20, and two in three of those victims were aged 10 to 19 (UNICEF, 2017). Differences even emerge within a life stage; adolescents between 15 and 19 years of age are three times more likely to die violently than younger adolescents aged 10 to 14 (UNICEF, 2017).

How can we tackle young people's exposure to and participation in crime? An after-school program (ASP) is an intervention that can protect children by preventing victimization and delinquent behavior (Gottfredson et al., 2007; Mahoney et al., 2001). These programs can also be an alternative source of learning and social development when they include a specific curriculum oriented to foster socioemotional skills (Taheri and Welsh, 2016; Durlak et al., 2010; Eccles and Templeton, 2002). They are often implemented in vulnerable schools where children have a high risk of engaging in or being affected by criminal activities. However, despite an increase in the number of programs implemented over the past years and the high incidence and economic costs of violence in the developing world, evidence of the impact of ASPs on social skills, crime, and violence is mixed and inconclusive (Taheri and Welsh, 2016); most of the literature does not use data from places where the issue of youth involvement with violence is most relevant (Goldschmidt et al., 2007; Hirsch et al., 2011; Biggart et al., 2014).

This article aims to contribute to the current economic literature in three dimensions. First, following recent evidence of how psychological interventions such as cognitive behavioral therapy

---

[1]In Latin America and the Caribbean—the most violent region in the world—crimes cost 3.5% of GDP on average and up to 7% in Central America (Jaitman et al., 2017), with an average cost of crime of around US$300 per capita for each country. These costs break down into 42% in public spending (mostly police services), 37% in private spending, and 21% in social costs of crime, mainly victimization (Jaitman et al., 2017).

[2]Recent papers show this exposure can occur in all domains such as at children's homes (Baker and Hoekstra, 2010), through their interaction with other peers at school (Sousa et al., 2011; Herrenkohl et al., 2008), in their neighborhoods (Damm and Dustmann, 2014; Chetty et al., 2016), and more recently on the Internet.

(CBT) can efficiently improve individuals' misbehavior and criminal outcomes (Blattman et al., 2017; Heller et al., 2017),[3] this paper provides experimental evidence designed to measure the effect of a CBT-inspired ASP to enhance impulse control and conflict management on participants' violent behaviors and academic outcomes in the context of a developing and highly violent country. Second, we also study differences in the ASP's impacts by students' initial propensity for violence. Third, by exploiting the quasi-exogenous exposure to treated participants, we study spillover effects on misbehavior and academic attainment of at-risk children who were not enrolled in the ASP. Finally, by collecting neurophysiological recordings, we estimate the ASP's impact on participant emotional regulation, which in our context is a proxy of motivation or disposition to act, and stress.[4] We discuss how this effect may be a mechanism of how the ASP modifies behaviors and academic performance. To our knowledge, this is the first experimental assessment of a CBT-inspired ASP's impact from a country with these characteristics and on these categories of outcomes.

The intervention was performed in five public schools located in violent communities in El Salvador. This context is key for two reasons. First, El Salvador is a lower-middle-income country defined as a victim of an "epidemic of violence" since 2009 (WHO, 2011) and categorized as one of the top five deadliest places for adolescent boys in 2015 (UNICEF, 2017).[5] Second, El Salvador's high violence levels and homicide rates have significantly affected its educational system in the last years. The country has witnessed a 13% reduction in its education enrollment rate (MINED, 2015),[6] with approximately 18% of students reporting they dropped out of school due to delinquency.[7] Also, in the past five years, more children and adolescents there have been victims of homicide than in the previous two decades (EPCD 2014).[8]

The ASP we study in this article consists of clubs implemented after school hours within school

---

[3]For example, CBT is a therapeutic approach used to treat harmful beliefs and behaviors by making people aware of these patterns and trying to disrupt them through a "learning by doing process" (Blattman et al., 2017).

[4]As we explain below, "emotional regulation" is defined as a mixture of cognitive and emotional processes that affect a disposition to act (Salzman and Fusi, 2010). It involves consciously affecting one's own emotional and physical responses to given stimuli.

[5]Between 2009 and 2012, the country's average homicide rate was 69 murders per 100,000 inhabitants (PNUD, 2013). Over half of those killed during this time period were 15 to 34 years old; approximately 80% of the victims were male; 70% were executed using firearms; and nearly 40% of the murders took place in public spaces (FUN-DAUNGO, 2013). In 2015, El Salvador was the country with the highest murder rate in the world: 103 killings per 100,000 inhabitants. As a reference, the worldwide homicide rate is 6.2 murders per 100,000 inhabitants (PNUD, 2013).

[6]In 2013, primary and secondary net enrollment rates were 93.4% and 61.6% respectively, after a relevant drop in 2015, when primary and secondary net enrollment rates were only 86.2% and 37.9% respectively (MINED, 2015).

[7]This may be a lower bound because 28.6% of students abandoned school due to change of address, which since 2010 has been highly correlated with gang threats according to testimonies elicited by local newspapers (LPG, 2016).

[8]From 2005 to 2013, approximately 6,300 youth were homicide victims. In 2013, 458 adolescents were charged with extortion and 321 with aggravated homicide (CSJ, 2014), which are crimes mainly related to gangs (PNC, 2014).

3

facilities from April to mid-October of the 2016 academic year. The study sample includes 1,056 enrolled students aged 10 to 16 years old.[9] This age range is relevant in the Salvadoran context because that is when children and adolescents are likely to be recruited by gangs.

All randomly selected participants attended two sessions per week that lasted 1.5 hours each. The intervention was implemented by volunteers of the Salvadoran branch of Glasswing International, a nongovernmental organization (NGO) working in Central America and Mexico. Every session is a combination of: (1) a discussion oriented toward fostering children's conflict management, violence awareness, and social skills; and (2) the implementation of curricula including activities such as scientific experiments, artistic performances, and others.

In the first part, instructors discuss concrete methods for regulating participants' violent behavior using experiential learning or role-playing. This allows participants to take action in their lives, face difficulties, and learn how to solve problems proactively. To apply this therapy in the ASP, a tutor implements an experiential learning activity in which she mentions day-to-day difficulties that children typically face in low-income contexts. She asks them to brainstorm solutions to those problems and evaluate them. The second part of the session includes the implementation of ludic activities related to each club category. This section aims to motivate students to participate in the intervention and increase ASP attendance.

To measure the program's overall impact and exploit excess demand, we rely on the first stage of Dinarte (2018)'s experimental setting. We randomly granted some of the enrolled students participation in the intervention while the remaining students were assigned to a comparison group.

Before the intervention, we collected self-reported data on personal and family characteristics from enrolled students. Follow-up self-reported data included questions to measure the intervention's impact on attitudes, violence, and crime; exposure to risky spaces; and educational or personal expectations. We combined this self-reported information with administrative records on math, reading, and science grades; behavioral reports; and absenteeism data from both enrolled and nonenrolled students. Schools provided these data before and after the intervention.

Additionally, we collected neurophysiological evidence from a random subsample of enrolled students at the end of the intervention. By using the James-Lange theory of valence and arousal (Davidson et al., 1990; Harmon-Jones et al., 2010; Ramirez and Vamvakousis, 2012; Verma and Tiwary, 2014), and following the methodology developed by Egana-delSol (2016b), we proxied participants' emotional regulation and stress directly from their brain activity. We used portable

---

[9]As we explain in detail later, there are two samples in this study. The first one is enrolled students—those who decided to participate in the ASP and were randomly assigned to treatment or control groups (1,056 students). The second sample of nonenrolled students includes those who were not registered for the ASP but were in the same schools and classrooms as treated students (1,364 students).

electroencephalograms within a lab-in-the-field setting.[10]

We find that this low-intensity intervention works in the context of a developing and highly violent country, and that its short-term effects are similar in magnitudes and signs to those of middle-intensity interventions in the United States. (Durlak et al., 2010; Cook et al., 2010). For example, our estimations indicate that students assigned to treatment have better attitudes toward school and reduce absenteeism by 23%. Moreover, we find a reduction in violence and misbehavior at school in both student and teacher reports.

In line with the evidence that emotional and behavioral—that is, noncognitive—skills promote and indirectly influence cognitive development (Cook et al., 2011; Cunha and Heckman, 2008), we also find the ASP successfully increases participant academic achievement. On average, after seven months of intervention, grades are 0.11–0.13 standard deviations higher for treated students. The intervention also reduces the probability of course repetition by 2.8 percentage points.

Overall, these effects are consistent with the expected results from learning and protection services that the ASP can deliver. Specifically, these interventions can provide an innovative learning structure for students, affecting their disposition toward school and learning. Additionally, these programs can promote skills such as resilience and control over automatic responses and bad behavior. Finally, the ASP can provide protection from unsafe neighborhoods by reducing the time children and adolescents may spend with delinquent peers. Unfortunately, the experimental design does not allow us to disentangle these mechanisms, so we can only provide suggestive evidence that the learning channel is likely driving all the effects.

Differences in ASP effects by students' initial propensity for violence show the program benefits both high- and low-violence children but in different categories of outcomes. Highly violent children are driving the effects on academic performance and less violent students are performing better with regard to misbehavior at school and attitudes toward school and learning. We argue that automatic responses that generate bad behavior may be harder to modify, especially for those accustomed to acting that way or whose individual and family characteristics limit their improvement. In other words, a highly violent child may find it easier to attend school more regularly than to stop hitting classmates.

We then study spillover effects in this context. Our estimations indicate the ASP also has indirect short-term effects on nonenrolled children. Exploiting the exogenous percentage of treated students within each classroom, we find positive spillover effects from the exposure of nonenrolled students to a higher proportion of treated classmates on both academic and violence outcomes. The magnitudes of these effects are 0.08–0.09 standard deviations on academic performance and

---

[10]The EEG devices and the experimental paradigm were piloted before going to the field in a lab setting at the Laboratory for Intelligent Imaging and Neural Computing (LIINC) at Columbia University (Egana-delSol, 2016b).

0.15 standard deviations on misbehaviors at school. Thus, the direct results previously described seem to be lower bounds of the total effect of the intervention.

Further analysis of heterogeneous spillover effects by intensity and proximity to treated classmates indicates that: (1) the greater the exposure of nonenrolled children to their treated classmates, the higher the spillovers; and (2) the spillover effects are greater if there is an intermediate proximity—that is, neither overly similar nor radically different—regarding misbehavior between treated and nonenrolled students within classrooms. This last result indicates that diversity can play an important role in enhancing these positive externalities.

A plausible explanation for the effects on participant behavior and academic performance is that the ASP modifies the psychological factors that prompt those attitudes and violent behaviors such as stress or automatic responses. Considering this, we move on to study the effects of the intervention on such neurophysiological aspects as emotional regulation and stress.

We present two main results on neurophysiological measures. First, for the estimations of the overall intention-to-treat (ITT) effects of program participation, we found an effect on emotional regulation in club participants as well as an impact on responsiveness to positive emotionally laden stimuli, compared to the control group. In particular, the program reduces their valence outcome by 0.35 standard deviations, indicating that participants become more phlegmatic and move more toward a withdrawal attitude or behavior, relative to the comparison group. Additionally, treated students report a reduction in their internal locus of control test by 0.25 standard deviations, compared to what students in the control group reported. Thus, treated students perceive that they can manage or control what happens in their lives at a greater magnitude than nontreated ones.

Second, comparing low-violence students assigned to treatment versus similar enrolled children assigned to a control group, we find that the effects on treated students are driving both the reductions in valence and in the perception that they can control their circumstances by 0.57 and 0.47 standard deviations, respectively. Unexpectedly, we also find highly violent treated children to have higher stress levels compared to both similarly violent peers in the control group and to treated children with low propensity for violence.

We argue that composition of peers within the ASP in terms of violence can explain this unintended result. Using the same sample, Dinarte (2018) finds that tracking by violence increases both stress level and probability of misbehavior at school for the most violent children, compared to treating them in a more violence-diverse group. In this sense, to avoid unintended neurophysiological effects on the highly violent children and, therefore, to increase the ASP's total effectiveness, a well-designed ASP must include the definition of an optimal distribution of participants with unlike

6

propensity for violence.

This paper relates to a wide literature that aims to measure the effects of ASPs on academic outcomes and violence (Gottfredson et al., 2004; Goldschmidt et al., 2007; Hirsch et al., 2011; Taheri and Welsh, 2016). As mentioned before, gaps remain despite extensive analysis of this topic. First, the literature has focused on the effects of these interventions in developed countries, mainly in the United States, a context that may have limited applicability for low- and middle-income countries. Thus one way we contribute to this literature is by providing evidence of the effect of this intervention in a developing and highly violent country, where these programs can be more relevant.[11]

Second, results from this paper can also contribute to debate on the quality of schooling and learning in Latin America—a discussion of foremost importance in recent years. Part of the lack of quality education in the region can be associated with problems such as violence, which are part of the learning context, rather than with classroom practices.

Third, the paper also engages with a recent and novel literature that studies the effects of psychological interventions such as CBT on youth and adult crime and violence patterns. Seminal papers are those of Heller et al. (2017) in Chicago and Blattman et al. (2017) in Liberia. Our paper differs mainly in our test of a hybrid structure of CBT plus ludic ASP activities.[12] This mixed structure may be more effective in the context of Salvadoran schools since a full therapy may be hard to implement if the target group consists of children and adolescents, or if enrollment and participation in the program are not mandatory.

To our knowledge, this is the first paper to assess how these programs change youth emotional regulation, which can be one mechanism explaining the ASP's behavioral and academic impacts. Except for Heller et al. (2017), there is no hard evidence of how these programs change participant behavior considering automaticity and impulse control (Kahneman, 2011; Fudenberg and Levine, 2006). These outcomes are relevant to the economics perspective since an individual's criminal actions may reflect how s/he manages her or his automatic responses and self-control when facing different events or levels of violence exposure. Additionally, there is evidence that the type of emotions a person faces is relevant to many cognitive and behavioral outcomes such as attention, memory, and perception (Damasio, 1994; Salzman and Fusi, 2010; Fuster, 2013).[13] In that sense,

---

[11]Additionally, most of the ASP literature measures heterogeneous effects only by initial academic attainment, gender, or household income (Marshall et al., 1997; Durlak et al., 2010), without considering variables such as violence that may affect this kind of intervention. In this sense, our results are novel because the ASP in this particular context generates a differential impact according to participant violence levels, most positively impacting the misbehavior and attitudes of the most vulnerable children.

[12]The program we analyze is more similar to one of the interventions in Heller et al. (2017) that included a CBT approach and additional activities like sports and dancing, among others.

[13]According to DellaVigna (2009), even slight manipulations of an individual's mood have a substantial impact

7

people exposed to highly risky environments might suffer more substantial differences compared to their less exposed peers when learning and developing cognitive and socioemotional skills. This in turn may create or widen a gap in educational or labor market outcomes.

The methodology for measuring neurophysiological outcomes proposed in this paper has many benefits. First, it offers a way to incorporate objective proxies of emotional dimensions into the fields of education and violence economics. The importance of emotional regulation in life satisfaction and labor market outcomes has recently been highlighted for both developed and developing economies (Deming, 2017; OECD, 2015). This study shows there are neurophysiological approaches to proxy emotional disposition and responsiveness with a high level of accuracy and at relatively low cost in violent communities. Second, the results may also aid evaluation of similar programs oriented to improving noncognitive skills.

The remainder of the paper is organized as follows: Section 2 describes the intervention, data collection, and study design. Section 3 summarizes the specifications used to estimate effects of the intervention on academic behavior, violence outcomes, and spillover effects in this context. These results are presented in Section 4. Section 5 discusses emotional regulation as a driver to reduce violence and improve academic attainment. Section 6 reviews the protection and learning mechanisms. Finally, we present preliminary conclusions in Section 7. All tables, appendix tables, and appendix figures are at the end of this paper.

2. Intervention, Experimental Design, and Data

2.1 Intervention

After-School Clubs

Since 2013, the NGO Glasswing International has implemented extracurricular clubs as part of its program *Community Schools*, which has taken place in 95 schools in Central America through 560 clubs benefiting approximately 20,000 children between 8 and 15 years old. According to the intervention approach, the main objective is to successfully modify children's violence, misbehavior, and attitudes through the acquisition of life skills, and therefore improve their academic performance (Glasswing International, 2012a).

Clubs meet twice a week for approximately 1.5 hours per session just after school ends. Each session includes two parts: social skills development and the traditional club curriculum. The first section is common to all participants and includes activities oriented to foster socioemotional skills,

---

on his or her behavior, both in the short and medium term. These emotions are often defined by the environment in which individuals are involved, such as their communities, schools, or homes.

inspired by CBT activities. Specifically, it tries to raise participant awareness of certain behaviors, to disrupt these patterns, and to promote better ones using experiential learning.

An example of strategies used in this first section is "cognitive development." It aims to foster the development and use of thinking as a process and tool to solve problems and transform conflicts. This strategy can be applied to topics such as conflict and impulsiveness management, self-discipline, school violence reduction, and "soft skills." For example, in the topic self-discipline, the instructor organize participants in a students' circle and, by applying strategies of cognitive development, she asks them to provide alternatives to get a ball from a club mate. Some of them suggest forcibly retrieving it either by hitting the ball or the club mate. Then the tutor engage them to think deeper on additional alternatives such as negotiation or simply asking for the ball. Implementation of this first section is uniform across schools.

The second part of the session includes ludic activities related to each club category. The NGO offers four categories of clubs by education level: leadership, art and culture, sports, and science. The objective of the club curriculum section is to motivate students to participate, make the learning process more fun and interactive, and increase attendance at the ASP. For example, in the science category, discovery clubs offer students opportunities to do experiments such as mock volcano eruptions. The art and culture category includes dancing, singing, and other activities to develop fine motor skills and creativity. In the sports category, children play soccer or basketball. Finally, leadership clubs attract those who want to develop social and leadership skills.

The combination of social skills development and ludic activities is another innovation of this intervention. This mixed approach can be more effective than full-therapy interventions in the Salvadoran context because the greatest manifestation of violence starts during childhood or adolescence—around 10 to 16 years of age—when individuals are more prone to gang recruitment (Rivera, 2013). Adolescents may find it unappealing to learn impulsiveness management and self-discipline alone. Thus, to guarantee participant attendance, it was necessary to complement the therapy with some recreational activities.

The ASP is organized by a school coordinator who verifies attendance at and attrition from the program, manages club materials, and assigns volunteers as tutors. These tutors have no formal training in social work or psychology and, like those at the Chicago program *Becoming a Man* (BAM), they do not necessarily have similar backgrounds to the participants (Heller et al., 2017).[14]

To our knowledge, there are only two qualitative and nonexperimental reports on this ASP that

---

[14]There are three categories of volunteers: community volunteers are tutors living in the community who stand out for their leadership skills; corporate volunteers are part of a particular firm that has a social project with Glasswing; and independent volunteers are usually college students doing social work. The NGO assessed these volunteers; even though they did not follow a pure random allocation procedure, there is nonetheless balance in observable tutor characteristics such as gender, age, and tutor type among treatment and control groups.

show improvements on participant social skills. In such assessments, treated students report greater tolerance of others, a reduction in interactions with bad peers, and an improvement in the overall classroom environment. In addition, some students report the program reinforces their academic experience in a more enjoyable way (Glasswing International, 2012b).

2.2 Recruitment and Enrollment Process

During 2016, the NGO offered and implemented this version of the ASP in five public schools in El Salvador. Using data from the country's 2015 Educational Census, we found that these schools were similar to the underlying population of public educational centers in El Salvador. Tests for differences between participant and nonparticipant schools are shown in Table A1 in the Appendix section.[15]

As part of the recruitment process, the NGO visited the schools to advertise the program and provided informational brochures and videos. Additionally, NGO staff invited participants from other schools to join them for these visits and share their experiences with prospective attendees. During these visits, the team gave a parental authorization form to children interested in participating in both the intervention and study, asking them to return it with a signature from a parent or tutor. They also gave teachers extra forms for students who were absent during the visit but might want to participate.

Some weeks later, the NGO and research team returned to schools to register and enroll children. Out of a total of 2,420 students from the five schools, we recruited and enrolled 1,056 students between 10 and 16 years of age. Children were allowed to self-enroll as long as they had an authorization signed by an appropriate adult.

During the registration process, students completed a form with personal and family information and their application to participate in a club. They were next assigned to a group, taking into consideration their preferences and the aggregated demand for the club category. Originally, club enrollments varied according to the number of participants interested in each category. However, for methodological reasons, our club sizes are between 13 and 15 participants on average.

2.3 Sample and Randomization

As we mentioned before, this paper provides experimental evidence for two gaps in economic literature. While it measures the direct impact an ASP has on behavioral, academic, and violence-related

---

[15]Both groups of schools are similar on characteristics such as location area, violence level, number of students, and additional revenues. Similarly, in terms of programs, facilities, and equipment, participant and nonparticipant schools are similar on most of these dimensions, except in the share of schools with a breakfast program or Internet access; treated schools are more likely to have both benefits.

outcomes, it also studies the ASP's indirect effects on nonenrolled children and investigates emotional regulation as a plausible mechanism to explain the direct effects of the intervention. To address these questions, we rely on the first stage of Dinarte's (2018) experimental setting.[16]

The randomization process and sample distribution of this first stage are shown in Figure 1, panel A, which shows there are two groups in this study. The first sample of "enrolled" children comprises the 1,056 students who applied to participate in the program and then were randomly assigned to treatment (T) or control (C) groups. The second sample of "nonenrolled" students consists of 1,364 children who were not interested in the ASP. Using available administrative data for both groups, we are able to compare enrolled and nonenrolled children's characteristics. We return to this later.

Students from the "enrolled" sample were the unit of randomization and were randomly assigned to either C (25%) or T (75%) within school-by-educational-level "blocks," as shown in Figure 1, panel B. Each education level consists of three years of schooling: the first is from first to third grades, the second from fourth to sixth grades, and the third from seventh to ninth grades.

This uneven distribution between both T and C groups addresses the need to simultaneously reduce the number of children left out of the program during 2016 in light of the high out-of-school risk these children face, but it guarantees there is enough power to capture the effects (Dinarte, 2018).

Children in T group could participate in the two sections of the ASP. Students in C group were supposed to leave school facilities after classes ended. We were able to collect their information at follow-up because we gave them an "enrollment coupon" redeemable the next year that guaranteed their participation in the traditional clubs in 2017.

Since enrolled students were assigned at the school-by-education-level block, it is important to note that the share of treated children from each grade within each education level—after controlling by the share of enrolled children—was quasi-exogenous. This generates an important variation that allows us to estimate spillover effects on nontreated participants. We discuss this in more detail in an upcoming section.

---

[16]The second stage of the experimental approach is described in Dinarte (2018). The author incorporated additional features in the design to study how peer composition, according to a predicted violence level at baseline, changes the effectiveness of the intervention. Further details and results can be found in that study.

Figure 1: Sample Structure and Randomization Process

Panel A. Sample Distribution

2420 Students From 5 public schools in El Salvador	
1056 enrolled students	**1364 non-enrolled students**

Block randomization

Control Group 25% (258)	Treatment Group 75% (798)

Panel B. Block Randomization Process

School A		
Level 1-A (1-3 grades)	Level 2-A (4-6 grades)	Level 3-A (7-9 grades)
Randomization	Randomization	Randomization

Control Group 25%	Treatment Group 75%	Control Group 25%	Treatment Group 75%	Control Group 25%	Treatment Group 75%

This figure shows the samples composition and block randomization procedure applied in this design, adapted from Dinarte (2018). From the total of enrolled children in each educational level $\{1,2,3\} \in$ school A, we randomly assigned 25% to C and 75% to T groups. The same procedure was implemented in the remaining schools.

2.4 Data

A. Baseline

During the registration phase, after the first three months of the school year and before the intervention, we collected two strands of data. First, enrolled students provided personal and family information on registration forms. For instance, we gathered information on enrolled children's age, gender, household composition, mother's education level, and adult supervision exposure, among other characteristics. We also collected school records of math, reading, and science grades; behavior reports;[17] and absenteeism data from both enrolled and nonenrolled children. Appendix 1

---

[17]In El Salvador, teachers complete behavior reports each quarter, which are presented in the following discrete scale: Excellent (E), Very Good (MB), Good (B), and Regular (R). It can be translated in a continuous scale that is comparable to course grades. In this paper, we use a reversed continuous scale to facilitate the interpretation and comparability to self-reported measures of violence and crime.

presents more details on these variables.

Despite the relevance of participant characterization in terms of violence before the intervention, the dangerous public school context in El Salvador prevented us from directly asking students about violence during the registration phase. Specifically, it was impossible for the NGO or schools to ensure that this personal information would be kept confidential during the study, an impediment that could endanger both children and instructors.

School behavior reports for each student were the only available administrative data that could have been used as a proxy for violence. Although these reports describe delinquent and violent actions committed within school facilities, they are not a comprehensive measure of violence because they do not include actions committed by children in other relevant domains such as at home and in their communities.

To address this constraint and to estimate a comprehensive proxy for violence, we rely on Dinarte (2018)'s prediction of violence for each enrolled child. Using available baseline data, Dinarte followed Chandler et al. (2011) and estimated a predictive model of violence and crime from existing data (FUSADES, 2015) using a two-sample least-squares strategy. First, the author estimated the likelihood of having committed a violent act V_f as a function of a wide range of covariates X_f, such as student characteristics (e.g., age, gender, time spent alone at home, and education level); children's household variables (e.g., residence area, mother's education, household composition); and school-level controls (e.g., school location and commuting time to school). Then, exploiting the availability of these variables in the registration forms of enrolled students, she predicted the measure of propensity for violence (IVV) for each child, using the vector of estimated coefficients.

In Dinarte (2018) a full section explains and discusses this estimation, including some robustness checks. Since the IVV may not be a perfect measure of violence, she provides some evidence that it is clearly the best proxy of propensity for violence given this particular context. For example, the author finds a positive and statistically significant correlation between the IVV estimation and misbehavior reports.

B. Follow-Up

The contents and structure of the intervention are oriented to directly affect noncognitive outcomes such as children's violence and misbehaviors. The program may also have some indirect effects on academic outcomes, since changes in children's noncognitive skills could affect the learning process and thereby improve academic performance (Cunha and Heckman, 2008; Heckman and Kautz, 2012). Considering this, we collected data on these two categories of outcomes.

Follow-up data on noncognitive outcomes were collected only from enrolled participants in school

facilities at the end of October 2016, after all clubs completed their curricula. Students took the survey in classrooms especially set up for this purpose. Each survey took approximately 45 minutes. Most surveys were self-administered with assistance from staff trained in the survey methodology.

The follow-up survey included questions to measure the intervention's impact on the following topics: student attitudes toward learning, violence and crime; exposure to risky spaces; and educational or personal expectations. To measure attitudes toward school and approval of a friend's criminal behavior, we used items from the *Communities That Care* Youth Survey. Delinquency and violence measures were calculated using the Self-Reported Delinquency Scale (SRD).

In the quantification of exposure to violence or crime, we used the nationwide El Salvador Youth Survey (ESYS), which was developed and validated in at-risk youth populations in El Salvador by Webb et al. (2016). It includes questions related to children's and adolescents' risk and protective factors in three domains: family, school, and community. The final implemented instrument is available upon request.

To increase statistical power to detect effects for outcomes within a family, we used indexes of variables that are expected to move in a similar direction and also to reduce the number of hypothesis tests (Haushofer and Fehr, 2014; Heller et al., 2017).

Due to the limitations of self-reported data, we attempted to recheck and validate these behaviors using proxies for these outcomes obtained from administrative data. For example, we were able to find school attendance in school records, which is a good proxy of students' positive motivation toward school and learning. Additionally, this administrative data provided measures of participant academic performance. In November 2016, at the end of the academic year, schools provided math, science, and reading grades; behavior reports; and school absenteeism and grade promotion data for students who were enrolled and not enrolled in the clubs. Finally, club attendance data were collected by club instructors after each session. We present a detailed description of all follow-up outcome variables measured in this paper in Appendix 1.[18]

To sum up the overall process of this study, we present a timeline in Figure 2.

Figure 2: Intervention, Randomization, and Data Collection Timeline
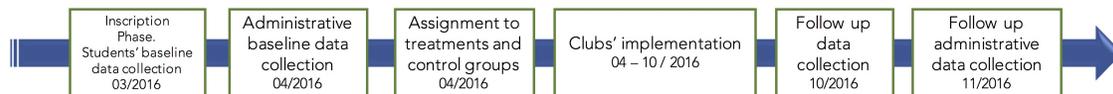
Timeline of the intervention and data collection.

---

[18]Although we aimed to match our sample to additional administrative datasets such as police records, employment, high-school/university enrollment, or any other longer-term follow-up outcomes for subsample aging out of school, it was not possible since there is no unique ID for individuals younger than 18 years of age in El Salvador.

14

C. Matching with Administrative Data and Attrition

The average matching rate of administrative data of enrolled children was 94% at baseline and 97% at follow-up. We present these estimations in Table A2 in the Appendix section. All the matching rates were balanced between T and C groups, except for the matching rate of science scores before the intervention. To account for this difference, we first impute the course average at baseline for those missing observations. Then, we include this variable with imputed values and a dummy of missing values in estimations of academic outcomes. Further details of these estimations are explained below.

As we show in Table A2 in the Appendix section, 968 of 1,056 baseline enrolled children (92%) were surveyed at endline. In the T group, 731 of 798 students (91%) were surveyed at endline, and in the C group, 237 of 258 (92%). This is a low attrition rate compared to those levels found in similar studies of at-risk youth. We regress the attrition dummy on the T dummy and the result shows no difference in the likelihood of attrition between the T and C groups. Therefore, results will not be driven by the absence of follow-up survey data for any group.

Finally, the average matching rate of administrative data of nonenrolled students was 85% at baseline and 98% at follow-up. In estimations using the nonenrolled sample, we follow the same imputation procedure for baseline variables as in the sample of enrolled children.

2.5 Summary Statistics

Descriptive statistics of the full sample and T and C groups are shown in Table 1. Column (1) exhibits statistics for the total sample and columns (2) and (3) are for C and T groups, respectively. Column (4) shows unadjusted p-values for tests of differences in means between T and C means, and column (5) presents adjusted p-values for multiple hypothesis testing of means and FWER following Jones et al. (2018). Each group of variables per panel constitutes a family for FWER control. All tests of differences in means control for randomization blocks and standard errors are clustered at the course-by-school level.

Panel A presents the summary statistics of the violence determinants. Participants are on average 12 years old, 49% are male, and 73% live in an urban area. Regarding family composition, 91% of the students live with at least one parent, and 9% live with a relative or an unrelated adult. On average, 62% of students' mothers have an intermediate education level (7–12 years), and 31% have less than six years of schooling. Regarding risk exposure, only 5% of students reported being alone at home when they are not at school. However, on average they travel around 18 minutes to school. Additionally, 29% of students are enrolled in the afternoon shift, increasing the probability

of being without adult supervision while their parents are at work.

The last row of panel A shows that the average propensity for violence for T and C groups is 0.038, with a standard deviation of 0.029, ranging from 0.001 to 0.215. This average propensity for violence is 14 times the mean probability that a given student will be vulnerable to violence in Chicago (Chandler et al., 2011). Even when both estimations are not completely comparable because we use fewer violence determinants than Chandler et al. (2011), this difference sheds light on the tremendous propensity for violence among the children in this study.

Panel B shows academic scores and absenteeism for the first quarter of the 2016 school year. We imputed average course grades for missing observations and present the same models using only subsamples with nonmissing information in Table A3 in the Appendix. In a grade scale of 0–10, requiring a minimum grade of 5 to pass each course, enrolled students have between 6.5 and 6.7 points, similar to average grades at national level. Additionally, 25% of enrolled children have a misbehavior report before the intervention. The mean absenteeism rate during the first quarter of the academic year—that is, before the intervention—was 4.8% (1.69 out of 35 days).

Panel C summarizes club characteristics: mean club size was 13 students and community tutors ran approximately 31% of these clubs. The average take-up, defined as the share of sessions attended by each student out of the total number, was 57%. Finally, the mean fraction of treated students by course was 42%.

2.6 Compliance

In terms of compliance, attendance records indicate that no children in C group received treatment. Additionally, school coordinators were instructed to make sure that only children in T group attended the courses. However, since we did not implement an additional compliance check, and considering that treated students attended approximately 60% of all club sessions, we use an intent-to-treat approach to measure the ASP impacts, as we explain in the following section. We consider all children assigned to participate in the program as T group regardless of whether they were part of the ASP at the time of the endline survey.

3. Estimation Framework

In this section, we describe the empirical strategy used to measure the ASP's effects on student behavior, violence, and academic outcomes, and to assess the heterogeneity of the intervention by individual violence levels. Additionally, we study how the intervention affects nonenrolled children's behaviors.

3.1 Measuring the ASP's Direct Impact

A. Intent-to-Treat Direct Effects of ASP Participation

Given our randomized experimental design, it is straightforward to measure the Intent-to-Treat (ITT) effects of ASP on noncognitive and academic outcomes. We use the random variation from the design and estimate the following equation:

$$y_{ijt} = \theta_0 + \theta_1 T_{ij} + \theta_2 X_{ij(t-1)} + S_j + \epsilon_{ijt} \tag{1}$$

where y_{ijt} is the post-ASP outcome for student i at school and education level j during postrandomization period t. T_{ij} is a dummy indicating that the student was randomly offered participation in the ASP. $X_{ij(t-1)}$ is a vector of control variables measured at baseline, including a second-order polynomial of the student's IVV percentile. As mentioned before, for the academic outcomes regressions, we included standardized grades at baseline (including imputed values) and a missing baseline grades indicator (i.e., equal to 1 when the student's grade was missed) as control variable. Finally, we also control for the "randomization block" with school-by-education-level fixed effects S_j. In this model, θ_1 captures the short-term ITT effect of being assigned to participate in an ASP compared to being randomly allocated to a control group.

Due to the possible bias in the estimation of the IVV, standard errors are adjusted using a cluster bootstrapped at the course-school level (Treiman 2009). As a sensitivity analysis, we also calculate p-values that account for multiple hypothesis testing and FWER following Jones et al. (2018). For the adjustment of p-values, we defined the following families of outcomes: (i) positive attitudes towards school, (ii) violence and misbehavior at school, (iii) intensive margin of academic outcomes, and (iv) extensive margin of academic attainment.

As an additional robustness check of the accuracy of the predicted IVV as a proxy for misbehavior, we estimate specification (1), yet instead of controlling by a second order polynomial of students' IVV percentile, we control by a similar polynomial specification of the student's percentile in the misbehavior distribution function.

B. Heterogeneity of the Intervention by Baseline Violence

Most of the existing ASP literature consistently finds differential effects of this type of interventions by individuals' gender, age, or household income (Marshall et al., 1997; Durlak et al., 2010). However, there is no empirical evidence of how ASP may affect individuals with unlike levels of violence. On the basis of this evidence, we can learn which children benefit more from the intervention and, in further studies, understand why their level of violence enhances gains from the intervention.

Exploiting the availability of a propensity-for-violence measure estimated by Dinarte (2018), we analyze if the ASP in this particular context generates a differential impact according to participant violence levels at baseline. As we explained before, this IVV measure does not capture effective violence, but a propensity. In addition, it is more comprehensive than misbehavior reports from schools because the IVV includes vulnerability to violence in different children's domains such as school, home, and communities.

To study these heterogeneous treatment effects, we include in equation (1) an interaction between T_{ij} and an indicator $IVV\_high_{ij(t-1)}$. This dummy classifies students whose percentile in the IVV distribution before the intervention was greater than the median at the group (C and T) and stratum level. Specifically, we estimate:

$$y_{ijt} = \theta_0 + \theta_1 T_{ij} + \theta_2 T_{ij} \times IVV\_high_{ij(t-1)} + \theta_3 X_{ij(t-1)} + S_j + \epsilon_{ijt} \qquad (2)$$

where θ_2 indicates the marginal impact of the intervention between treated students with high and low levels of propensity for violence. The rest of the variables are defined as in specification (1), and $X_{ij(t-1)}$ includes the dummy $IVV\_high_{ij(t-1)}$.

C. Spillovers on Nonenrolled Children

Besides the ASP's direct effects, spillovers from treated students on their nontreated classmates can occur via at least two channels: (1) better school climate: if treated children are less disruptive during classes and/or behave better, this can reduce violence within schools and improve the learning process for all; (2) John Henry effect: the interaction between treated and nontreated students can allow the latter to imitate or learn some skills from the former. If any of these situations occurs, estimations from the specification (1) may be lower bounds of the ASP's total impact due to the presence of spillovers from the program.

To calculate spillover effects, we have to use the sample of nontreated students. However, a plausible concern in this estimation is that students assigned to C group can be more motivated than their nonenrolled peers, proxied by their registration decision, and may have better socioemotional skills or behavior at school. Thus, by measuring spillovers on the C group we may be underestimating them, since the motivation effect seems to go in the same direction as the spillovers. Therefore, to capture the total indirect effect while accounting for a nonobservable motivation effect, we use only the sample of nonenrolled children.

Are the enrolled and nonenrolled samples similar on observables? Using administrative data on academic performance and misbehavior at school, we compare the characteristics of these two groups and find the groups are similar regarding academic performance and behavior at school

before the intervention. Table A4 in the Appendix presents these results.

We use the nonexperimental variation in the share of students who were randomized to treatment across courses to estimate spillover effects. Recalling that (i) the assignment to treatment was done at the school-by-education-level block and (ii) each level includes three courses, we can assume that the share of enrolled children allocated to participate in the ASP at each course n—that is, the share of treated students Sh_n—was quasi-exogenous. Considering this, we can follow Carrell et al. (2013) to measure the ASP's spillover effects on nonenrolled students m.

A possible concern about this estimation is that nonenrolled participants may have influenced the enrollment decision, thus indirectly affecting the share of classmates assigned to treatment Sh_n. To address this concern, we include as a control the share of all enrolled students—T and C groups—from each course, E_n. The final specification will be the following:

$$y_{mnt} = \gamma_0 + \gamma_1 Sh_n + \gamma_2 X_{mn(t-1)} + E_n + \epsilon_{mnt} \tag{3}$$

where y_{mnt} is the academic or misbehavior outcome measured at the end of the academic year. $X_{mn(t-1)}$ is a vector of individual controls, including imputed grades at the baseline and missing grades dummies.

4. Results

In this section we present reduced-form estimates and heterogeneity analysis of the ASP's impact on student grades, violence, bad behavior at school, and attitudes toward school and learning. We also describe the ASP's spillover effects on nonenrolled students. There are three main conclusions. First, this ASP successfully modifies children's violence and academic performance in magnitudes that are similar to those found in the context of developed countries. Second, both highly and less-violent children gain from this CBT-based ASP. Finally, the intervention has positive spillover effects on nonenrolled children's academic outcomes and misbehavior.

4.1 Measuring the Overall ASP's Impact

A. Intent-to-Treat Effects of ASP Participation

Table 2 shows results of equation (1). We split them into two sets of outcomes: positive attitudes toward school, violence, and misbehavior at school (panel A), and academic outcomes (panel B).

First, in columns (1) – (4) in panel A, we present the ASP's effects on students' pro-learning attitudes from both self-reports and administrative data. Compared to students in the C group, ASP

19

participants report having better attitudes toward school by 0.17 standard deviations and spending 16% more time (20.4 minutes approximately) each day doing their homework. Moreover, 7.9% report that they pay more attention during classes, compared to the C group. This improvement in attitudes is also confirmed using administrative data: treated students are absent 1.6 days fewer per week than students in the comparison group. This implies a reduction of 23% in school absenteeism. These effects shed light that the ASP directly affects students' positive attitudes toward school as the program may allow them to be involved in a different, potentially more interesting learning approach, or to be exposed to a new category of role models—tutors—along with their teachers.

Then, we estimate the ITT effect on misbehavior and violence-related outcomes, using measures from student and teacher reports. As we can see in columns (5) – (9) in panel A, after seven months of intervention, students report having committed fewer delinquent actions and being less violent compared to reports of students in the C group (in magnitudes of 0.19 and 0.14 standard deviations, respectively). Similar effects are found using teacher reports. Students randomly assigned to participate in the ASP reduced both their bad behavior at school by 0.17 standard deviations and their probability of having a misbehavior report by 6.4 percentage points.[19]

Although our two sets of measures are not completely comparable since they differ by report source and the items and domains included—family, school, or community—results from both are consistent with an increase in participant willingness to reduce bad behavior and tendencies to violence. Combining these two groups of results, the effects we find from the intervention are similar to those previously identified in the literature. For example, Durlak et al. (2010) find a reduction in criminal behavior by 0.19–0.30 standard deviations in a meta-analysis of ASPs implemented in the United States. Similarly, Heller et al. (2017) find that the BAM program for youth in Chicago reduced violent-crime arrests, improved school engagement, and increased graduation rates.

The existence of such impacts from this ASP is not surprising to the extent that the neuroscience literature suggests it is possible to affect noncognitive skills during adolescence. Existing work suggests that noncognitive investments during that life stage can have a positive impact on the development of noncognitive skills such as behavior. For instance, Heckman and Kautz (2012) find that "soft skills" are malleable during adolescence; given that skills produced at one stage raise the productivity of investments in subsequent stages, they can predict and causally affect individuals' success in life. In addition, Mahoney et al. (2010) and Cassel et al. (2000) posit that extracurricular involvement helps to dissuade students from involvement in delinquency and crime.

Although ASP activities do not relate directly to academic outcomes, there is evidence of a

---

[19]Differences in number of observations in noncognitive outcomes is because of variation in the response rate for each outcome. We estimated these results using the smallest sample (836 observations) and there were not significant differences in the results.

positive correlation between academic results and social skills (Durlak et al., 2011). Moreover, according to Cunha and Heckman (2008)'s model of skill formation technology, noncognitive skills foster cognitive skills. Thus, orienting resources toward their improvement is particularly relevant for individuals from disadvantaged families or backgrounds.

In this sense, we also present ITT results of the intervention on academic outcomes in Table 2, panel B, columns (1) – (4). Grades have been standardized at the course-school level. At the end of the academic year, the ASP has a positive effect on math and science grades, with a magnitude of 0.11 and 0.13 standard deviations, respectively (intensive margin).

Using the data on grades, we can also assess the ASP's short-term effect on the extensive margin, that is, on the probability of passing each course. Exploiting the fact that the minimum grade to pass a course in El Salvador is 5, we create for each child an indicator of having a score above that value in each course. Our estimations indicate that the intervention increases the probability of passing reading and science courses, and reduces the probability of grade repetition by 2.8 percentage points (panel B, column (8)). Under some equilibrium effects assumptions, from a total 36,491 children enrolled in public schools that offered repeat courses in El Salvador in 2015 (MINED 2016), this last effect might imply that 1,022 children could pass their course if this type of intervention were scaled up.

As a robustness check, we estimate the ASP's effects on the relevant outcomes, controlling by a second-order polynomial of student misbehavior at school via teacher reports. The estimated effects using this alternative specification are similar in magnitude and sign to those presented in Table 2. This result strengthens the argument that the predicted propensity for violence indeed measures student misbehavior. Estimations are presented in Table A5 of the Appendix.

Since this is a low- to intermediate-intensity ASP, the effects on academic outcomes are in between those results from high- and low-intensity programs. For example, Durlak et al. (2010) find that ASPs in the United States have an average positive impact of 0.12 standard deviations on school grades. However, Shulruf (2010) concludes that extracurricular activities with a duration of three hours per session, five times per week—that is, high-intensity programs—have an average effect of 0.30 standard deviations on math and science grades.

How does an intervention that only teaches social skills indirectly affect participants' academic grades? There can be at least two explanations. First, the ASP can modify students' classroom misconduct, reducing disruptions that affect their learning or that of their classmates. (Scott-Little et al., 2002; Durlak et al., 2010). This improved learning environment will therefore benefit both treated and nontreated children.

Second, as we briefly explained before, a large body of theoretical and empirical work in eco-

21

nomics and psychology (Borghans et al., 2008; Cunha and Heckman, 2008; Dodge et al., 1990; Heckman et al., 2006; Moffitt et al., 2011) shows that cognitive skills and academic performance are defined by noncognitive skills such as future orientation and attitudes toward school and learning. In this sense, an effective investment in socioemotional skills can directly foster cognitive ability of beneficiaries, especially if those improvements occur at an early life stage.

B. Heterogeneity of the Intervention by Baseline Violence

Table 3 summarizes the estimated effects from specification (2). Coefficients in row [i] in each panel show the ASP's effects on low-violence treated students compared to low-violence children in the C group, coefficients in row [ii] show the differences in effects between high-violence treated students and similar children in the C group, and coefficients in row [iii] point to the difference in effects between high- and low-violence treated students. Row [iv] indicates adjusted p-values of the test for difference in effects between high- and low-violence treated students.

From our estimations presented in panel A, we can draw two main conclusions. First, while the ASP successfully modifies attitudes toward school and learning behaviors of both high- and low-violence treated children, compared to their respective peers in the C group, most of the effects are driven by the least violent children. Moreover, the difference in the intensive margin of misbehavior at school—reported by teachers—between high- and low-violence students shows the reduction is greater for the low-violence group of treated children, as we can see in column (8), row [iii].

Evidence of individuals' violence trends might help to explain these results. For example, the automatic responses that generate bad behavior are harder to modify, particularly for those who are accustomed to acting that way or whose individual and family characteristics limit the probability of improvement. From a neurophysiological perspective, Lewis et al. (1979) find that more violent individuals may have greater brain damage, which increases their tendency toward violence. As a policy implication, we may need more intensive interventions to successfully modify misbehaviors of individuals with the highest propensity for violence.

An additional interpretation is related to Akerlof and Kranton (2002)'s ideal student theory. They state that teachers and coaches reward or disapprove of students according to a "school's ideal student." In this sense, teachers may have already tagged students by initial violence level and, despite observing a reduction in bad behavior, they report that this decrease is greater for those who already been seen as the ideal low-violence student.

The second conclusion from this heterogeneous effects analysis is that estimations of the comparison between highly violent participants in T and C groups show that the ASP was more effective at reducing absenteeism of the most vulnerable children. We present this estimation in column (4)

row [ii]. The improvement on this attitude has important implications in terms of human capital accumulation, as we can see in Table 3, panel B. Our results indicate that highly violent treated children have greater academic performance compared to both similar peers in the control group and to low-violence treated children. Specifically, as we can see in row [iii], differences between high- and low-violence treated students' grades are 0.19–0.24 standard deviations.

Although there are no statistically significant differences on the extensive margin between both groups, a notable result from row [ii], column (9) in panel B is that the total effect on the probability of school promotion for highly violent treated students increases by 7.9 points, which accounts for approximately 10% of average course promotion from the C group.[20]

Summing up the results from the heterogeneity analysis presented in Table 3, the second novel contribution from this experiment is that both tails of the propensity for violence distribution function seem to gain from this ASP on different sets of outcomes.

Exploiting the lack of correlation between students' school grades and their propensity for violence at baseline presented in Table A6 in the Appendix, we also explore heterogeneous effects by initial academic attainment on the outcomes of interest. Estimation approach and results are summarized in Appendix 2. Our findings indicate that the ASP also benefits students with lower academic grades before the intervention. Particularly, low-performing treated children at baseline face a greater effect on school absenteeism and on the extensive margin of academic grades after the intervention, compared to initially high-performing treated children.

Finally, as previous studies have found (Durlak et al., 2010), it may be expected that this ASP affects boys and girls differently. However, since the predicted IVV includes gender as a determinant, the difference of the effects among boys and girls may be due either to sex alone or to the combination of this variable and all determinants included in the IVV estimation. To account for this, we use an alternative specification to show that most of the differences in the effects we find in this section are driven mostly by student propensity for violence. A detailed description of the equation and estimations is presented in Appendix 3.

C. Effects on NonEnrolled Children: Spillovers

Using the sample of nonenrolled children, we estimate specification (3) to measure how exposure to a higher share of treated classmates affects academic and behavioral outcomes of nonenrolled students. This model controls by the proportion of enrolled children and includes school-by-education-level

---

[20]Further heterogeneous effects by initial level of violence are depicted in Appendix Figure A1. The graph shows the estimations of a local polynomial fit of standardized endline score grades by predicted IVV for T and C groups. There are statistical differences between both groups for students in the 55th to 95th percentiles in the IVV distribution.

fixed effects. Since we rely only on administrative data on nonenrolled students, spillover results are limited to school grades and behavior reports.

Table 4 shows the results of spillover effects. We find evidence that the interaction of students with a greater share of ASP participants generates positive effects on their reading, math and science grades, and reduces their bad behavior at school. Estimations indicate that adding 3 treated students in a classroom of 22 (almost a 1 standard deviation increase in treated students, 13.4% on average) increases academic achievement up to 0.095 standard deviations, (e.g., on math grades: $13.4\% \times 0.007 = 0.095$), and reduces bad behavior reports by 0.15 standard deviations $(13.4\% \times 0.011)$.[21]

These results have similar signs to some evidence previously found in the literature. For example, Carrel and Hoekstra (2010) use the share of classmates coming from troubled families—that is, the share of children exposed to domestic violence—to measure its effect on grades and classroom misbehavior. They found that making 5% of a class troubled students (1 standard deviation) significantly decreases reading and math test scores by 0.69 percentage points, and increases misbehavior in the classroom by 0.09 more infractions.

To sum up, the spillover results yield two findings. First, these positive spillovers on nonenrolled students indicate that the ASP's direct effects described in subsection 4.1-A are lower bounds of the total effect of the intervention in the context of these highly violent schools. Second, combining these results with those from Carrel and Hoekstra (2010), we can conclude that it is possible to outweigh the negative effects of misbehaving children by incorporating students with improved or positive behavior in their classrooms. This result particularly contributes to the evidence of optimal class design, indicating that finding the optimal share of bad-to-good peers can maximize the total effect of a program (Krueger, 2003; Lazear, 2001).

In Appendix 4 we present a more complete analysis of the structure and characteristics of these spillover effects, such as optimal combination of treated with high and low violence level, intensity of exposure, and proximity on misbehavior between enrolled and nonenrolled children within classrooms. What do we find?

First, from the results in Table 4 we learn that exposure to a greater share of treated children increases nonenrolled student performance. However, that group of treated classmates can have different levels of violence before the intervention. Therefore, an important missing factor in these results is whether the high- or the low-violence group is causing the main spillover effects. In other

---

[21]After adding individual controls, estimated coefficients are similar in magnitude and statistical significance except for bad behavior reports, which are no longer statistically significant due to the increase in the standard errors. Despite this, the sign of the effect is negative, indicating that a higher share of treated classmates reduces the effect on bad behavior reports and provides additional evidence of reduction in the formation of violence networks or disruption during classes.

words, is there an optimal combination of high- and low-violence treated children that maximized the aggregated effect?

As we describe in Appendix 4, the separated effects of the shares of high- and low-violence treated children are not statistically different from zero, particularly due to an increase in the standard errors. However, coefficients signs indicate that spillover effects on academic outcomes may be driven by the share of treated students with low levels of violence, while the reduction in misbehavior at school may be caused mainly by the share of treated students with higher propensity for violence. In this sense, exposing nonenrolled children to both groups with unlike levels of violence can increase the total spillover effect on different outcomes, either because treated children are role models for the rest of their class or due to a reduction in the violence focal points or disruptive behaviors.

Second, the intensity of these spillovers may change due to the exposure level—in terms of time length—of nonenrolled children to treated participants. Since nonenrolled children usually spend more time with treated students in their own class, one may expect that the spillover effects of exposure to a greater share of treated classmates are larger. In fact, our estimations indicate that spillovers on a nonenrolled student's outcomes are led mostly by the share of treated students from her own classroom, providing evidence for our hypothesis of better school climate.

Finally, spillover effects may be different by misbehavior closeness of nonenrolled with treated students. Since the ASP's effects are different by initial propensity for violence of treated participants, there may also exist heterogeneity in the spillover effects by initial nonenrolled students' misbehavior at school. The estimations we present in Appendix 4 show that most of the effects are greater for students whose bad behavior at school is intermediately away—between 1 and 2 standard deviations—from the average misbehavior of the share of treated students within their classroom. Particularly, the effects of this medium closeness are greater on bad behavior reports. Thus, this result highlights that only certain levels of similarity to treated students can have positive spillover effects. Similar to Dinarte (2018)'s findings, this last result indicates that diversity plays an important role in enhancing these positive externalities because it allows children to learn from peers with different levels of misbehavior.

5. Emotional Regulation Drives Reductions in Violence

Why did this intervention have such behavioral and academic impacts? We argue that the enhancement of participant capacity for emotional regulation is the main path through which to observe the ASP's impacts. In this section, we first present neurophysiological and economic evidence that

indicates how the emotional state of an individual can drive her behavioral responses to particular stimuli, which in turn affect her economic decisions. Then, we study how the ASP affects the subject's ability to regulate her emotions or disposition to act, in ways that enable her to make more measured and rational decisions that lead to more productive academic performance. Finally, we provide experimental evidence that the intervention enhances students' emotions and reduces their automatic responses.

5.1 Emotional Regulation and Economic Decision Making

Emotional regulation (ER) can be defined as a mixture of cognitive and emotional processes that shape a mental state and thus a disposition to act (Salzman and Fusi, 2010). It involves developing the ability to consciously affect one's own emotional and physical responses to given stimuli. Thus, self-control, internal locus of control, grit, and emotional intelligence differ from emotional regulation. For example, emotional intelligence is traditionally defined as a synthesis of four capabilities or competencies: self-awareness, self-management, social awareness, and social skills (Goleman, 2010), but it does not necessarily involve a disposition to act.

Papers in the behavioral economics literature argue that ER matters because emotions influence decision making and economic behavior (Haushofer and Fehr, 2014; DellaVigna, 2009; Loewenstein, 2000). For example, in the context of poverty, stress and emotional instability may lead to short-sighted and risk-averse decision making. In that sense, instead of adopting behaviors that might generate greater returns, individuals end up favoring habitual low-yield ones. This in turn can generate vicious cycles or psychological poverty traps (Haushofer and Fehr, 2014).

Similarly, recent economic evidence indicates a relationship between psychological factors, emotions, and exposure to violent shocks (Baysan et al., 2018; Card and Dahl, 2011; Moya, 2018; Baysan et al., 2018). Immediate emotions (such as stress or uncertainty) that at-risk individuals experience at the time of decision making can lead them to take extreme actions with long-lasting consequences that are less efficient than actions of agents less exposed to violence (Loewenstein, 2000).

Kahneman (2011) argues that individuals often respond to situations automatically according to their experience with situations they commonly face. For at-risk people, the inability to regulate emotion can make them more susceptible to respond to certain stimuli with violence. How can the ASP affect participants' emotions and thus change their misbehavior at school and violence responses? By fostering ER, the program was designed to reduce participant automaticity—similar to Heller et al. (2017)—and nurture conflict and violence management. As we mentioned before, the ASP included some experiential exercises that allow children to expose their traditional responses but also introduce them to nonviolent alternative responses.

5.2 Measuring Emotional Regulation

To measure the effects of the ASP on ER and automaticity, we follow an alternative approach to Heller et al. (2017). In that paper, authors implemented a dictator game to measure if at-risk treated youth made more deliberate decisions compared to individuals in the control group. In our study, we estimate objective ER measures collected from participants' brains to provide evidence of changes in their automatic responses and other neurophysiological outcomes, such as stress. Appendix 4 offers a more detailed explanation about previous evidence showing correlations of EEG-based, arousal and valence indexes with behaviors or outcomes.

As explained in Egana-delSol (2016a), self-reported measures are suboptimal for estimating the impact of interventions because they can be affected by other factors that may impact individuals' emotional states and be confounding with the treatment itself. As such, ER can be proxied using the emotion-detection theory from affective neuroscience literature. In this study, we use the James-Lange theory of valence and arousal. In this model, arousal is a proxy of an individual's stress, estimated directly from her brain activity. Valence can be interpreted as a positive or negative mood, as well as an attitude of either approach toward or withdrawal from a stimulus (Davidson et al., 1990; Harmon-Jones et al., 2010).

To generate these measures, we follow the methodology developed by Egana-delSol (2016b). We acquired portable EEG headsets to obtain a proxy measure of students' emotional states and responsiveness to stimuli in the arousal-valence locus (Ramirez and Vamvakousis, 2012). These devices provide an average accuracy or correct detection that is similar to research grade instruments and an emotion-detection accuracy of about 79% (Egana-delSol, 2016b; Martinez-Leon et al., 2016).

Due to the relatively high costs necessary to collect neurophysiological data, we randomly selected a subsample of 598 individuals from the total 1,056 enrolled children, with a similar distribution of C and T groups (25% and 75%, respectively). Descriptive statistics of this subsample are presented in Table A7 in the Appendix. Then, following Egana-delSol (2016a) we established a lab-in-the-field setting to collect three streams of neurophysiological data for each student: (1) pretest resting emotional state from EEG recordings, (2) psychometric tests to measure noncognitive, creative, and cognitive skills, and (3) emotional responsiveness to both positive and negative stimuli. (See Figure A2 in the Appendix as a reference of the lab-in-the-field setting).

In the first stage, the pretest resting emotional state, we collected EEG recordings to estimate emotional arousal and valence indexes at resting state while students watched a black cross in the center of a gray screen for a period of 30 seconds.

In the second section of data collection, students responded to a battery of psychometric tests, which included the Rotter Locus of Control Scale (Rotter, 1966), Raven-like progressive matrices,

and the Cognitive Reflection Test (CRT). Our measure of locus of control indicates children think they are unable to control what happens in their lives. More specifically, a decrease in locus of control indicates that students feel they can manage their experiences, thus demonstrating an increase in self-efficacy. Raven is a measure of abstract reasoning and a nonverbal estimate of intelligence. It is implemented as a set of matrices in progressive order. Finally, CRT is a test designed to measure if an individual tends to automatically choose an initially incorrect response and then engage in deeper reasoning to find a correct answer.

During the third stage, right after the students finished the battery of psychometric tests, we obtained emotional response intensity for negative and positive stimuli in terms of valence locus. Here, we exposed students to alternate series of images selected to elicit positive and negative emotional responses in order to estimate poststimuli valence indexes.

Then, using this response intensity measure and the valence-at-resting-state index from the first stage, we estimated emotional responsiveness as the difference between both levels. Specifically, the positive (negative) valence difference indexes measure the variation in valence index recorded in the third stage when the stimulus was positive (negative) net of the individual's baseline resting state valence index recorded in the first stage. Both differences can be interpreted as a lower level of overreaction of participants—they become more phlegmatic or cold headed—or that individuals move toward a greater withdrawal behavior or attitude.

5.3 Attrition Analysis

The final sample, after filtering the EEG data and accounting for attrition, is 308 valid EEG recordings of students; that is, an average attrition share of 49%. In this subsection, we present some checks to verify that this attrition rate was not correlated with the intervention. Our argument is that attrition was caused mainly from the quality of the data recordings. For example, long, dense, or unclean hair, and/or freezing computers were the most common troubles for the Matlab toolbox in logging the EEG recordings.

As the first check, we compare the share of nonvalid EEG recordings between T and C groups and find there are no statistical differences among them. (See the "Missing Share" row at the end of Table A7). Therefore, their missing status is not correlated with the treatment.

How different are the attrited and nonattrited subsamples? We provide information on means of variables for each subgroup in Table A8 in the Appendix. Estimations indicate that both groups are similar in most of the baseline variables except for the number of boys and area of residence (both are statistically different at 5%) and a category of household composition and mother's education. However, after adjusting the tests for differences by the number of hypotheses we tested and FWER,

these differences are no longer statistically relevant.

Finally, we face 13 attrited observations from the psychometric tests (locus of control, Raven, and CRT) due to technical problems with the computers; they stopped working and did not record student responses. As we show in Table A9 in the Appendix, the average characteristics of these 13 observations are close to those of the other 295 observations with valid psychometric measures, and are similarly distributed among T and C groups.

As we will explain later, we run two robustness checks. First, we also estimate our main specification using Heckman's correction for selection bias and compare if the coefficients are similar after accounting for missing observations. Second, we estimate Lee's (2009) treatment effects bounds for nonrandom sample selection.

5.4 ASP Effects on Emotional Regulation

To measure the overall impact of the intervention on emotional regulation, we use a similar specification to (1), but now our main outcomes of interest are E_{ijt}, the emotional regulation and socioemotional measures of the student i in school-by-education-level block j, expressed in standard deviations from the outcomes distributions of students in the control group. The remaining variables are defined as before. θ_1 is now the ITT effect of being randomly assigned to participate in an ASP on emotional resilience outcome E_{ijt}, compared to the C group.

Table 5 shows the ASP's effects on measures of emotional regulation and socioemotional skills. We find that ASP participants face a reduction in their valence-at-resting-state outcome by 0.35 standard deviations compared to children in the C group, as we can see in column (2). In terms of emotional responsiveness, we find a reduction in participant overreactions to positive stimuli by 0.44 standard deviations, as we present in column (3). This result can be interpreted as participants moving toward a greater withdrawal behavior or attitude, relative to the C group.

This improvement in emotional regulation is also complemented by student self-reports on locus of control. We can see in column (5) that treated children report a reduction in their locus of control by 0.25 standard deviations compared to children in the C group. Thus, treated students are perceiving that they can manage or control what happens in their lives by a greater magnitude than nontreated children. In the rest of the outcomes, we find no statistical difference between treated and control students at the conventional levels due to power limitations, as we show in the MDE estimations row.

These results shed light on the conclusion that the ASP directly affects student reactions to some stimuli. The program is oriented to teach them how to manage their automatic responses that produce violent behaviors across the different domains where children perform. Moreover, the

results are consistent with recent evidence on how some interventions in CBT can positively impact behavior (Blattman et al., 2017; Heller et al., 2017).

Although valid EEG recordings were approximately 52% of the total subsample, we showed in the attrition subsection that this rate was not different across T and C groups. Therefore, assuming that the process governing recordings validity follows a standard monotonicity property, we can compare outcomes across both groups, which will be valid estimates of the impact of treatment on outcomes for the full randomly selected subsample. However, as a check, we computed Heckman selection-corrected regression results for the same subsample outcomes (Heckman, 1979). These results are presented in Table A10 in the Appendix. We find these estimated ITT effects using the 598 observations. After correcting for the selection bias—using a dummy that indicates if the student was selected or not selected to participate in the neurophysiological experiments as selection variable—we notice they are very similar to the results obtained from the 308 valid EEG recordings. We also estimated Lee's treatment effects bounds for nonrandom sample selection (Lee, 2009) and present these results in Table A11 in the Appendix. We find that zero belongs to the treatment effects interval for all variables except for valence, positive valence difference, and locus of control. Thus, attrition is unlikely to have biased the results reported above.

There is a lack of evidence of heterogeneous effects of ASPs on socioemotional skills and no evidence of differences in the impact on ER by participants' initial propensity for violence. Therefore, we contribute to the literature testing if ASPs affect high- and low-violence children's neurophysiological outcomes differently. In Table 6, we present the estimated effects using specification (2) for ER measures. Coefficients in row [i] show the ASP's effects on low-violence treated students compared to low-violence children in the C group. Coefficients in row [ii] show the differences in effects between highly violent treated students and similar children in the control group. Then, coefficients in row [iii] point to the difference in effects between high- and low-violence treated students. Finally, row [iv] indicates p-values for the test of difference in effects between high- and low-violence treated students.

When we compare low-violence students assigned to T or C groups (row [i]), we find treated children drive both the reductions in the valence-at-resting-state measure and in the perception that they can control their circumstances by 0.57 and 0.47 standard deviations, respectively.

Then, looking at the estimations on the ASP's differential effects on arousal, we unexpectedly find that treating highly violent children increases their stress levels compared to both highly violent nontreated peers and with low-violence treated students. What can be causing this unexpected result? Due to the novelty of the complete experimental design, evidence from Dinarte (2018) on group composition in terms of participants' initial propensity for violence helps to pro-

vide an explanation. The author identifies that treating vulnerable children—here, those with high violence—with similarly violent peers increases their stress levels and their probability of misbehaving at school, compared to treatment in a more violence-diverse group. Dinarte (2018) argues that exposure to risky environments usually increases individuals' stress levels, either because they have to avoid danger or learn how to face it—for instance, by defending themselves.

In summary, these results on emotional regulation reinforce our previous conclusions on the ASP's effects on misbehavior and attitudes. This low-intensity intervention improves noncognitive and cognitive outcomes of at-risk youth, and the effects are different according to individuals' initial propensity for violence. For highly violent children, the impacts are mostly on absenteeism and academic performance. The intervention helps low-violence participants to regulate their emotions, which contributes to better attitudes and reduction of violent behaviors. Similar to our previous results, it is harder to change emotions in those who are more prone to violence. As a result, more intensive interventions may be more effective for these individuals.

Finally, an important contribution through these results is that the distribution of participants with unlike propensity for violence is relevant for this ASP and can increase its effectiveness. Specifically, segregation of participants might cause unintended neurophysiological effects on the highly violent children who are supposed to benefit most from these programs.

6. Learning versus Protection Mechanisms

At least two mechanisms through this ASP may have changed behavioral and academic outcomes. First, students may have learned social skills and conflict management directly from the club curricula, through their interaction with other children, or both. We call this the *learning mechanism*. Second, children may have reduced their violent behaviors because the ASP kept them busy and off the streets when they might otherwise be left alone and exposed to external risks (Gottfredson et al., 2004; Jacob and Lefgren, 2003; Newman et al., 2000). This will be the *protection mechanism* or, as previous papers have called it, the "involuntary incapacitation" channel (Heller et al., 2017). Although this experimental design does not allow us to perfectly disentangle the two mechanisms, in this section we provide suggestive evidence that students are indeed learning social skills. Therefore, the first mechanism is more likely driving the effects.

First, we exploit the availability of baseline data on adult supervision after school hours to test for differences between both mechanisms. Our assumption is that treated students who reported being without adult supervision after school receive both effects from the intervention, and that the effects for students who are with an adult after school may be caused only by the learning

mechanism. To estimate these differential effects, we included in specification (1) an interaction between the treatment variable T_{ij} and a dummy of being alone after school.

We present these estimations in Table 7. Row [i] presents the learning mechanism effects alone, row [ii] includes both effects, and row [iii] shows the protection effect alone. Estimated coefficients indicate that most of the effects are mainly related to the learning mechanism, on both cognitive and noncognitive outcomes. An interesting result drawn from row [iii] is that only protecting children may have an unintended effect compared to teaching them life skills. As we can see in columns (6) and (7), the net effect of protection alone increases the violence index and approval of antisocial peer behavior. These results illustrate that the main mechanism of the intervention can be social skills learning.[22]

As an additional attempt to study the protection mechanism, we use students' self-reports of exposure to crimes, as either victims or witnesses, and their awareness of risk within their communities or at home.[23] The assumption here is that if the protection channel is operating, they may perceive changes in their vulnerability to risky environments. We do not find statistically significant effects on most of those outcomes, except an increase in students' awareness of risk in their communities, which can also be interpreted as a skill developed through the learning channel. These results are available upon request.

7. Conclusions

This paper provides the first experimental evaluation of the direct and indirect impact of an ASP implemented in a developing and highly violent country. By experimentally granting the participation of 1,056 at-risk students in the intervention, our analysis examines whether the program generates direct impacts on participants' academic, behavioral, and violence outcomes. We also present some evidence of spillover effects from the intervention on nonenrolled children. Finally, a remarkable contribution from our paper is to provide evidence of how this program modifies participants' emotional regulation and stress.

The first novel result is that this low-intensive ASP is effective in the context of a developing and highly violent country. Moreover, the magnitude of the effects is between those found by Durlak et al. (2010); Cook et al. (2015) for developed countries, and those found by Heller et al. (2017)

---

[22]We also find that effects are greater when we estimate them using only the sample of students who participated in at least one session. These results are exhibited in Table A12 in the Appendix and shed light on how effective participation strengthens the impact from both mechanisms.

[23]These last estimations are only an approximation. We should be cautious in interpreting them because the question regards crimes witnessed or experienced after school hours, a period usually from 12:30 p.m. to 2:00 p.m. However, most crimes in El Salvador occur after 5:00 p.m.

32

from a CBT intervention implemented in the United States.

Our results also find a positive effect on both the intensive and extensive margin of student academic performance. This piece of evidence is relevant in the context of educational systems in developing countries, which are characterized by lack of resources and poor quality. Thus, these interventions can act as remedial programs to enhance student educational attainment through the acquisition of noncognitive skills.

An additional result is that all participants benefit from the program, but those gains are particular to specific outcome categories. For example, students with a lower propensity for violence drive the effects on misbehavior and violence, while children with a greater propensity for violence are more likely to increase academic achievement and reduce school absenteeism, compared to both the less violent and C groups.

The most notable contribution of this paper is that the program likely affects mental states and self-reported test scores related to socioemotional skills. First, we found a significant impact on emotional regulation and socioemotional skills, which is part of our main hypothesis of how the intervention affects behavior and academic performance.

However, neurophysiological results also indicate that the most vulnerable treated students experience an increase on our objective measure of stress, compared to high-violence nontreated children. The complete experimental design presented in Dinarte (2018) allows us to conclude that the manner of treatment is likely causing these unintended effects. Thus, with this evidence we also contribute to the optimal design and implementation of ASPs.

The methodology proposed in this paper to collect data on emotions has many benefits. First, it offers a way to incorporate emotion into the fields of education and violence economics. The importance of emotional regulation to life satisfaction has recently been highlighted for both developed and developing countries (Deming, 2017; OECD, 2015). This study shows there are neurophysiological approaches to proxy emotional disposition and responsiveness with a high level of accuracy and at relatively low cost in violent communities. The results may also aid evaluation of similar programs oriented to improving noncognitive skills.

These results have implications for public policy discussions of interventions oriented to improve academic outcomes and reduce violence within schools. First, participation in an ASP, where students learn about life skills and conflict management, benefits academic and noncognitive outcomes for both the least and most vulnerable students. Additionally, increasing adult supervision of students for some hours during the week reduces their exposure to risk and, particularly for boys at this age, may reduce the probability of gang recruitment (Cruz, 2007).

Since the intervention keeps students away from potential risk contexts and under supervision

while they also learn life skills, positive effects can stem either from the acquisition of these skills or from decreased involvement with bad peers outside school. We provide suggestive evidence that the socioemotional skills-learning mechanism drives the results. However, further rigorous research on these two channels is still necessary and would have significant implications for the design of these programs.

Further research should also examine if these results will persist over time. Due to the high out-of-school risk these children face, students in the control group were allowed to participate in the intervention the following academic year. This makes it more difficult to measure the ASP's long-term effect.

Finally, in the literature of interventions aimed at reducing crime and violence, one important aspect of these programs concerns the development of new, healthier social ties that foster a sense of belonging for participants and positively influence identity (Heller et al., 2017). In this aspect, more evidence is needed on how to improve this intervention if students participate in the program within their closer network, exploiting their preferences for similar peers.

References

Baker, A. and M. L. Hoekstra (2010). Externalities in the classroom: How children exposed to domestic violence affect everyone's kids. *American Economic Journal: Applied Economics 2*(1), 211–228.

Baysan, C., M. Burke, F. González, S. Hsiang, and E. Miguel (2018). Economic and non-economic factors in violence: Evidence from organized crime, suicides and climate in mexico. Technical report, National Bureau of Economic Research.

Biggart, A., S. Sloan, and L. O?Hare (2014). A Longitudinal Follow Up Study of the Doodle Den After School Programme. Technical report, Childhood Development Initiative.

Blattman, C., J. C. Jamison, and M. Sheridan (2017, April). Reducing crime and violence: Experimental evidence from cognitive behavioral therapy in liberia. *American Economic Review 107*(4), 1165–1206.

Borghans, L., A. L. Duckworth, J. J. Heckman, and B. Ter Weel (2008). The economics and psychology of personality traits. *Journal of human Resources 43*(4), 972–1059.

Card, D. and G. B. Dahl (2011). Family violence and football: The effect of unexpected emotional cues on violent behavior*. *The Quarterly Journal of Economics 126*(1), 103–143.

Carrel, S. and M. L. Hoekstra (2010). Externalities in the classroom: How children exposed to domestic violence affect everyone's kids. *American Economic Journal: Applied Economics 2*(1), 211–228.

Carrell, S. E., B. I. Sacerdote, and J. E. West (2013). From natural variation to optimal policy? The importance of endogenous peer group formation. *Econometrica 81*(3), 855–882.

Cassel, R. N., P. Chow, D. F. Demoulin, and R. C. Reiger (2000). Extracurricular involvement in high school produces honesty and fair play needed to prevent delinquency and crime. *Education 121*(2), 247.

Chandler, D., S. D. Levitt, and J. A. List (2011). Predicting and preventing shootings among at-risk youth. *The American Economic Review 101*(3), 288–292.

Chetty, R., N. Hendren, and L. F. Katz (2016). The effects of exposure to better neighborhoods on children: New evidence from the moving to opportunity experiment. *The American Economic Review 106*(4), 855–902.

Cook, P. J., K. Dodge, G. Farkas, R. G. Fryer Jr, J. Guryan, J. Ludwig, S. Mayer, H. Pollack, and L. Steinberg (2015). Not Too Late: Improving Academic Outcomes for Disadvantaged Youth. *Institute for Policy Research Northwestern University.*.

Cook, P. J., D. C. Gottfredson, and C. Na (2010). School crime control and prevention. *Crime and Justice 39*(1), 313–440.

Cook, P. J., J. Ludwig, and J. McCrary (Eds.) (2011). *Controlling Crime*. University of Chicago Press.

Cruz, J. M. (2007). *Street Gangs in Central America*. UCA Editores.

CSJ (2014). Estadisticas de la Corte Suprema de Justicia de El Salvador.

Cunha, F. and J. J. Heckman (2008). Formulating, identifying and estimating the technology of cognitive and non-cognitive skill formation. *Journal of Human Resources 43*(4), 738–782.

Damasio, A. R. (1994). Descartes' error: Emotion, reason, and the human brain.

Damm, A. P. and C. Dustmann (2014). Does growing up in a high crime neighborhood affect youth criminal behavior? *The American Economic Review 104*(6), 1806–1832.

Davidson, R., C. Schaffer, and C. Saron (1983, July). Frontal and Parietal Electroencephalogram Asymmetry in Depressed and Nondepressed Subjects. *Biol Psychiatry 18*(7), 753–762.

Davidson, R. J., P. Ekman, C. D. Saron, J. A. Senulis, and W. V. Friesen (1990). Approach-withdrawal and cerebral asymmetry: Emotional expression and brain physiology: I. *Journal of personality and social psychology 58*(2), 330.

DellaVigna, S. (2009). Psychology and economics: Evidence from the field. *Journal of Economic literature 47*(2), 315–72.

Deming, D. J. (2017). The growing importance of social skills in the labor market. *The Quarterly Journal of Economics 132*(4), 1593–1640.

Dinarte, L. (2018). Peer effects in after-school programs. experimental evidence in el salvador.

Dodge, K. A., J. E. Bates, and G. S. Pettit (1990). Mechanisms in tie cycle of violence. *Science 250*, 1678–1683.

Durlak, J. A., R. P. Weissberg, A. B. Dymnicki, R. D. Taylor, and K. B. Schellinger (2011). The impact of enhancing students? social and emotional learning: A meta-analysis of school-based universal interventions. *Child development 82*(1), 405–432.

Durlak, J. A., R. P. Weissberg, and M. Pachan (2010). A meta-analysis of after-school programs that seek to promote personal and social skills in children and adolescents. *American journal of community psychology 45*(3-4), 294–309.

Eccles, J. S. and J. Templeton (2002). Chapter 4: Extracurricular and other after-school activities for youth. *Review of research in education 26*(1), 113–180.

Egana-delSol, P. (2016a). Affective neuroscience meets labor economics: Assessing non-cognitive skills on late stage investment on at-risk youth. In *Skills For Sustainable Development: Essays On How Creativity, Entrepreneurship And Emotions Foster Human Development*, Chapter 1. Columbia University Academic Commons.

Egana-delSol, P. (2016b). Can we trust self-reporting as a measurement of non-cognitive skills? In *Skills For Sustainable Development: Essays On How Creativity, Entrepreneurship And Emotions Foster Human Development*, Chapter 2. Columbia University Academic Commons.

Fudenberg, D. and D. K. Levine (2006). A dual-self model of impulse control. *American economic review 96*(5), 1449–1476.

FUNDAUNGO (2013). Atlas de la violencia en el salvador (2009-2012). *San Salvador*.

FUSADES (2015). El Salvador Youth Survey Database.

Fuster, J. M. (2013). *The neuroscience of freedom and creativity: Our predictive brain*. Cambridge University Press.

Glasswing International (2012a). Estudio Interno sobre impacto de programa extracurricular de seis centros escolares. Technical report, Fundacin Crislida. Glasswing International.

Glasswing International (2012b). Mi Escuela, Mi espacio. un programa educativo extracurricular para el desarrollo integral de la niez, la adolescencia y la juventud de El Salvador. Technical report, Fundacin Crislida. Glasswing International.

Goldschmidt, P., D. Huang, and M. Chinen (2007). The long-term effects of after-school programming on educational adjustment and juvenile crime: A study of the LA's BEST After-School Program. *Los Angeles: UCLA/CRESST. Retrieved September 8*, 2008.

Goleman, D. (2010). Emotional intelligence: Why it can matter more than iq. *Bantam Books*.

Gottfredson, D. C., A. Cross, and D. A. Soulé (2007). Distinguishing characteristics of effective and ineffective after-school programs to prevent delinquency and victimization. *Criminology & Public Policy 6*(2), 289–318.

Gottfredson, D. C., S. A. Gerstenblith, D. A. Soulé, S. C. Womer, and S. Lu (2004). Do After-School Programs reduce delinquency? *Prevention Science 5*(4), 253–266.

Harmon-Jones, E., P. A. Gable, and C. K. Peterson (2010). The role of asymmetric frontal cortical activity in emotion-related phenomena: A review and update. *Biological psychology 84*(3), 451–462.

Haushofer, J. and E. Fehr (2014). On the psychology of poverty. *Science 344*(6186), 862–867.

Heckman, J. J. (1979). Sample selection bias as a specification error. *Econometrica 47*(1), 153–161.

Heckman, J. J. and T. Kautz (2012). Hard evidence on soft skills. *Labour economics 19*(4), 451–464.

Heckman, J. J., J. Stixrud, and S. Urzua (2006). The effects of cognitive and non-cognitive abilities on labor market outcomes and social behavior. *Journal of Labor economics 24*(3), 411–482.

Heller, S. B., A. K. Shah, J. Guryan, J. Ludwig, S. Mullainathan, and H. A. Pollack (2017). Thinking, fast and slow? Some field experiments to reduce crime and dropout in Chicago. *The Quarterly Journal of Economics 132*(1), 1–54.

Herrenkohl, T. I., C. Sousa, E. A. Tajima, R. C. Herrenkohl, and C. A. Moylan (2008). Intersection of child abuse and children's exposure to domestic violence. *Trauma, Violence, & Abuse 9*(2), 84–99.

Hirsch, B. J., L. V. Hedges, J. Stawicki, and M. A. Mekinda (2011). After-School Programs for high school students: An evaluation of After School Matters. *Evanston, IL: Northwestern University*.

Hughes, D., M. Yates, E. Morton, and L. Smillie (2014, December). Asymmetric frontal cortical activity predicts effort expenditure for reward. *Soc Cogn Affect Neurosci 10*(7).

Jacob, B. A. and L. Lefgren (2003). Are idle hands the devil's workshop? Incapacitation, concentration, and juvenile crime. *The American Economic Review 93*(5), 1560–1577.

Jaitman, L., D. Caprirolo, R. Granguillhome Ochoa, P. Keefer, T. Leggett, J. A. Lewis, J. A. Mejía-Guerra, M. Mello, H. Sutton, and I. Torre (2017). The costs of crime and violence: New evidence and insights in latin america and the caribbean. *Inter-American Development Bank*.

John, O. and S. Srivastava (1999). The Big-Five Trait Taxonomy: History, Measurement, and Theoretical Perspectives. In L. Pervin and O. John (Eds.), *Handbook of Personality: Theory and Research*, pp. 102–138. New York: Guilford Press.

Jones, D., D. Molitor, and J. Reif (2018). What do workplace wellness programs do? evidence from the illinois workplace wellness study. Technical report, National Bureau of Economic Research.

Kahneman, D. (2011). *Thinking, fast and slow*. Macmillan. New York and London.

Kassam, K. S., A. R. Markey, V. L. Cherkassky, G. Loewenstein, and M. A. Just (2013, June). Identifying Emotions on the Basis of Neural Activation. *PLoS One 8*(6), e66032.

Korjus, K., A. Uusberg, H. Uusberg, N. Kuldkepp, K. Kreegipuu, J. Allik, R. Vicente, and J. Aru (2015, February). Personality Cannot Be Predicted from the Power of Resting State EEG. *Front Hum Neurosci 9*(63), 1–7.

Krueger, A. B. (2003). Economic considerations and class size. *The Economic Journal 113*(485), F34–F63.

Krug, E. G., J. A. Mercy, L. L. Dahlberg, and A. B. Zwi (2002). The World Report on Violence and Health. *The Lancet 360*(9339), 1083–1088.

Lazear, E. P. (2001). Educational Production. *The Quarterly Journal of Economics 116*(3), 777–803.

Lee, D. S. (2009). Training, wages, and sample selection: Estimating sharp bounds on treatment effects. *The Review of Economic Studies 76*(3), 1071–1102.

Lewis, D. O., S. S. Shanok, J. H. Pincus, and G. H. Glaser (1979). Violent juvenile delinquents: Psychiatric, neurological, psychological, and abuse factors. *Journal of the American Academy of Child Psychiatry 18*(2), 307–319.

Loewenstein, G. (2000). Emotions in economic theory and economic behavior. *The American Economic Review 90*(2).

LPG (2016). Amenazas de pandillas es la quinta causa de desercin escolar.

Mahoney, J. L., M. E. Parente, and E. F. Zigler (2010). After-School Program participation and children's development. *Handbook of Research on Schools, Schooling, and Human Development*, 379–397.

Mahoney, J. L., H. Stattin, and D. Magnusson (2001). Youth recreation centre participation and criminal offending: A 20-year longitudinal study of swedish boys. *International journal of behavioral development 25*(6), 509–520.

Marshall, N. L., C. G. Coll, F. Marx, K. McCartney, N. Keefe, and J. Ruh (1997). After-school time and children's behavioral adjustment. *Merrill-Palmer Quarterly (1982-) 20*(3), 497–514.

Marshall, P., B. Reeb, Fox, C. Nelson, and C. Zeanah (2008). Effects of early intervention on eeg power and coherence in previously institutionalized children in romania. *Dev Psychopathol*, 861–880.

Martinez-Leon, J.-A., J.-M. Cano-Izquierdo, and J. Ibarrola (2016). Are low cost brain computer interface headsets ready for motor imagery applications? *Expert Systems with Applications 49*(1), 136–144.

MINED (2015). Estadisticas del Ministerio de Educacin de El Salvador.

Moffitt, T. E., L. Arseneault, D. Belsky, N. Dickson, R. J. Hancox, H. Harrington, R. Houts, R. Poulton, B. W. Roberts, S. Ross, et al. (2011). A gradient of childhood self-control predicts health, wealth, and public safety. *Proceedings of the National Academy of Sciences 108*(7), 2693–2698.

Moya, A. (2018). Violence, psychological trauma, and risk attitudes: Evidence from victims of violence in colombia. *Journal of Development Economics 131*, 15–27.

Newman, S. A., J. A. Fox, E. A. Flynn, and W. Christeson (2000). America's After-School Choice: The Prime Time for Juvenile Crime, or Youth Enrichment and Achievement. Technical report, Fight Crime: Invest in Kids, Washington, DC.

OECD (2015). Skills for social progress: The power of social and emotional skills. Technical report, OECD Center for Research and Innovation (CERI).

PNC (2014). Estadisticas de la Policia Nacional Civil de El Salvador.

PNUD (2013). Informe Regional de Desarrollo Humano 2013-2014. Seguridad ciudadana con rostro humano: diagnóstico y propuestas para América Latina. Technical report, Programa de Naciones Unidas para el Desarrollo.

Ramirez, R. and Z. Vamvakousis (2012). Detecting emotion from eeg signals using the emotive epoc device. *Brain Informatics*, 175–184.

Rivera, J. (2013). Las maras. el fenmeno criminal del siglo xxi.

Rotter, J. B. (1966). Generalized expectancies for internal versus external control of reinforcement. *Psychological monographs: General and applied 80*(1), 1.

Salzman, C. D. and S. Fusi (2010). Emotion, Cognition, and Mental State Representation in Amygdala and Prefrontal Cortex. *The Annual Review of Neuroscience 33*, 173–202.

Scott-Little, C., M. S. Hamann, and S. G. Jurs (2002). Evaluations of After-School Programs: A meta-evaluation of methodologies and narrative synthesis of findings. *American Journal of Evaluation 23*(4), 387–419.

Shulruf, B. (2010). Do extra-curricular activities in schools improve educational outcomes? A critical review and meta-analysis of the literature. *International Review of Education 56*(5-6), 591–612.

Soares, R. R. and J. Naritomi (2010). Understanding high crime rates in Latin America: The role of social and policy factors. In *The Economics of Crime: Lessons for and from Latin America*, pp. 19–55. University of Chicago Press.

Sousa, C., T. I. Herrenkohl, C. A. Moylan, E. A. Tajima, J. B. Klika, R. C. Herrenkohl, and M. J. Russo (2011). Longitudinal study on the effects of child abuse and children?s exposure to domestic violence, parent-child attachments, and antisocial behavior in adolescence. *Journal of interpersonal violence 26*(1), 111–136.

Taheri, S. A. and B. C. Welsh (2016). After-school programs for delinquency prevention: A systematic review and meta-analysis. *Youth violence and juvenile justice 14*(3), 272–290.

UNICEF (2017). A familiar face: Violence in the lives of children and adolescents.

Verma, G. and U. S. Tiwary (2014). Multimodal Fusion Framework: A Multiresolution Approach for Emotion Classification and Recognition from Physiological Signals. *Neuroimage 102*, 162–172.

Webb, V., L. Nuno, and C. Katz (2016). Influence of risk and protective factors on school aged youth involvement with gangs, guns, and delinquency: Findings from the El Salvador Youth Survey. Technical report, Center for Violence Prevention and Community Safety. Arizona State University.

WHO (2011). World Health Organization Statistics.

WHO (2016). *World Health Statistics 2016: Monitoring Health for the SDGs Sustainable Development Goals.* World Health Organization.

TABLE 1. SUMMARY STATISTICS: MEANS OF VARIABLES BY GROUP PRIOR TO TREATMENT

	(1) Full Sample	(2) Control Group (C)	(3) Treatment Group (T)	(4) Unadjusted p- values	(5) Adjusted p- values
PANEL A: IVV DETERMINANTS					
Student is male	0.49	0.51	0.49	0.659	0.990
Student's age	11.94	11.86	11.97	0.247	0.825
Student lives in urban area	0.73	0.72	0.74	0.502	0.995
Student's household composition					
Student lives with both parents	0.53	0.47	0.55	0.073	0.465
Student lives only with one parent	0.32	0.36	0.31	0.125	0.745
Student lives with one parent and stepparent	0.06	0.07	0.06	0.658	0.995
Student lives with other relatives/adults	0.09	0.10	0.08	0.551	0.990
Student's mother's level of education:					
Basic education (1–6 years)	0.31	0.34	0.30	0.283	0.935
Intermediate education (7–12 years)	0.62	0.59	0.63	0.379	0.985
University or higher (13 years +)	0.07	0.07	0.07	0.775	0.995
Student's travel time from house to school (min.)	17.64	16.98	17.85	0.459	0.985
Student is alone at home after school	0.05	0.05	0.05	0.805	0.995
Student's school year	5.75	5.67	5.78	0.072	0.450
Student is enrolled in the morning shift	0.71	0.71	0.72	0.667	0.995
Student's violence index	0.04	0.04	0.04	0.790	0.995
PANEL B: ACADEMIC OUTCOMES					
Academic scores Q1 2016 (Baseline)					
Reading scores	6.66	6.48	6.73	0.034	0.110
Math scores	6.50	6.44	6.51	0.628	0.565
Science scores	6.61	6.47	6.66	0.222	0.185
Behavior scores	7.39	7.28	7.43	0.085	0.185
Absenteeism Q1 2016	1.69	2.21	1.52	0.010	0.110
PANEL C: CLUBS' CHARACTERISTICS					
Average club size	-	-	13.4	-	-
Average club take-up	-	-	0.57	-	-
Community tutors	-	-	0.31	-	-
Club category					
Leadership	-	-	0.29	-	-
Art and Culture	-	-	0.16	-	-
Sports	-	-	0.26	-	-
Science	-	-	0.29	-	-
Share of treated by course	-	-	0.42	-	-
Observations	1056	258	798		

Table 1 shows descriptive statistics of the available variables at baseline for the sample of enrolled children. Panel A summarizes information obtained from the enrollment form that was used as determinants in the IVV estimation. Panel B presents administrative data provided by schools. This data is from the first quarter of academic year 2016, before the clubs were implemented. We imputed average course grade for missing observations. For the absenteeism data, we have no records for some courses; thus we imputed the average school absenteeism. The scale of grades in El Salvador is 0–10 points. Panel C presents club characteristics. Take-up is estimated as number of hours attended by student i/total hours in each club. p-values in column (5) have been adjusted for multiple hypothesis testing of means and FWER using a command developed by Jones et al. (2018), which follows the free step-down resampling methodology of Westfall and Young (1993). In the estimations of adjusted p-values, we separated the variables into the two main outcomes families (Panel A and Panel B).

TABLE 2. OVERALL EFFECTS OF THE ASP

PANEL A: VIOLENCE AND ATTITUDES

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
	Attitudes toward school and learning				Violence and Behavior				
	Positive attitudes toward school	Time to do homework (hours)	Pay attention in class	Absenteeism (days)	Delinquency (Index)	Violent actions (Index)	Approval of antisocial behavior	Behavior reports (-)	Probability of misbehavior report (-)
Any treatment	0.172**	0.336**	0.079**	-1.593***	-0.198*	-0.143*	-0.104***	-0.172**	-0.064**
	(0.062)	(0.098)	(0.022)	(0.272)	(0.067)	(0.045)	(0.024)	(0.052)	(0.021)
Observations	948	935	962	836	916	956	956	1010	1010
Mean control group	-0.13	2.12	0.59	7.16	0.00	0.00	0.17	7.09	0.67
SD - control group	1.49	1.89	0.49	9.20	0.97	0.97	0.38	1.29	0.47

PANEL B: ACADEMIC OUTCOMES

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
	Grades				Probability of passing				Passing course
	Reading	Math	Science	Score	Reading	Math	Science	Score	
Any treatment	0.015	0.108**	0.131*	0.057	0.037**	0.019	0.029**	0.026	0.028**
	(0.039)	(0.037)	(0.047)	(0.038)	(0.010)	(0.017)	(0.014)	(0.015)	(0.014)
Observations	1023	1023	1023	1023	1023	1023	1023	1023	1023
Mean control group	6.47	6.23	6.37	6.37	0.87	0.87	0.88	0.87	0.90
SD - control group	1.75	1.76	1.66	1.63	0.34	0.33	0.32	0.33	0.29

***, **, * significant at 1%, 5%, and 10%, respectively. Bootstrapped standard errors at the course-school level are in parentheses. Panel A is effects on noncognitive outcomes. Positive attitudes, time spent on homework, paying attention in class, criminal and violent action indexes, and approval of antisocial behavior were estimated using collected self-reported data at follow-up. Absenteeism is the number of days student missed school between April–October of the 2016 academic year. It was obtained from school administrative data. Bad behavior reports are administrative school reports. They were standardized using the control group at the school-grade level. Panel B presents results on academic outcomes. Reading, math, and science grades are standardized values from control groups at the school-grade level at follow-up. Score is an average of the three courses. See Appendix 1 for a detailed description of the outcome variables. All regressions include as controls: a second-order polynomial of student's IVV, and school-by-education-level block fixed effect (stratification level). Additionally, in estimations for academic outcomes, absenteeism, and bad behavior reports, we also include the corresponding imputed outcome at the baseline and a dummy indicating a missing value at baseline. Differences in number of noncognitive outcome observations are due to variation in the response rate for each outcome. Coefficients remain statistically significant after adjusting p-values for multiple hypothesis testing of means and FWER following Jones et al. (2018). For this adjustment, the families of outcomes are: (i) attitudes toward school and learning, (ii) violence and misbehavior, (iii) academic performance (intensive margin), and (iv) academic performance (extensive margin).

42

TABLE 3. HETEROGENEOUS TREATMENT EFFECTS BY IVV

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)

PANEL A: VIOLENCE AND ATTITUDES

	Attitudes toward school and learning				Violence and Behavior				
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
	Positive attitudes toward school	Time to do homework (hours)	Pay attention in class	Absenteeism (days)	Delinquency (Index)	Violent actions (Index)	Approval of antisocial behavior	Behavior reports (-)	Probability of bad behavior report
[i] Effect on low-violence students	0.097	0.428***	0.078**	-0.851	-0.229*	-0.095	-0.070**	-0.282***	-0.084***
	(0.095)	(0.153)	(0.039)	(0.678)	(0.133)	(0.077)	(0.029)	(0.081)	(0.031)
[ii] Effect on high-violence students	0.255*	0.244	0.080	-2.329**	-0.167	-0.194	-0.143***	-0.061	-0.043
	(0.141)	(0.210)	(0.052)	(0.964)	(0.109)	(0.129)	(0.042)	(0.082)	(0.040)
[iii] Difference between high and low violence	0.158	-0.184	0.002	-1.478	0.062	-0.099	-0.073	0.221**	0.041
	(0.187)	(0.225)	(0.064)	(1.050)	(0.176)	(0.139)	(0.045)	(0.098)	(0.046)
[iv] p-value: high-violence effect = low-violence effect	0.397	0.412	0.980	0.159	0.727	0.476	0.105	0.025	0.375
Observations	948	935	962	836	916	956	956	1010	1010

PANEL B: ACADEMIC OUTCOMES

	Grades				Probability of passing				Passing course
	Reading	Math	Science	Score	Reading	Math	Science	Score	
[i] Effect on low-violence students	-0.039	-0.009	0.031	-0.036	0.033**	-0.007	0.015	0.027	-0.029*
	(0.053)	(0.055)	(0.058)	(0.050)	(0.015)	(0.023)	(0.020)	(0.023)	(0.015)
[ii] Effect on high-violence students	0.072	0.228***	0.234***	0.153***	0.040	0.044	0.045	0.026	0.079**
	(0.078)	(0.069)	(0.080)	(0.067)**	(0.029)	(0.030)	(0.030)	(0.029)	(0.036)
[iii] Difference between high and low violence	0.111	0.237***	0.203***	0.189***	0.008	0.050	0.030	-0.001	0.098***
	(0.073)	(0.071)	(0.075)	(0.064)	(0.026)	(0.032)	(0.038)	(0.030)	(0.031)
[iv] p-value: high-violence effect = low-violence effect	0.126	0.001	0.007	0.003	0.774	0.116	0.429	0.972	0.001
Observations	1023	1023	1023	1023	1023	1023	1023	1023	1056

TABLE 4. ASP SPILLOVERS. EFFECTS ON NONENROLLED STUDENTS.

	(1)	(2)	(3)	(4)	(5)
	Grades				Behavior reports (-)
	Reading	Math	Science	Score	
[i] Percentage of club participants within student's n classroom (γ_1)	0.007** (0.003)	0.007*** (0.003)	0.005** (0.003)	0.007*** (0.002)	-0.011** (0.005)
[ii] Spillover effect of adding 3/22 treated students (1 sd)	0.094	0.094	0.080	0.094	-0.147
Observations	1357	1358	1357	1356	1194
Mean of nonenrolled	6.82	6.58	6.77	6.73	7.33
sd of nonenrolled	1.68	1.67	1.70	1.55	1.23

***, **, * significant at 1%, 5%, and 10%, respectively. Standard errors—clustered at course-by-school level—are in parentheses. Outcome variables are standardized grades at course-by-school level at follow-up. Percentage of club participants was estimated as $Sh_n = (Treated_n \div Enrolled_n) \times 100$. All regressions include as main control the percentage of enrolled/total students from each course. Individual controls include imputed grades in the course at baseline and a dummy indicating a missing value in the grade at baseline. Average percentage of enrolled/total children per classroom is 13.2% and average percentage of treated/enrolled children per course is 26.5%. Row [i] indicates the coefficients of specification (3) γ_1. Row [ii] indicates the average effect of adding 3 treated students in a classroom of 22 students who enrolled in the ASP (a standard deviation of treated students share×100 = 13.4%) on nonenrolled academic grades and bad behavior reports. Description of outcome variables is available in Appendix 1.

TABLE 5. ASP EFFECTS ON EMOTIONAL REGULATION

	(1) Arousal (stress)	(2) Valence	(3) Positive Valence Difference	(4) Negative Valence Difference	(5) Locus of Control	(6) CRT	(7) Raven
Treatment	0.069	-0.352**	-0.437**	-0.244	-0.249**	-0.012	-0.035
	(0.121)	(0.159)	(0.195)	(0.187)	(0.110)	(0.127)	(0.121)
Observations	308	308	308	308	295	295	295
Mean control group	0.00	0.00	0.00	0.00	0.00	0.00	0.00
SD - control group	1.00	1.00	1.42	1.28	1.00	1.00	1.00
MDE T = C	0.21	0.21	0.21	0.21	0.21	0.21	0.21

***, **, * significant at 1%, 5%, and 10%, respectively. Bootstrapped standard errors at the course-school level are in parentheses. All outcomes have been standardized at the control-course level with a mean of 0 and standard deviation 1.0. All regressions include as controls: a second-order polynomial of student's propensity for violence and stratification-blocks fixed effects. Differences in number of observations are due to variation in the response rate for each outcome.

TABLE 6. HETEROGENEOUS EFFECTS OF THE ASP ON EMOTIONAL REGULATION BY PARTICIPANTS' INITIAL PROPENSITY FOR VIOLENCE

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	Arousal (stress)	Valence	Positive Valence Difference	Negative Valence Difference	Locus of Control	CRT	Raven
i. Effect on low-violence treated students	-0.202	-0.566**	-0.569*	-0.487**	-0.466***	-0.019	0.041
	(0.176)	(0.232)	(0.345)	(0.207)	(0.123)	(0.185)	(0.168)
ii. Effect on high-violence treated students	0.340*	-0.137	-0.312	0.001	-0.039	-0.001	-0.110
	(0.179)	(0.197)	(0.257)	(0.316)	(0.200)	(0.263)	(0.184)
iii. Difference of effects between high-low treated students [ii] - [i]	0.542**	0.429	0.257	0.488	0.426*	0.019	-0.151
	(0.219)	(0.332)	(0.425)	(0.356)	(0.249)	(0.289)	(0.235)
iv. p-value Ho: Difference = 0	0.013	0.196	0.546	0.170	0.087	0.948	0.522
Observations	308	308	308	308	295	295	295

***, **, * significant at 1%, 5%, and 10%, respectively. Bootstrapped standard errors at the course-school level are in parentheses. All outcomes have been standardized at the control-course level with a mean of 0 and standard deviation 1.0. All regressions include as controls: a second-order polynomial of student's propensity for violence and stratification-blocks fixed effects. Differences in number of observations are due to variation in the response rate for each outcome.

TABLE 7. LEARNING AND PROTECTION MECHANISMS

PANEL A: VIOLENCE AND ATTITUDES

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
	Attitudes toward school and learning					Violence and Behavior			
	Positive attitudes toward school	Time to do homework (hours)	Pay attention in class	Absenteeism (days)	Delinquency (Index)	Violent actions (Index)	Approval of antisocial behavior	Behavior reports (-)	Probability of bad behavior report
[i] Effect on "protected"	0.192***	0.323***	0.077***	-1.603***	-0.199***	-0.163***	-0.109***	-0.178***	-0.062***
	(0.060)	(0.108)	(0.023)	(0.291)	(0.075)	(0.047)	(0.026)	(0.057)	(0.023)
[ii] Effect on "unprotected"	-0.255	0.534	0.103	-1.190	-0.185	0.363***	0.010	-0.035	-0.082
	(0.380)	(0.604)	(0.109)	(1.110)	(0.403)	(0.134)	(0.025)	(0.202)	(0.099)
[iii] Net protection effect	-0.447	0.212	0.026	0.410	0.013	0.527***	0.120***	0.143	-0.020
	(0.375)	(0.646)	(0.112)	(1.204)	(0.436)	(0.141)	(0.034)	(0.222)	(0.106)
Observations	948	935	962	836	916	956	956	1,010	1,010

PANEL B: ACADEMIC OUTCOMES

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
	Grades				Probability of passing				Failing at least one course
	Reading	Math	Science	Score	Reading	Math	Science	Score	
[i] Effect on "protected"	0.011	0.089**	0.118**	0.045	0.034***	0.017	0.027**	0.028*	-0.027***
	(0.041)	(0.038)	(0.047)	(0.039)	(0.010)	(0.017)	(0.013)	(0.015)	(0.009)
[ii] Effect on "unprotected"	0.082	0.449*	0.376	0.308	0.089	0.039	0.077	-0.023	-0.049
	(0.223)	(0.272)	(0.341)	(0.280)	(0.098)	(0.092)	(0.102)	(0.101)	(0.081)
[iii] Net protection effect	0.072	0.360	0.258	0.264	0.055	0.022	0.050	-0.051	-0.023
	(0.233)	(0.279)	(0.348)	(0.286)	(0.101)	(0.091)	(0.101)	(0.102)	(0.081)
Observations	1,023	1,023	1,023	1,023	1,023	1,023	1,023	1,023	1,023

***, **, * significant at 1%, 5%, and 10%, respectively. Bootstrapped standard errors at the course-school level are in parentheses. Panel A presents effects on noncognitive outcomes. Panel B present results on academic outcomes. Row [i] indicates the effect on academic outcomes, or those without adult supervision after school hours (i.e., the learning effect). Row [ii] shows results of the ASP on "unprotected" students, or those without adult supervision after school hours (i.e., both learning and protection mechanisms). And row [iii] shows the net protection effect, i.e., the difference between unprotected and protected students. All regressions include as controls: a second-order polynomial of student's IVV, and ciclo-school fixed effect (stratification level). Additionally, in estimations for academic outcomes, absenteeism, and bad behavior reports, I also include the corresponding imputed outcome at the baseline and a dummy indicating a missing value at the baseline.

Appendix section

1. Appendix tables

**TABLE A1. TESTS FOR DIFFERENCES BETWEEN PARTICIPANT
AND NONPARTICIPANT SCHOOLS**

	(1)	(2)	(3)	(4)	(5)
Panel A. School characteristics					
	School is located in urban area	School is in a top ten most violent municip.	Enrollment (number of students)	Share of indigenous students	Additional school revenues
Participant school	0.125	0.000	130.41	-0.003	575.973
	(0.182)	(0.000)	(107.25)	(0.003)	(2,109.54)
Mean nonparticipant schools	0.245***	0.093	256.14***	0.041***	1,798.6***
	(0.000)	(0.000)	(0.104)	(0.000)	(2.054)
Panel B. School programs					
School has a	EITP Program	School kits program	*Vaso de leche* Program	Food Program	Psychologist
Participant school	-0.109	-0.184	0.191***	0.034	0.148
	(0.070)	(0.134)	(0.043)	(0.024)	(0.115)
Mean nonparticipant schools	0.149***	0.979***	0.572***	0.983***	0.035***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Panel C. School facilities or equipment					
School has access to	Computer	Water	Electricity	Sanitation	Internet
Participant school	22.462	-0.072	0.003	0.172	0.416**
	(13.949)	(0.183)	(0.002)	(0.199)	(0.205)
Mean nonparticipant schools	9.024***	0.774***	0.976***	0.031***	0.217***
	(0.014)	(0.000)	(0.000)	(0.000)	(0.000)
Observations	5,134	5,134	5,134	5,134	5,134

Data source: El Salvador Educational Census (2015). For these estimations, we restricted the sample to public schools only and estimated the following specification: $y_{ij} = \alpha_0 + \alpha_1 P_{ij} + F_j + \epsilon_{ij}$, where y_{ij} is the characteristic of interest of school i in department j, geographic division; α_0 is the mean of nonparticipant schools; P_{ij} is an indicator for participant schools; and F_{ij} are department fixed effects. *Vaso de leche* corresponds to a breakfast program and EITP is an acronym for *Escuela Inclusiva a Tiempo Pleno*, which is the full-time school program. ***, **, * indicates that coefficients are significant at 1%, 5%, and 10%, respectively. Robust standard errors at course-school level are in parentheses.

**TABLE A2. MATCHING RATE WITH ADMINISTRATIVE DATA
AND ATTRITION RATE.**

	(1) Full Sample	(2) Control Group (C)	(3) Treatment Group (T)
Fraction of students with matched administrative data, Q1 2016 - Baseline			
Reading scores	0.98	0.97	0.98
Math scores	0.93	0.91	0.94
Science scores	0.95	0.94	0.96*
Behavior scores	0.95	0.93	0.95
Absenteeism	0.68	0.68	0.67
Fraction of students with matched administrative data, Q4 2016 - Follow-up			
Reading scores	0.97	0.98	0.97
Math scores	0.97	0.98	0.97
Science scores	0.97	0.98	0.97
Behavior scores	0.96	0.96	0.95
Absenteeism	0.80	0.79	0.80
Number of students at baseline and follow-up			
Number of students present at baseline	1056	258	798
Number of students present at follow-up	968	237	731
Retention rate (1-attrition)	0.92	0.92	0.91

The table provides the match rate with administrative data, calculated as the fraction of students present at the survey during the baseline who could be matched with administrative data from schools. Attrition rate refers to the share of children for whom we have baseline self-reported data but who were not available during follow-up data collection (survey). In comparing T and C, * denotes difference significant at the 10% level.

TABLE A3. SUMMARY STATISTICS: MEANS OF VARIABLES BY GROUP
PRIOR TO TREATMENT
Only subsample with nonmissing values at baseline.

	(1) Full Sample	(2) Control Group (C)	(3) Treatment Group (T)	(4) Unadjusted *p*-values	(5) Adjusted *p*-values	(6) Observations
ACADEMIC OUTCOMES						
Academic scores Q1 2016 (Baseline)						
Reading scores	6.70	6.46	6.73	0.051	0.140	1034
Math scores	6.48	6.42	6.50	0.591	0.580	984
Science scores	6.62	6.46	6.67	0.097	0.205	1007
Behavior scores	7.38	7.25	7.43	0.098	0.205	1000
Absenteeism Q1 2016	2.10	2.85	1.85	0.041	0.140	713

Table A3 shows descriptive statistics of the available variables from administrative datasets provided by schools at baseline for the sample of enrolled children using sample with nonmissing information. This data corresponds to the first quarter of academic year 2016, before the clubs were implemented. For the absenteeism data, we have no records for some courses. The scale of grades in El Salvador is 0–10 points. *p*-values in column (5) have been adjusted for multiple hypothesis testing of means and FWER using a command developed by Jones et al. (2018), which follows the free step-down resampling methodology of Westfall and Young (1993). All these variables constitute a family for the FWER adjustment. In tests for means at baseline, we control by stratification variable, and standard errors are clustered at the school-by-course level.

TABLE A4. BALANCE BETWEEN ENROLLED AND NONENROLLED STUDENTS

	(1)	(2)	(3)	(4)	(5)
	Grades				Behavior reports (-)
	Reading	Math	Science	Score	
Enrolled students	-0.106	-0.041	-0.051	-0.051	0.048
	(0.101)	(0.147)	(0.163)	(0.111)	(0.081)
Mean nonenrolled students	6.82	6.56	6.67	6.69	7.55
	(0.131)	(0.130)	(0.174)	(0.110)	(0.089)
Observations	2,415	2,415	2,415	2,415	2,334

The sample includes all students from the 5 participating public schools. The estimated specification was the following: $y_{ij} = \alpha_0 + \alpha_1 E_{ij} + F_j + \epsilon_{ij}$, where y_{ij} is the nonstandardized grades or misbehavior report of student i in the school course j at baseline, including imputed missing data; α_0 is the mean of nonenrolled children; E_{ij} is an indicator of student's decision to participate in the ASP at baseline, i.e., if they and their parents signed a consent form; and F_{ij} are school-by-education-level fixed effects. We also control with a missing data indicator. These data were obtained from school administrative records.

***, **, * indicate that the estimation is significant at 1%, 5%, and 10%, respectively. Clustered standard errors at the course-school level are in parentheses.

TABLE A5. OVERALL EFFECTS OF THE ASP CONTROLLING BY STUDENTS' BEHAVIOR AT SCHOOL

PANEL A: VIOLENCE AND ATTITUDES

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
	Attitudes toward school and learning				Violence and Behavior				
	Positive attitudes toward school	Time to do homework (hours)	Pay attention in class	Absenteeism (days)	Delinquency (Index)	Violent actions (Index)	Approval of antisocial behavior	Behavior reports (-)	Probability of Bad Behavior report
Any treatment	0.174***	0.343***	0.080***	-1.622***	-0.205***	-0.147***	-0.104***	-0.239***	-0.086***
	(0.067)	(0.101)	(0.023)	(0.250)	(0.065)	(0.047)	(0.024)	(0.043)	(0.018)
Observations	1010	935	962	836	916	956	962	1010	1010
Mean control group	-0.13	2.12	0.59	7.16	0.00	0.00	0.174	7.18	0.72
SD - control group	1.49	1.89	0.49	9.20	0.973	0.971	0.379	1.24	0.45
MDE T = C	0.108	0.109	0.108	0.173	0.108	0.108	0.131	0.135	0.123

PANEL B: ACADEMIC OUTCOMES

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
	Grades				Probability of passing				Passing course
	Reading	Math	Science	Score	Reading	Math	Science	Score	course
Any treatment	0.039	0.126***	0.156***	0.080**	0.041***	0.022	0.029**	0.031**	0.030***
	(0.040)	(0.037)	(0.047)	(0.040)	(0.009)	(0.015)	(0.013)	(0.014)	(0.008)
Observations	1010	1010	1010	1010	1010	1010	1010	1010	1010
Mean control group	6.47	6.23	6.37	6.37	0.865	0.873	0.884	0.873	0.904
SD - control group	1.75	1.76	1.66	1.63	0.342	0.334	0.319	0.334	0.294
MDE T = C	0.096	0.092	0.100	0.096	0.088	0.103	0.104	0.097	0.080

****, **, * significant at 1%, 5%, and 10%, respectively. Bootstrapped standard errors at the course-school level are in parentheses. Panel A is effects on noncognitive outcomes. Panel B presents results on academic outcomes. All regressions include as controls: a second-order polynomial of student's bad behavior at school using teacher reports before the intervention and stratification-blocks fixed effects. Additionally, in estimations for academic outcomes, absenteeism, and bad behavior reports, we also include the corresponding imputed outcome at the baseline and a dummy indicating a missing value at baseline. Differences in number of noncognitive outcome observations are caused by the differences in the response rate for each outcome.

TABLE A6. CORRELATION BETWEEN IVV, ACADEMIC GRADES, AND MISBEHAVIOR REPORTS AT BASELINE

	(1)	(2)	(3)	(4)	(5)
	GRADES				Behavior
	Reading	Math	Science	Score	

Panel A. Standardized and imputed grades

IVV	-0.013	0.021	-0.021	-0.011	0.056***
	(0.017)	(0.039)	(0.020)	(0.020)	(0.021)
Constant	0.176*	-0.048	0.179*	0.143	0.304***
	(0.096)	(0.150)	(0.104)	(0.087)	(0.104)
Observations	1,056	1,056	1,056	1,056	1,056

Panel B. Standardized grades at the course level

IVV	-0.015	-0.007	-0.021	-0.018	0.050**
	(0.019)	(0.028)	(0.018)	(0.021)	(0.020)
Constant	0.059	0.025	0.078	0.067	0.190*
	(0.103)	(0.104)	(0.097)	(0.090)	(0.101)
Observations	1,034	984	1,007	970	1,000

Panel C. Nonstandardized grades

IVV	-0.029	-0.005	-0.031	-0.024	0.066**
	(0.031)	(0.042)	(0.026)	(0.027)	(0.026)
Constant	6.772***	6.499***	6.740***	6.723***	7.202***
	(0.161)	(0.164)	(0.143)	(0.118)	(0.130)
Observations	1,034	984	1,007	970	1,000

We estimated the correlation between the IVV prediction with academic grades and misbehavior reports before the intervention using administrative data. The estimated specification was the following: $y_{ij} = \alpha_0 + \alpha_1 IVV_{ij} + \epsilon_{ij}$, where y_{ij} is the academic grade or misbehavior report for student i in school j and IVV_{ij} is the estimated propensity for violence. ***, **, * indicate that coefficients are significant at 1%, 5%, and 10%, respectively. Robust standard errors at course-school level are in parentheses.

TABLE A7. SUMMARY STATISTICS: MEANS OF VARIABLES BY GROUP FOR THE SELECTED SUBSAMPLE

	(1) Full Sample	(2) Control Group (C)	(3) Treatment Group (T)	(4) Unadjusted p-values (C = T)
PANEL A: IVV DETERMINANTS				
Student is male	0.49	0.50	0.49	0.797
Student's age	11.9	11.8	12.0	0.074
Student lives in urban area	0.75	0.74	0.75	0.760
Student's household composition				
Student lives with both parents	0.53	0.48	0.55	0.155
Student lives with only one parent	0.32	0.37	0.31	0.250
Student lives with one parent and stepparent	0.06	0.06	0.06	0.885
Student lives with other relatives/adults	0.08	0.10	0.08	0.488
Student's mother's level of education:				
Basic education (1–6 years)	0.31	0.34	0.29	0.230
Intermediate education (7–12 years)	0.61	0.56	0.63	0.181
University or higher (13 years +)	0.08	0.10	0.08	0.385
Student's travel time from house to school (min.)	17.2	15.8	17.6	0.192
Student is alone at home after school	0.05	0.06	0.05	0.885
Student's school year	5.72	5.62	5.75	0.011
Student is enrolled in the morning shift	0.72	0.74	0.71	0.488
Student's violence index	0.038	0.037	0.039	0.502
PANEL B: ACADEMIC OUTCOMES				
Academic scores Q1 2016 (Baseline)				
Reading scores	6.71	6.49	6.72	0.108
Math scores	6.55	6.52	6.56	1.000
Science scores	6.66	6.50	6.72	0.150
Behavior scores	7.40	7.27	7.44	0.198
Absenteeism Q1 2016	1.83	2.20	1.71	0.329
Observations	598	145	453	
Missing share	-	0.48	0.52	0.482

 This table shows descriptive statistics of the available variables at baseline for the randomly selected subsample for emotional regulation data collection. Panel A summarizes information obtained from the enrollment form that was used as determinants in the IVV estimation. Panel B presents administrative data provided by schools only from students who had consented. This data is from the first quarter of academic year 2016, before the clubs were implemented. The scale of grades in El Salvador is 0–10 points. In the comparison between treated and control groups within this subsample, students in the treatment group are older and enrolled in higher courses, on average. The only statistical difference at 10% in the comparison between the randomly selected subsample and the rest of observations was on the academic grades on math at baseline.

TABLE A8. MEANS OF VARIABLES BY SUBSAMPLES WITH VALID AND NONVALID EEG

	(1) Valid EEG	(2) Nonvalid EEG	(3) Unadjusted p-values (1) = (2)	(4) Adjusted p-values (1) = (2)
PANEL A: IVV DETERMINANTS				
Student is male	0.44	0.54	0.022	0.36
Student's age	13.1	10.7	0.903	1.00
Student lives in urban area	0.72	0.77	0.015	0.31
Student's household composition				
Student lives with both parents	0.50	0.56	0.821	1.00
Student lives with only one parent	0.33	0.31	0.807	1.00
Student lives with one parent and stepparent	0.06	0.07	0.074	0.72
Student lives with other relatives/adults	0.10	0.06	0.267	0.89
Student's mother's level of education:				
Basic education (1–6 years)	0.32	0.29	0.207	0.88
Intermediate education (7–12 years)	0.59	0.63	0.911	1.00
University or higher (13 years +)	0.09	0.08	0.086	0.72
Student's travel time from house to school (min.)	17.0	17.4	0.158	0.81
Student is alone at home after school	0.05	0.06	0.130	0.80
Student's school year	6.84	4.53	0.942	1.00
Student is enrolled in the morning shift	0.64	0.80	0.257	0.89
Student's violence index	0.04	0.03	0.389	0.97
PANEL B: ACADEMIC OUTCOMES				
Academic scores Q1 2016 (Baseline)				
Reading scores	6.66	6.76	0.094	0.42
Math scores	6.30	6.83	0.999	1.00
Science scores	6.41	6.93	0.190	0.52
Behavior scores	7.43	7.37	0.097	0.43
Absenteeism Q1 2016	1.68	1.99	0.590	0.88
Observations	308	290	-	-

This table shows descriptive statistics of the available variables at baseline for the subgroups with valid and nonvalid EEG measures. Panel A summarizes information obtained from the enrollment form that was used as determinants in the IVV estimation. Panel B presents administrative data provided by schools only from students who had consented. This data is from the first quarter of academic year 2016, before the clubs were implemented. The scale of grades in El Salvador is 0–10 points. Take-up is estimated as number of hours attended by student i/total hours in each club. p-values have been adjusted for multiple hypothesis testing of means following Jones et al. (2018).

TABLE A9. *p*-VALUES OF MISSING VARIABLES FOR 13 OBSERVATIONS

	(1) Locus of Control	(2) CRT	(3) Raven
Student is male	0.457	0.457	0.457
Student's age	0.573	0.573	0.573
Student lives in urban area	0.524	0.524	0.524
Student's travel time from house to school (min.)	0.544	0.544	0.544
Student's violence index	0.041	0.041	0.041
Student's school year	0.189	0.189	0.189
Student's household composition			
Student lives with both parents	0.851	0.851	0.851
Student lives with only one parent	0.320	0.320	0.320
Student lives with a parent and a stepparent	0.632	0.632	0.632
Student lives with other relatives/adults	0.381	0.381	0.381
Student's mother's level of education:			
Basic education (1–6 years)	0.784	0.784	0.784
Intermediate education (7–12 years)	0.765	0.765	0.765
University or higher (13 years +)	0.963	0.963	0.963
Homog. Vs Control	0.181	0.181	0.181
Heterog. Vs Control	0.129	0.129	0.129
Homog. Vs Heterog.	0.522	0.522	0.522
Homog. High vs Homog. Low	0.627	0.627	0.627
Arousal	0.775	0.775	0.775
Valence	0.162	0.162	0.162
Positive Valence	0.213	0.213	0.213
Negative Valence	0.397	0.397	0.397
Positive Valence Difference	0.248	0.248	0.248
Negative Valence Difference	0.166	0.166	0.166

***, **, * significant at 1%, 5%, and 10%, respectively. Bootstrapped standard errors at the course-school level are in parentheses. All outcomes have been standardized at the control-course level with a mean of 0 and standard deviation 1.0. All regressions include as controls: a second-order polynomial of student's propensity for violence and education-level-by-school fixed effects (stratification level). Differences in number of observations are due to variation in the response rate for each outcome.

TABLE A10. OVERALL EFFECTS OF THE ASP ON EMOTIONAL REGULATION

Heckman correction

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	Arousal (stress)	Valence	Positive Valence Difference	Negative Valence Difference	Locus of Control	CRT	Raven
Any treatment	0.048	-0.330**	-0.415**	-0.237	-0.240**	-0.004	-0.006
	(0.129)	(0.172)	(0.196)	(0.251)	(0.133)	(0.131)	(0.135)
Observations	598	598	598	598	585	585	585
Lambda	-0.425	0.429	0.445	0.126	0.213	0.177	0.648
	(0.423)	(0.978)	(0.971)	(1.192)	(0.506)	(0.510)	(0.654)
rho	-0.528	0.258	0.243	0.062	0.213	0.182	0.636
sigma	0.804	1.658	1.827	2.022	1.000	0.973	1.018

***, **, * significant at 1%, 5%, and 10%, respectively. Bootstrapped standard errors at the course-school level are in parentheses. All outcomes have been standardized at the control-course level with a mean of 0 and standard deviation 1.0. All regressions include as controls: a second-order polynomial of student's propensity for violence and education-level-by-school fixed effect (stratification level). Differences in number of observations are due to variation in the response rate for each outcome.

TABLE A11. OVERALL EFFECTS OF THE ASP ON EMOTIONAL REGULATION
Lee (2009) treatment effects bounds

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	Arousal (stress)	Valence	Positive Valence Difference	Negative Valence Difference	Locus of Control	CRT	Raven
Lower	-0.030	-0.556***	-0.664***	-0.583**	-0.338	-0.215	-0.215
	(0.135)	(0.188)	(0.221)	(0.239)	(0.212)	(0.132)	(0.182)
Upper	0.234*	-0.035	-0.018	0.205	-0.101	4.656***	0.005
	(0.135)	(0.188)	(0.237)	(0.251)	(0.178)	(0.119)	(0.171)
Obs.	598	598	585	585	585	598	598
Selected obs.	308	308	308	308	295	295	295

This table show bounds estimations using the leebounds STATA command developed by Tauchman (2012), which computes treatment effect bounds for samples with nonrandom sample selection/attrition as proposed by Lee (2009). Bootstrapped standard errors are in parentheses. All outcomes have been standardized at the control-course level with a mean of 0 and standard deviation 1.0.

TABLE A12. OVERALL EFFECTS OF THE ASP – LEARNING MECHANISM

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)

PANEL A: VIOLENCE AND ATTITUDES

	Attitudes toward school and learning				Violence and Behavior				
	Positive attitudes toward school	Time to do homework (hours)	Pay attention in class	Absenteeism (days)	Delinquency (Index)	Violent actions (Index)	Approval of antisocial behavior	Behavior reports (-)	Probability of bad behavior report
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Any treatment	0.154**	0.360***	0.064***	-1.474***	-0.209***	-0.161***	-0.107***	0.187***	0.071***
	(0.061)	(0.092)	(0.024)	(0.252)	(0.071)	(0.046)	(0.025)	(0.056)	(0.024)
Observations	845	830	857	741	814	850	850	897	897

PANEL B: ACADEMIC OUTCOMES

	Grades				Probability of passing				Failing at least one course
	Reading	Math	Science	Score	Reading	Math	Science	Score	
Any treatment	0.037	0.124***	0.146***	0.076**	0.048***	0.025	0.034**	0.032**	-0.031***
	(0.041)	(0.041)	(0.047)	(0.037)	(0.008)	(0.016)	(0.014)	(0.015)	(0.009)
Observations	907	907	907	907	907	907	907	907	907

*** , ** , * significant at 1%, 5%, and 10%, respectively. Bootstrapped standard errors at the course-school level are in parentheses. Panel A presents effects on noncognitive outcomes. Panel B presents results on academic outcomes. Estimations are restricted to sample that attended at least one session of the ASP. All regressions include as controls: a second-order polynomial of student's IVV and ciclo-school fixed effect (stratification level). Additionally, in estimations for academic outcomes, absenteeism, and bad behavior reports, we also include the corresponding imputed outcome at the baseline and a dummy indicating a missing value at the baseline. Differences in number of observations of noncognitive outcomes are due to variation in the response rate for each outcome.

2. Appendix estimations

Appendix 1. Description of Outcome Variables.

In our follow-up survey, we have multiple variables that measure some behavioral outcomes, such as attitudes, delinquency, and violent behavior. In order to have a single continuous measure that can be compared to previous evidence in the literature, we have built for some of them a standardized index that is an average of the multiple variables measured in the survey.

In the following section, we provide details of each outcome variable. For the index outcomes construction, we provide information of the main items included.

A. Behavior and Academic Outcomes

1. *Positive attitudes towards school and learning:* standardized index estimated using principal components analysis. We used 5 items from the self-reported follow-up survey, each of them was on a 1-4 scale, where 1 = very important, and 4 = no important. The items were the following: (i) Indicator of how important is to learn for them, (ii) indicator of how they like their school, (iii) indicator of studying is an opportunity for having a better future, (iv) indicator of whether hard work at schools pays off, and (v) indicator of whether what they learn at school is relevant for their future.

2. *Time spent on homework*: This variable was a self report from students in time units (minutes). The question was: *During the last 3 months, how much time did you spend doing your homework aside from the time you were at school or in classes?*

3. *Paying attention in class*: This outcome was a self report from students. The question was: *When I am in class, I pay attention* The answers were on a 1-4 scale, in which 1 = always and 4 = never. Using this scale, we created a dummy that takes the value of 1 if student reports that she pay attention most of the time (1 or 2) and 0 if she barely pays attention during classes (3 or 4).

4. *Delinquency Index*: corresponds to a standardized sum of self reported delinquent actions, such as theft, mugging someone, etc. The question was: *During the last 3 months, did you ...?* This variable can be interpreted as an aggregation of delinquency actions.

5. *Violent Actions Index*: is the standardized sum of other violent acts such as fighting at school, damage of municipal property, fight with siblings, etc. The question was: *During the last 3 months, did you ...?* As before, this index variable can be interpreted as an aggregation of violent actions.

6. *Approval of peers' antisocial behavior* is a binary indicator that takes the value of 1 if students approve any peer behavior related to alcohol and drugs consumption, fighting, etc. The question was: *What do you think if someone of your closest friends ...?*

7. *Absenteeism*: corresponds to the number of days the student did not attended school between April-October of the 2016 academic year. This was administrative data provided by schools at follow-up.

8. *Bad behavior reports*: In El Salvador, these are reported by teachers each quarter. They are presented on the following discrete scale: Excellent (E), Very Good (MB), Good (B), and Regular (R). It can be translated in a continuous scale that is comparable to course grades. In this paper, we used a reversed continuous scale to facilitate the interpretation and comparability to the self-reported measures of violence and crime.

9. *Reading, math and science grades*: Variables that indicate performance on each course. It is a 0-10 scale, where 0 is the worst performance and 10 is the best. We have standardized these values from control groups at the school-grade level. Score is an average of the three courses.

10. *Passing course*: Is a dummy variable that takes the value of 1 if student has been promoted to the following course and 0 otherwise.

B. Neurophysiological Outcomes

1. *Arousal*: pre-test resting measure of individual's stress, estimated directly from her brain activity using EEG recordings measured while children were watching a black cross in the center of a gray screen for a period of 30 seconds.

2. *Valence*: pre-test resting state measure estimated directly from participants' brain activity using EEG recordings. As the arousal measure, this recordings were estimated while children were watching a black cross in the center of a gray screen for a period of 30 seconds. This variable can be interpreted as a positive or negative mood, as well as an attitude of either approach or withdrawal towards/from a stimulus (Harmon-Jones et al., 2010; Kassam et al., 2013).

3. *Locus of control*: Psychometric test developed by Rotter (1966). This indicates that children think that they are not able to control what happens in their lives.

4. *Cognitive Reflection Test (CRT)*: Is a test designed to measure if an individual tends to automatically choose an initially incorrect response and then engage in a deeper reasoning to find a correct answer.

5. *Raven*: Is a measure of abstract reasoning and a non-verbal estimate of intelligence. It is implemented as a set of matrices in progressive order.

6. *Positive Valence Difference*: Corresponds to the difference between the response intensity measure after exposure to positive stimuli and the valence-at-resting-state index described before.

7. *Negative Valence Difference*: Is a measure of the variation in the valence index recorded when the stimulus was negative net of the individual's baseline resting state valence index. Both differences can be interpreted as a lower level of overreaction of participants –they become more phlegmatic or cold headed– or that individuals move towards a more withdrawal behavior or attitude.

Appendix 2. Heterogeneity of the ASP by baseline grades.

As mentioned before, since the intervention provides life skills training and promotes positive attitudes towards school and learning, according to the NGO's theory of change it may also improve children's academic attainment. Previous papers have shown that interventions with similar approaches to the ASP analyzed in this paper have greater impact to the most vulnerable students –defined as those with lower academic performance– compared to the rest of their class (Durlak et al., 2010).

The main concern in the estimation of heterogeneous effects by baseline academic performance under our experiment design is that the differences in the impacts from the ASP can be caused mostly by children's propensity for violence (IVV) than by their initial academic attainment.

However, as we showed in table A6 in the Appendix, the predicted IVV for this particular sample is not correlated with grades at the baseline, that is, children with lower academic performance are not necessarily those with greater probability of being violent. This result is policy-relevant because it indicates that we have actually two vulnerable groups in schools located in highly violent contexts.

Exploiting this lack of correlation between grades and estimated IVV in this sample, we assess the heterogeneous effects by initial academic achievement. We include a dummy variable A_{ij}, which indicates whether child i was in the bottom half of the baseline score distribution in their course. This score is an average of the grades achieved by the student in her three main courses: math, reading and science during the first quarter of the 2016 academic year, before the intervention. We also add an interaction between this dummy and the treatment dummy. The resulting equation used to identify differential effects of the program by academic performance at baseline is the following:

$$y_{ijt} = \theta_0 + \theta_1 T_{ijt} + \theta_2 T_{ijt} \times A_{ij(t-1)} + \theta_3 A_{ij(t-1)} + \theta_4 X_{ij(t-1)} + S_j + \epsilon_{ijt} \tag{4}$$

The rest of variables are defined as before. Results are shown in table A13. As before, Panel A shows violence and attitudes outcomes and Panel B shows academic performance results. Row [i] in both panels shows the results for students with low academic performance before the intervention (θ_1) and row [ii] shows the results for students with a score higher than the median within her course (θ_3) compared to their similar peers in the control group. The differential effect of the intervention between treated subgroups with different initial academic performance (θ_2) is presented in row [iii].

We find that treated students with lower initial academic achievement reduce their absenteeism by 1.8 days more than treated students with high academic performance. There are no differences in the effects on the rest of behavioral outcomes for either group. Regarding academic outcomes, results indicate that the effects on the extensive margin are higher for those students in the bottom of the grade distribution, including an increase in the probability of course promotion.

Combining these results with the heterogeneous effects results by initial IVV presented before, we can conclude that the ASP is benefiting the most vulnerable children, which are those with either higher propensity for violence or lower academic performance.

TABLE A13. HETEROGENEOUS TREATMENT EFFECTS BY INITIAL ACADEMIC ACHIEVEMENT

PANEL A: VIOLENCE AND ATTITUDES

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
	Attitudes towards school and learning				Violence and Behavior				
	Positive attitudes towards school	Time to do homework (hours)	Pay attention in class	Absenteeism (days)	Delinquency (Index)	Violent actions (Index)	Approval of antisocial behavior	Behavior reports (-)	Probability of bad behavior report
[i] Effect on low performance students	0.222	0.270	0.072*	-2.706***	-0.126	-0.137	-0.116***	-0.187***	-0.097***
	(0.151)	(0.167)	(0.040)	(0.616)	(0.103)	(0.094)	(0.030)	(0.046)	(0.027)
[ii] Effect on high performance students	0.157**	0.439***	0.091***	-0.898**	-0.285***	-0.166**	-0.095***	-0.235***	-0.059*
[iii] Difference between high and low performers	-0.065	0.170	0.020	1.808**	-0.160	-0.029	0.021	-0.048	0.038
	(0.199)	(0.206)	(0.058)	(0.910)	(0.145)	(0.146)	(0.042)	(0.099)	(0.046)
[iv] p-value: effect on Top half = effect on bottom half	0.746	0.410	0.737	0.047	0.270	0.842	0.622	0.627	0.412
Observations	948	935	962	833	916	956	962	1010	1010

PANEL B: ACADEMIC OUTCOMES

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
	Grades				Probability of passing				Passing course
	Reading	Math	Science	Score	Reading	Math	Science	Score	
[i] Effect on low performance students	0.154***	0.183***	0.235***	0.211***	0.098***	0.061***	0.089***	0.103***	-0.073***
	(0.043)	(0.037)	(0.052)	(0.037)	(0.014)	(0.021)	(0.026)	(0.029)	(0.018)
[ii] Effect on high performance students	0.076**	0.121***	0.184***	0.128***	0.002	-0.002	-0.003	-0.005	0.001
[iii] Difference between high and low performers	-0.077	-0.062	-0.050	-0.083	-0.096***	-0.064***	-0.092***	-0.108***	0.073***
	(0.058)	(0.055)	(0.066)	(0.051)	(0.019)	(0.024)	(0.028)	(0.034)	(0.019)
[iv] p-value: effect on Top half = effect on bottom half	0.182	0.259	0.449	0.106	0.000	0.009	0.001	0.001	0.000
Observations	1023	1023	1023	1023	1023	1023	1023	1023	1023

*** ** * significant at 1%, 5% and 10% respectively. Bootstrapped standard errors at the course-school level are in parentheses. Panel A present results on academic outcomes. Panel B presents effects on non-cognitive outcomes. Description of the outcome variables is available in Appendix 1. Row Total effects on High Scores is the sum of the coefficients of any treatment dummy and the coefficient of the interaction term. All regressions include as controls: a second order polynomial of student's IVV, and ciclo-school fixed effect (stratification level). Additionally, in estimations for academic outcomes, absenteeism and misbehavior reports, we also include the corresponding imputed outcome at the baseline and a dummy indicating a missing value at baseline.

Appendix 3. Gender vs. propensity for violence heterogeneous effects.

Previous studies have found that gender-mixed ASP usually impact differently to boys and girls (Durlak et al., 2010). They regularly identify this difference by incorporating an interaction between the gender variable and the treatment indicator.

Under this approach, we would estimate the following equation:

$$y_{ijt} = \theta_0 + \theta_1 T_{ijt} + \theta_2 T_{ijt} \times G_{ij(t-1)} + \theta_3 G_{ij(t-1)} + \theta_4 X_{ij(t-1)} + S_j + \epsilon_{ijt} \tag{5}$$

where G_{ij} is a dummy that takes the value of 1 if the child is a boy.

Estimated coefficients of this equation are presented in table A14. Row [i] in both panel A and B shows the results for girls before the intervention (θ_1) compared to other girls in the C group. Row [ii] shows the results for treated boys (θ_3) compared to other boys in C group. The coefficient of the interaction term (θ_2) would indicate the difference in the effects of the ASP between boys and girls. This estimation is presented in row [iii].

Following this approach we find higher effects on absenteeism for treated boys compared to treated girls (a reduction of 2.1 days of absenteeism). Additionally, the impact on the extensive margin of school grades is more significant for treated boys on math and score, compared to treated girls.

However, the aforementioned approach may provide naive estimates of the intervention differential effects because the estimation of the IVV includes sex as a determinant. Thus, a concern in the study of heterogeneous effects among boys and girls is that they may be caused either by gender alone or by the combination of it and the rest of determinants included in the IVV estimation.

For example, as we can see from the previous results, most of the differences by gender are found on the same outcomes as the differences by initial propensity for violence: absenteeism and intensive margin of academic performance.

To verify which of the measures –gender or propensity for violence– are generating the differences, we run "horse-races" between them, as shown in the following alternative specification:

$$y_{ijt} = \theta_0 + \theta_1 T_{ijt} + \theta_2 T_{ijt} \times G_{ij(t-1)} + \theta_3 T_{ijt} \times IVV_{ij(t-1)} + \theta_4 X_{ij(t-1)} + S_j + \epsilon_{ijt} \tag{6}$$

where θ_2 indicates the difference of the ASP impacts by gender (boys versus girls) and θ_3 shows the difference of the impact by the propensity for violence (highly versus low violent children). In the control variables vector $X_{ij(t-1)}$, we include gender, high-IVV dummy and a second order polynomial of students' percentile of initial IVV. The rest of variables are defined as before.

Appendix table A15 shows the results, separated in the two main panels. We present only the estimated coefficients θ_2 and θ_3 in rows [i] and [ii] respectively. Results reinforce the previous conclusion that the heterogeneous effects on academic attainment and misbehavior at school reported in Table 3 are in fact driven by students' initial propensity for violence, except for absenteeism. Moreover, we are finding that there are some heterogeneous effects by gender on attitudes towards school and the extensive margin of reading grades.

TABLE A14. HETEROGENEOUS TREATMENT EFFECTS BY GENDER.

PANEL A: VIOLENCE AND ATTITUDES

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
	Attitudes towards school and learning				Violence and Behavior				
	Positive attitudes towards school	Time to do homework (hours)	Pay attention in class	Absenteeism (days)	Delinquency (Index)	Violent actions (Index)	Approval of antisocial behavior	Behavior reports (-)	Probability of bad behavior report
[*i*] Effect on girls	0.064	0.300**	0.031	-0.604	-0.255**	-0.101	-0.100***	-0.241***	-0.085***
	(0.094)	(0.138)	(0.034)	(0.576)	(0.112)	(0.064)	(0.029)	(0.087)	(0.029)
[*ii*] Effect on boys	0.296***	0.376**	0.127**	-2.692***	-0.133	-0.168**	-0.106***	-0.100	-0.041
[*iii*] Difference between girls and boys	0.232*	0.076	0.096	-2.088**	0.122	-0.067	-0.006	0.141	0.044
	(0.128)	(0.229)	(0.062)	(0.985)	(0.168)	(0.105)	(0.038)	(0.132)	(0.061)
Observations	948	935	962	836	916	956	962	1010	1010

PANEL B: ACADEMIC OUTCOMES

	Reading	Math	Science	Score	Reading	Math	Science	Score	Passing course
	Grades				Probability of passing				
[*i*] Effect on girls	-0.031	0.010	0.097	-0.007	0.010	0.002	0.022	0.017	-0.010
	(0.055)	(0.061)	(0.064)	(0.055)	(0.015)	(0.023)	(0.023)	(0.020)	(0.016)
[*ii*] Effect on boys	0.057	0.201***	0.163**	0.119**	0.063***	0.034	0.037	0.033	-0.045***
[*iii*] Difference between girls and boys	0.088	0.191**	0.066	0.126*	0.053**	0.032	0.015	0.016	-0.035
	(0.078)	(0.083)	(0.090)	(0.077)	(0.021)	(0.032)	(0.037)	(0.031)	(0.023)
Observations	1023	1023	1023	1023	1023	1023	1023	1023	1023

***, **, * significant at 1%, 5% and 10% respectively. Bootstrapped standard errors at the course-school level are in parentheses. Panel A present results on academic outcomes. Panel B presents effects on non-cognitive outcomes. Description of outcome variables is available in Appendix 1. Row "Total effects on Boys" is the sum of the coefficients of any treatment dummy and the coefficient of the interaction term. All regressions include as controls: a second order polynomial of student's IVV, and education level fixed effect (stratification level). Additionally, in estimations for academic outcomes, absenteeism and bad behavior reports, we also include the corresponding imputed outcome at the baseline and a dummy indicating a missing value at the baseline.

TABLE A15. HETEROGENEOUS TREATMENT EFFECTS BY GENDER AND VIOLENCE.

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
PANEL A: VIOLENCE AND ATTITUDES									
	Attitudes towards school and learning				Violence and Behavior				
	Positive attitudes towards school	Time to do homework (hours)	Pay attention in class	Absenteeism (days)	Delinquency (Index)	Violent actions (Index)	Approval of antisocial behavior	Behavior reports (−)	Probability of bad behavior report
[*i*] T x Boy	0.202	0.272	0.140**	-1.831**	0.128	-0.014	0.053	-0.024	-0.032
	(0.201)	(0.312)	(0.066)	(0.854)	(0.201)	(0.123)	(0.035)	(0.151)	(0.071)
[*ii*] T x High IVV	0.048	-0.339	-0.077	-0.449	-0.012	-0.092	-0.103**	-0.205*	-0.021
	(0.261)	(0.302)	(0.068)	(0.951)	(0.210)	(0.156)	(0.046)	(0.109)	(0.052)
Observations	948	935	962	836	916	956	956	1010	1010

PANEL B: ACADEMIC OUTCOMES									
	Grades				Probability of passing				Failing at least one course
	Reading	Math	Science	Score	Reading	Math	Science	Score	
[*i*] T x Boy	0.031	0.076	-0.078	0.024	0.071**	0.005	-0.004	0.024	-0.018
	(0.123)	(0.118)	(0.139)	(0.121)	(0.035)	(0.047)	(0.056)	(0.052)	(0.027)
[*ii*] T x High IVV	0.094	0.193*	0.245*	0.175	-0.032	0.047	0.032	-0.014	-0.029
	(0.118)	(0.104)	(0.126)	(0.108)	(0.039)	(0.047)	(0.056)	(0.051)	(0.031)
Observations	1023	1023	1023	1023	1023	1023	1023	1023	1023

***, **, * significant at 1%, 5% and 10% respectively. Bootstrapped standard errors at the course-school level are in parentheses. Panel A present results on academic outcomes. Panel B presents effects on non-cognitive outcomes. Description of outcome variables is available in Appendix 1. Row Total effects on Boys is the sum of the coefficients of any treatment dummy and the coefficient of the interaction term. All regressions include as controls: a second order polynomial of student's IVV, and ciclo-school fixed effect (stratification level). Additionally, in estimations for academic outcomes, absenteeism and bad behavior reports, we also include the corresponding imputed outcome at the baseline and a dummy indicating a missing value at the baseline.

Appendix 4. Further analysis and evidence on spillovers.

In this Appendix, we present further evidence of spillover characteristics in the context of this ASP. First, in the primary analysis of the intervention impact, we find that mainly students with a higher propensity for violence benefit more from the program. However, the results of group composition effects found by Dinarte (2018) indicate that the gains of highly violent students stem mainly from exposure to a diversity of peers regarding violence. Therefore, treating both groups of students—with high and low propensity for violence—may maximize overall results.

To test if this is also true for spillover estimations, we divide the share of treated students into groups of high and low propensity for violence. The estimation equation is the following:

$$y_{mnt} = \gamma_0 + \gamma_1 ShH_n + \gamma_2 ShL_n + \gamma_3 X_{mn(t-1)} + E_n + \epsilon_{mnt} \tag{7}$$

where ShH_n and ShL_n are the share of treated students with high and low IVV at the classroom level, respectively. As before, E_n corresponds to the share of all enrolled students—T and C groups—from each course, and $X_{mn(t-1)}$ is a vector of individual controls, including grades at the baseline and a missing grades dummy.

Results are shown in Table A16. Unfortunately, we do not have enough power to find statistically significant differences between the shares of treated students with low and high levels of violence. However, the signs of the estimated coefficients may suggest that spillover effects on academic outcomes can be driven by the share of treated students with low levels of violence. However, the reduction in misbehavior at school is caused mainly by the share of treated students with high propensity for violence.

The second analysis we implemented was to test if the intensity of these spillovers may change due to the level of exposure—in terms of time—of nonenrolled children to treated participants. To measure intensity of exposure, we exploit the fact that nonenrolled children usually spend more time with students from their own classroom compared to treated students from other classrooms. To study this between-classrooms closeness, we estimate the following equation:

$$y_{mnt} = \gamma_0 + \gamma_1 Sh_n + \gamma_2 Sh_{n-1} + \gamma_3 Sh_{n+1} + \gamma_4 X_{mn(t-1)} + E_n + \epsilon_{mnt} \tag{8}$$

where Sh_n is the share of treated children at own student's classroom n, and Sh_{n-1} and Sh_{n+1} are the share of treated students in the previous and next course, respectively. The rest of the variables are defined as before.

TABLE A16. SPILLOVERS BY STUDENT PROPENSITY FOR VIOLENCE

	(1)	(2)	(3)	(4)	(5)
	Grades				Bad behavior
	Reading	Math	Science	Score	reports (-)
[i] Proportion of treated students with high propensity for violence	0.004 (0.004)	0.007* (0.004)	0.006 (0.004)	0.005 (0.004)	-0.014** (0.005)
[ii] Proportion of treated students with low propensity for violence	0.009*** (0.003)	0.008** (0.003)	0.005* (0.003)	0.008*** (0.003)	-0.010 (0.008)
[iii] p-value [i] = [ii]	0.219	0.980	0.948	0.594	0.667
Observations	1357	1357	1357	1357	1194

***, **, * significant at 1%, 5%, and 10%, respectively. Robust standard errors at course-school level are in parentheses. Outcome variables are standardized grades at school-grade level at follow-up. All regressions include as main control the share of enrolled students from each course. Individual controls include imputed grades in the course at baseline and a dummy indicating a missing value in the grade at baseline. Row [i] indicates the effect of the share of treated students with high propensity for violence within each classroom. Similarly, row [ii] indicates the effect of the proportion of treated students with lower propensity for violence. Row [iii] is the p-value of the hypothesis that the difference between both coefficients is statistically different from 0.

As we can see in Table A17, spillovers on a nonenrolled student's academic outcomes are led only by the share of treated students from her own classroom. Nevertheless, a novel result here is that the effect on bad behavior at school is caused by both the percentage of treated from her classroom and one course below. To better understand this last result, further analysis of the social interactions within schools is necessary, using sociograms, for example. However, from the results, we can infer that most of the interactions seem to come from treated children with whom nonenrolled students spend relatively more time.

Finally, spillover effects may be different by misbehavior closeness of nonenrolled with treated students within the same classroom. Since the ASP effects are modified by the initial propensity for violence of treated participants, there may also be heterogeneity in spillover effects by nonenrolled students' misbehavior at school before the intervention.

Since we rely on only administrative data for nonenrolled students—i.e., we do not have an IVV measure for them—to test this within-classroom closeness, we use misbehavior reports at school for all children. Then we create dummies indicating if each nonenrolled student is less than i standard deviations away from the average misbehavior of her treated classmates before the intervention, with $i \in \{1, 2, +2\}$. Finally, we estimate the following specification:

$$y_{mnt} = \gamma_0 + \gamma_1 Sh_n + \gamma_2 Sh_n \times C1_{mn(t-1)} + \gamma_3 Sh_n \times C2_{mn(t-1)} + \gamma_4 X_{mn(t-1)} + E_n + \epsilon_{mn} \qquad (9)$$

where $Ci_{mnj(t-1)}$ are dummies indicating whether student m has a bad behavior level that is less than i standard deviations from the average behavior of treated children in her classroom n. Other variables are defined as before.

Results are presented in Table A18. We find the effects are more significant for students whose bad behavior at school is between 1 and 2 standard deviations away from the mean of misbehavior of the share of treated students from their classroom. Notably, the effects of this intermediate closeness are more significant on bad behavior reports. Thus, similar to Dinarte (2018)'s finding regarding the importance of diversity in this context, this result also highlights that only certain levels of similarity to treated students can have positive spillover effects.

TABLE A17. RELATIVE SPILLOVER EFFECTS

	(1)	(2)	(3)	(4)	(5)
		Grades			Bad behavior
	Reading	Math	Science	Score	reports (-)
[i] Proportion of treated students at classroom m (own classroom)	0.007**	0.007***	0.006**	0.007***	-0.009*
	(0.003)	(0.003)	(0.003)	(0.003)	(0.005)
[ii] Proportion of treated students at classroom $m-1$	-0.001	0.001	0.001	0.001	-0.005*
	(0.001)	(0.001)	(0.001)	(0.001)	(0.002)
[iii] Proportion of treated students at classroom $m+1$	-0.001	0.000	-0.002	-0.001	0.005
	(0.001)	(0.001)	(0.002)	(0.001)	(0.003)
p-value $[i] = [ii]$	0.0349	0.0342	0.0785	0.0386	0.4485
p-value $[i] = [iii]$	0.0485	0.0254	0.0164	0.0253	0.0392
p-value $[ii] = [iii]$	0.9835	0.8009	0.2131	0.5130	0.0352
Observations	1357	1.327	1.326	1.356	1135

***, **, * significant at 1%, 5%, and 10%, respectively. Robust standard errors at course-school level are in parentheses. Outcome variables are standarized grades at school-grade level at follow-up. All regressions include as main control the share of enrolled students from each course. Individual controls include imputed grades in the course at baseline and a dummy indicating a missing value in the grade at baseline. Row [i] indicates the affect of the share of treated students within own student's classroom (m). Row [ii] indicates the effect of the proportion of treated students within one course lower ($m-1$) than student's own classroom. Row [iii] is similar to the previous row but related to the share of treated students one course greater ($m+1$). p-values are related to the null hypothesis that the difference between each pair of coefficients is different from 0.

TABLE A18. RELATIVE SPILLOVER HETEROGENEOUS EFFECTS

	(1)	(2)	(3)	(4)	(5)
		Grades			Bad behavior
	Reading	Math	Science	Score	reports (-)
[i] Spillovers on nonenrolled whose bad behavior report is within 1 sd from treated students	0.007** (0.003)	0.007*** (0.003)	0.005 (0.003)	0.007** (0.003)	-0.005 (0.005)
[ii] Spillovers on nonenrolled whose bad behavior report is within 1-2 sd from treated students	0.006** (0.002)	0.007*** (0.002)	0.008*** (0.002)	0.006** (0.002)	-0.019*** (0.006)
[iii] Spillovers on nonenrolled whose bad behavior report is more than 2 sd from treated students	0.001 (0.004)	0.008** (0.004)	-0.001 (0.005)	0.004 (0.004)	0.002 (0.007)
p-value $[i] = [ii]$	0.867	0.858	0.417	0.979	0.076
p-value $[i] = [iii]$	0.036	0.623	0.168	0.366	0.121
p-value $[ii] = [iii]$	0.286	0.700	0.127	0.578	0.018
Observations	1.357	1.327	1.326	1.356	1.135

***, **, * significant at 1%, 5%, and 10%, respectively. Robust standard errors at course-school level are in parentheses. Outcome variables are standardized grades at school-grade level at follow-up. All regressions include as main control the share of enrolled students from each course. Individual controls include imputed grades in the course at baseline and a dummy indicating a missing value in the grade at baseline. Row [i] shows spillover effect on outcomes for nonenrolled students with a bad behavior level of 1 sd from their treated classmates (at baseline). Row [ii] shows the spillover effect on those nonenrolled who had a bad behavior level of 2 sd from the average of their treated classmates. Row [iii] exhibits spillovers for nonenrolled students with a bad behavior level at baseline that was 3 or more sd from their treated classmates.

3. Appendix figures

Figure 3: A1. Heterogeneous effects by propensity for violence before the ASP
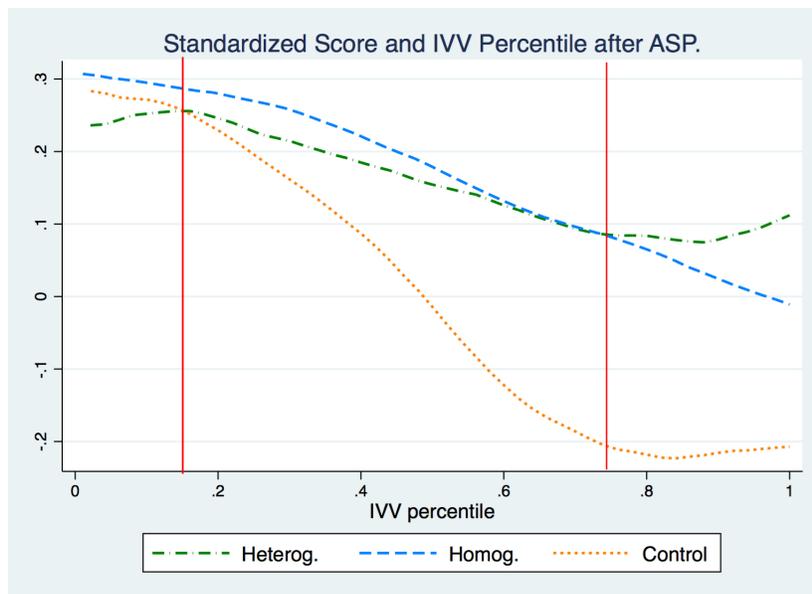
Figure 4: A2. Lab-in-the-field Setting

This photo was taken at the "Juan Ramon Jimenez" School in La Libertad, El Salvador during the EEG recordings collection. At the time, students were in the third stage of the data-collection process. While they watched images, the headset recorded brain responses after the stimuli generated through the images.

5. Appendix EEG and Behaviors

How does arousal of valence correlate with behavior or actions?

The literature interprets arousal as the inverse of alpha power in the prefrontal cortex and as a proxy for stress or alertness of the individual. The valence index aims to proxy the positive or negative nature of the emotion that the person experiences. The literature features a varied discussion due to the process of building these indexes, usually using external stimuli. On one hand, researchers argue they are able to proxy the actual emotion that the individual experiences (Davidson et al., 1990; Harmon-Jones et al., 2010). On the other, authors maintain that EEG recordings capture the motivation to the stimuli that the individual feels during exposure. In this regard, we can interpret positive and negative valence indexes as approach and withdrawal motivations to stimuli, respectively (Davidson et al., 1990; Harmon-Jones et al., 2010). Evidence from the psychology and neurophysiology literature points out that frontal EEG asymmetry is associated with different emotional and psychological states. In a seminal work, Davidson et al. (1983) suggested a model, called approach/withdrawal theory, to investigate frontal EEG asymmetry during emotional states. They posited that the left prefrontal cortex (PFC) activity is involved in a system that facilitates approach behavior to positive stimuli, while the right PFC activity participates in a system that facilitates withdrawal behavior from aversive stimuli. This model maintains that processing related to emotional valence itself is not lateralized in the PFC. Rather, emotion-related lateralization is observed because emotions contain approach and/or withdrawal components. Approach/withdrawal motivational states have frequently been linked to asymmetries in left/right frontal cortical activation, especially using EEGs, although meta-analyses of functional magnetic resonance imaging (fMRI) data have failed to find consistent localizations (Kassam et al., 2013).

A number of papers consider our measure of valence—that is, left relative to right frontal cortical activity (LFA)—to build emotion-related (or approach/motivation-related) indexes and correlate them with behavioral or performance outcomes. For instance, Hughes et al. (2014) examined the relation between LFA and effort expenditure for reward, a behavioral index of approach motivation. They found that subjects with greater resting LFA were more willing to expend greater effort in the pursuit of larger rewards, particularly when reward delivery was less likely.

Relative to outcomes in different settings, nascent literature identifies correlations between arousal and valence from EEG recordings and outcomes or behaviors. For instance, Egana-delSol (2016a) used a similar protocol to measure arousal and valence using EEG. He found that EEG features of arousal and valence correlate with an intervention to impact socioemotional skills and resilience for potential entrepreneurs, and with some program outcomes such as enrollment for a high-stakes test to apply to college. Moreover, Marshall et al. (2008) found significant effects of early intervention on EEG alpha power and coherence in previously institutionalized children in Romania. Supplementary analyses examined whether the EEG measures mediated changes in intellectual abilities within the foster care children, but no clear evidence of mediation was observed. Thus, EEG does not directly

relate to changes in cognitive abilities, but, as we argue here, plausibly relates to socioemotional skills, particularly emotional regulation.

We do not expect arousal and valence to correlate with other measures based on self-reported testing. This would imply that we should be able to proxy other socioemotional skills such as grit or locus of control based on EEG features related to emotional state and responsiveness to stimuli. Currently, there is no evidence of this relationship. Moreover, it is important to note that these EEG recordings are not intended to describe and/or predict personality traits or character. For instance, in a recent study, Korjus et al. (2015) showed no correlation between resting-state EEG waves and any of the five personality dimensions of the self-reported Big Five Inventory (BFI, John and Srivastava, 1999), and they concluded that the extraction of personality traits from resting-state EEG power spectra is extremely noisy, if not impossible. In summary, EEGs can relate self-reported tests for noncognitive skills—e.g., BFI or Grit Scale— to transient emotional states during testing, but they do not enable us to predict psychometric test scores from a normal, resting-state EEG recording without stimuli.

To the best of our knowledge, there is no comprehensive analysis regarding either how these indexes of arousal and valence should change across situations, age, etc., or how large the expected impact is. This is also true for many of the self-reported measures typically used in the literature such as Grit, Locus of Control, or Big Five personality tests. In fact, the typical claim is that these tests—and thus, these traits—are stable over the life cycle and across situations. Again, to the best of our knowledge, there is no comprehensive analysis regarding these self-reported measures, either.