Gender-Based Violence in Schools and Girls' Education: Experimental Evidence from Mozambique*

Sofia Amaral
World Bank

Aixa Garcia-Ramos
University of Passau

Selim Gulesci
Trinity College Dublin

Sarita Oré
Teachers College, Columbia

Alejandra Ramos
Trinity College Dublin

Maria Micaela Sviatschi
Princeton University

March 26, 2026

Abstract

School-related gender-based violence (GBV) is pervasive, yet little is known about how to address it or its consequences for education in contexts with low accountability for perpetrators and limited agency among victims. We evaluate a large-scale randomized intervention in Mozambique that combined teacher training with student-focused sessions to strengthen school personnel's capacity to address GBV and to increase students' awareness and reporting of such incidents. To examine differential effects by degree of agency, the student training was randomly targeted to girls, boys, or both. The intervention substantially reduced sexual violence perpetrated by teachers against girls in all treated schools. Using administrative records, we find that when girls received the student training their school enrollment increased. Our evidence shows that these gains were driven by greater willingness of victims to report abuse. Administrative data from the national child helpline shows a rise in GBV-related calls and investigations. These findings highlight that reducing school-related GBV and improving girls' education requires a dual strategy: deterring potential perpetrators through strengthened accountability and empowering victims to report.

*We appreciate the valuable feedback from Janet Currie, Thomas Fujiwara, Ilyana Kuziemko, Eliana La Ferrara, Chris Neilson, Helmut Rainer, Christine Valente, Alex Eble, Ludger Woessmann, Cátia Batista, Pedro Vicente, Alex Armand, Sandra Sequeira, Amber Peterman, Wayne Sandholtz, Joao Montalvão, Eleonora Guarnieri, Girija Borker, Nishith Prakash, Karlijn Morsink, Sara Troiano, Eeshani Kandpal, André Silva, Wei Chang, Sophia Friedson-Ridenour, and seminar participants at the NBER, CEPR, Africa Gender Innovation Lab, KDI-DMI Development Impact Conference, Princeton, Trinity College, Dublin, Warwick University VAW Workshop and Yale University Kuznets Conference. We thank Judite Sambo, Sarita Sinai, and various staff at the Ministry of Education and Human Development of Mozambique for their support. We thank Miguel Jambo, Gizela Samuel, and Nyararai Magudo from Girl Child Rights for the support in the implementation and development of the program. We also thank Gaia Segola and partners from UNICEF and GIZ for their support in the development of the curriculum. This work was supported by J-PAL Post-Primary Education Initiative, Princeton Overdeck, Leibniz Association, UNICEF, GIZ, and Weiss Foundation. The project received IRB approval from the Bioethics Committee of Mozambique, Princeton University, Trinity College Dublin; and was registered at the American Economic Association RCT registry, ID: AEARCTR-0008361.

1 Introduction

Every year, more than 246 million children around the world are subjected to gender-based violence (GBV) in or around schools (UN Women and UNESCO, 2016). While both boys and girls are prone to such violence, girls are more likely to experience the most severe forms of it: sexual violence (Smarrelli et al., 2024). A recent review estimates that one quarter of girls aged 13 to 17 in low-income countries experienced sexual violence in the last year, with 20 percent of those incidents occurring at school (Hares and Smarrelli, 2023). These young adolescents often may lack the maturity to recognize and address abuse which, combined with the fact that violence is enforced by unequal teacher-student power dynamics and limited accountability of perpetrators, makes GBV against girls in schools a threat to their human capital accumulation. Yet, there is limited causal evidence on the effect of GBV in schools on girls' educational attainment and effective strategies to address GBV within the educational setting.

In this paper, we examine how fostering awareness and strengthening support systems for victims and accountability of perpetrators of GBV in schools affect the educational outcomes of adolescent girls at a critical age of transition from upper primary to secondary school. We provide empirical evidence that GBV in schools can be reduced by training designated teachers using a low-cost and scalable approach. We further show that when the teacher-focused training is combined with the encouragement of victims' willingness to report, it can lead to improvements in girls' school enrollment.

Our study takes place in Mozambique, where 7 out of every 10 girls report knowing of cases of sexual harassment cases in their school (Reyes et al., 2023), and yet very few know how to seek support (UNICEF, 2018). To address this issue, in partnership with the Ministry of Education and UNICEF, we designed and evaluated an intervention aimed at strengthening designated teachers' capacity to address GBV in the school and raising student awareness and proactiveness around violence.

The intervention followed a cascade model with two sequential components. First, Gender Focal Points (GFPs) – teachers designated by the Ministry to address gender issues at schools – received training to effectively respond to GBV by protecting victims and holding perpetrators accountable. Trained teachers then delivered four guided sessions to students in grades 6 and 7 (typically ages 13–14). The ses-

sions combined short animated videos with guided discussions. The videos highlighted distinct locally prevalent forms of GBV such as sexual abuse by teachers and peer-to-peer GBV. The discussions after the videos revolved around what actions constitute GBV, why such violence is not acceptable, how to seek support and other examples of proactive behaviors for victims and bystanders.

To evaluate the intervention, we implemented a clustered randomized controlled trial involving 326 primary schools in the province of Sofala. Following an in-school baseline survey of 9,107 students, 239 schools were randomly assigned to receive GFP training, while 88 schools served as controls. Building on the cascade model, we cross-randomized the student training within treated schools: 76 schools trained girls only (Treatment 1), 83 trained boys only (Treatment 2), and 80 trained both girls and boys in mixed-gender sessions (Treatment 3).

Our intervention aims to strengthen awareness and proactivity around GBV at multiple levels of the school hierarchy. In particular, our experimental design allows us to assess how activating different actors within the hierarchy contributes to reducing GBV and improving girls' educational outcomes. Because all treated schools receive the teacher training, similar outcomes across student-treatment arms would suggest that strengthening institutional capacity at the teacher level alone is sufficient to deter violence. In contrast, the variation in the gender composition of the student sessions isolates the contribution of mechanisms operating at the student level. Training girls only (T1) targets victims' proactive behaviors; training boys only (T2) targets bystander engagement and potential peer-level deterrence; mixed-gender sessions (T3) allow us to examine whether collective norm change can operate beyond individual behaviors. Distinguishing these mechanisms is essential for understanding whether reductions in GBV and improvements in schooling outcomes require empowering the most vulnerable actors to act, or whether strengthening institutional capacity and shifting collective norms can, on their own, restore the conditions under which victims safely invest in their education.

A key strength of our study is the combination of multiple, complementary data sources that provide insight into GBV and educational outcomes. To measure the effects of the intervention, we conducted follow-up surveys with students and interviewed the schools' GFPs and teachers. Our first innovation is that for students, we collected information not only on their own experiences of GBV but also on the

experiences of their classmates, enabling third-party reporting. For GFPs, we asked about the actions they had taken to address GBV and complemented this with case-level data from the national GBV helpline, mapped to the administrative units where the schools are located. The system of social investigation provides a unique window to corroborate GFP-reported actions with observed behavior, representing our second innovation. For teachers, we captured their knowledge of GBV laws and expected penalties, providing a measure of deterrence among staff. Third, to investigate the role of GBV on educational outcomes, we combined self-reported school attendance with previously undigitized student registration records, allowing us to track school enrollment objectively. Together, these sources create a uniquely rich dataset that enables us to link intervention exposure, GFP and student behaviors, teacher deterrence, and formal reporting.

Our first main finding is that improving the school's capacity to address GBV substantially reduces the prevalence of teacher-perpetrated violence, regardless of the gender composition of student training. One year after the intervention, drawing on both direct and third-party reports, we find that girls in all treated schools were 67 percent less likely to experience violence by teachers or school staff in the past month. Specifically, girls in treated schools were less likely to report being forced to perform sexual acts by a teacher or school staff compared to girls in the control group.

The decline in sexual violence by school staff in all treatment arms underscores the key role of strengthening school capacity to deter GBV by intervening at the high levels of the school hierarchy. Consistent with this interpretation, we find that GFPs in all treated schools were more proactive. They engaged more frequently in discussions about GBV with students of both genders and with other teachers, and were more diligent in reporting incidents to school authorities.

In contrast to the reductions in teacher-perpetrated abuse, we find no detectable effects on violence perpetrated by students against girls, even if the intervention also targeted peer-to-peer violence. Several factors may explain this result. First, although the intervention improved students' ability to recognize violence, it did not translate into measurable changes in the underlying attitudes toward GBV. Second, GFPs lacked both informal mechanisms and legal instruments to effectively address peer-to-peer violence among students —unlike cases of violence perpetrated

by school staff, which could be formally reported to relevant authorities or school's community leadership— thereby limiting the potential for deterrence. Moreover, GBV by students (e.g., pulling a girl's skirt) often face little sanction, whereas GBV by teachers is met with strong social disapproval. This result is consistent with recent evidence showing that severe forms of GBV are more likely to be mitigated by behavioral interventions (Amaral et al., 2023; Sharma, 2022). Third, dating relationships were relatively uncommon among students, potentially limiting the scope for the intervention to influence behaviors a year after the implementation, but it is still possible that could impact their future relationships.

Our second main finding is that reductions in teacher-perpetrated sexual violence translate into improved schooling outcomes only when girls receive the training. Using administrative records, we find that girls in schools where girls received training were 10 percent more likely to be enrolled at endline relative to the control group. While the effect is positive in schools where only boys were trained, it is small and imprecisely estimated.

We also provide evidence that reporting mechanisms play a central role in reducing teacher-perpetrated sexual violence and fostering educational improvements. Using our novel case-level data from the national GBV helpline, we find that districts with treated schools experienced a surge in calls after the intervention, while calls were similar across control and treated schools prior to the intervention. Moreover, a greater share of calls from treated schools led to formal social work investigations. According to our qualitative findings, these investigations triggered social sanctions by the community and increased teacher awareness of the consequences of their actions. In line with this, survey data from teachers show that staff in treated schools — particularly where girls were trained — were more likely to associate stricter penalties with GBV cases.

Using survey data from the GFPs, we find a corresponding increase in proactive behavior among girls. In schools where only girls were trained, girls were more likely to approach GFPs to report violence. GFPs' awareness and use of the national helpline was significantly higher in these schools, indicating active engagement by students in initiating formal complaints. Supporting this interpretation, GFPs in these schools were more likely to report GBV cases to the school council, which involves parents, community leaders, and principals.

We conduct several robustness checks to ensure our results are not driven by reporting bias or experimenter demand effects. First, when contrasting self-reported violence with third-party reported measures, we find similar effects. Second, following recent advancements in the literature (Bursztyn et al., 2025), we test whether social desirability bias drives our estimates by interacting treatment indicators with baseline social desirability scores and we find no significant heterogeneity. Third, we assess changes in reporting of pre-treatment episodes of violence, showing that students' reporting of GBV prior to 2021 was similar across treatment and control schools, ruling out recall bias. Moreover, the absence of effects on dating violence or attitudes towards GBV, which the intervention also targeted, provides additional evidence that our results are not driven by experimenter demand.

Taken together, our findings suggest that reducing GBV in contexts with asymmetrical power –such as schools– improves educational outcomes only when both top-level deterrence and bottom-up proactivity operate together. In teacher-perpetrated sexual violence, proactive engagement by GFPs deters perpetrators, but these gains translate into better schooling outcomes only if girls themselves report incidents, enabling school authorities to respond and potentially rebalance the teacher-student power dynamic. This parallels the accountability literature (e.g., Ferraz and Finan (2025)), which shows that in settings with limited enforcement capacity, top-down oversight alone is insufficient without complementary bottom-up engagement. Extending this framework to GBV, we show that credible deterrence and victim agency must operate together to allow girls to reinvest in their human capital.

Our findings are also consistent with economic models of crime and social interactions. Glaeser et al. (1996) show that crime responds not only to individual incentives but also to the strength of social interactions and sanctions. When sanctions are weak, minor infractions persist; when community disapproval and social penalties are strong, deterrence multipliers arise. Schools operate under similar dynamics: hierarchical and socially embedded environments where informal norms shape compliance with formal rules. This perspective helps explain why our intervention reduces teacher-perpetrated GBV, which is subject to strong institutional and social sanctions, while leaving peer-to-peer violence, where social sanctions are weaker, largely unchanged.

Our study also connects to research on GBV in workplace settings (e.g. Boudreau

et al., 2023; Adams-Prassl et al., 2023), where asymmetrical power is central and combining top-down enforcement with peer-level mechanisms reduces abuse and improves labor outcomes. In schools, we find a parallel pattern: teacher-perpetrated GBV can be deterred by strengthening the capacity of key personnel (GFPs), while improvements in girls' educational outcomes require activating proactive behaviors among victims. This highlights the effectiveness of interventions targeting violence at multiple institutional levels to protect vulnerable agents.

Our results contribute to the growing literature on the economic and social consequences of GBV (e.g. Borker, 2022; Amaral et al., 2023; Sanin, 2021; Siddique, 2022; Chakraborty et al., 2018; Garlick et al., 2025). Closely related to Adams-Prassl et al. (2024), who document large and persistent effects of sexual violence on employment and mental health in Finland, we quantify how sexual abuse at school disrupts girls' human capital accumulation by increasing dropout and widening gender gaps in education. While this literature highlights the substantial economic costs of GBV, our findings point to a complementary insight: effective prevention is both feasible and affordable. Our intervention is low-cost, scalable, and implementable in resource-constrained education systems, demonstrating that multi-level, proactive strategies can mitigate GBV and enhance economic outcomes. Our paper also relates to studies on interventions to reduce different forms of school-based violence (e.g Devries et al., 2015; Karmaliani et al., 2020; Gutierrez et al., 2018; Smarelli, 2023). We contribute by studying an intervention that specifically targets GBV in schools and by providing causal evidence on how reducing school-related GBV affects girls' schooling.

We also contribute to recent studies that evaluate the effectiveness of interventions aimed at empowering girls and/or women, specifically examining the impact of targeting girls versus boys (e.g Andrew et al., 2022; Cassidy et al., 2023; Shah et al., 2023; Edmonds et al., 2023). We contribute to this literature by demonstrating how, also in the school context, there are gender specific complementarities between interventions inducting changes at the individual level and interventions targeting violence at higher level of aggregation such as the schools. In particular, our results align with Andrew et al. (2022), who show that shifts in the behavior of key actors responsible for enforcing gender-related norms within communities are crucial for improving women's well-being. We complement this work by demonstrating that

an analogous shift among key enforcers at the school level is essential for reducing sexual violence against girls and improving their educational outcomes.

The remainder of the paper is organized as follows. Section 2 describes the study context and intervention. Section 3 outlines the experimental design, and the conceptual framework. Section 4 details the data sources and measurement approach. Section 5 presents the empirical results, along with a cost-effective analysis in Section 6. Section 7 concludes.

2 Intervention

Over the past two decades, Mozambique has achieved remarkable progress in expanding access to primary education. Gross (net) enrollment rates now exceed 100 (90) percent for both boys and girls, reflecting near-universal completion of primary school.[1] However, these achievements sharply decline at the secondary level: only about one in three adolescents is enrolled in secondary school, with girls disproportionately dropping out during the transition from upper primary to lower secondary education.

Gender-based violence (GBV) is increasingly recognized by Mozambican institutions and NGOs as a key factor contributing to girls' early school exit. According to a national survey by UNICEF (2018), roughly 60 percent of secondary school students report awareness of sexual violence in their schools, and more than half identify teachers as the main perpetrators. Yet, only a small fraction of students are aware of mechanisms to report such incidents or seek support.

For girls in particular, abuse by male teachers often involves coercion into sexual favors in exchange for better marks or exam passage—a practice widely known as "*sex for grades*". Such abuse can lead to early pregnancy and school dropout, further widening gender gaps in education. Teachers, however, are not the only source of risk: roughly one in five cases of sexual violence against young women before age 18 is perpetrated by friends, classmates, or other students (Reyes et al., 2023).

Given girls' vulnerability within the school system, we partnered with the Ministry of Education and UNICEF to design and implement a school-based curriculum

[1]Stat Bulk Data Download Service, UN Educational, Scientific and Cultural Organization (UNESCO).

aimed at addressing GBV at multiple levels of the school hierarchy: *Está na Hora de Agir* ("It's Time to Act").

2.1 The Curriculum: *Está na Hora de Agir*

Está na Hora de Agir ("It's Time to Act") is an intervention that promotes awareness and proactive behaviors around GBV by examining violence through the lens of victims, bystanders, and perpetrators. The overarching objective is to build the understanding that anyone within the school community can take proactive steps against GBV, to help participants identify and engage in safe and appropriate behaviors, and to strengthen the accountability of perpetrators.

Although the curriculum is based on the content of the Ministry of Education's official textbooks, it delivers these materials through interactive and engaging activities designed to enhance comprehension and retention. The intervention follows a cascade model with two sequential components: the '*Gender Focal Point (GFP) training*' and the '*Student Training*'.

Gender Focal Points Training

Gender Focal Points (GFPs) are teachers appointed by the Ministry of Education to address gender-related issues within schools. This role is an established feature of the Mozambican education system that predates our intervention. Upon appointment, GFPs continue their regular teaching duties and receive no specialized training tailored to their role.

The GFP training component of *Está na Hora de Agir* sought to strengthen teachers' knowledge of GBV and their capacity to respond to GBV incidents within their schools. The two-day training combined interactive sessions and applied exercises on the definition and forms of GBV, how it differs from other types of violence, victim support mechanisms, and the specific responsibilities of GFPs in preventing and addressing GBV (Figure A.1). GFPs were also trained to deliver the student component of *Está na Hora de Agir*.

Throughout the training, strong emphasis was placed on reporting procedures and the activation of a response system to support victims and hold perpetrators accountable. Following UNICEF's referral protocol (see Figure A.2), GFPs were in-

structed to report each identified case of GBV through two complementary reporting pathways: (i) the national toll-free children's helpline, 116 *Linha Fala Criança* (LFC),[2] and (ii) the local school authorities, such as the school council and school director.

These two pathways operate simultaneously and complementarily: while the LFC initiates a formal investigation, school authorities can take immediate preventive measures within the institution. However, the activation of formal proceedings requires the victim's consent to move forward with an official report.

Student Training

The student training component aimed to strengthen awareness of GBV and encourage proactive behaviors among students—whether as victims, bystanders, or potential perpetrators. The program consisted of four two-hour sessions in which students watched short animated videos, participated in reflection activities, and engaged in guided group discussions focused on GBV.

Two animated videos were specifically developed for the intervention, addressing sexual abuse by teachers and dating violence.[3] The first video depicted sexual harassment in the school setting, showing a male teacher inappropriately touching and harassing a female student. It then illustrated how the victim sought support from friends, family, and the GFP, and concluded with practical guidance on how students can seek help (Figure A.3). The second video focused on dating violence, portraying scenarios of verbal and physical aggression between adolescent students. For instance, one scene showed a male student pushing his girlfriend after losing a football game; another showed a male student shouting angrily at a female peer (Figure A.4).

After each video, the GFP guided discussions on the types and consequences of violence, why such behavior is unacceptable, and how students can seek support or engage in proactive bystander actions. Each session concluded with a reflection

[2]Upon receiving a call, the helpline conducts a preliminary investigation to gather evidence before referring the case to formal institutions (e.g., police, social services). LFC also maintains a registry of trained professionals and NGOs equipped to provide local assistance to victims. For a detailed description of the GBV referral protocol, see Appendix A.

[3]The use of animated videos was motivated by growing evidence that video-based edutainment can effectively influence social norms and behaviors related to fertility, domestic violence, and early marriage (La Ferrara et al., 2012; Banerjee et al., 2019a,b; Green et al., 2020; Cassidy et al., 2023). The videos also ensured uniformity in the delivery of the curriculum across schools.

exercise or activity to reinforce learning and prepare students for the following session.

2.2 Implementation

The intervention was implemented by the local NGO Girl Child Rights (GCR), in collaboration with the Ministry of Education and UNICEF in Mozambique.[4] GCR was responsible for conducting the GFP training and ensuring that all necessary equipment (e.g., projectors and video materials) for the student sessions was available on site. Following a cascade model, once trained, GFPs took full responsibility for implementing the student training within their schools.

The GFP training was delivered over two days, complemented by one-on-one review sessions and on-site technical support from GCR during the student sessions. The implementing partners covered all participation costs, including travel, lodging, and daily allowances. Attendance was 98.9 percent: 236 of the 239 GFPs from treatment schools completed the training.

During the GFP training, each participant received a printed implementation manual detailing step-by-step instructions for conducting the student sessions. GFPs also received distinctive t-shirts to wear during those sessions, which made them easily identifiable to students and allowed monitoring of implementation compliance. Specifically, GFPs were instructed to wear their t-shirts during every student session and to document each session with dated photographs taken on school grounds. When a video was shown, the projector also had to be visible in the picture. All photos were timestamped and transmitted to the research team through GCR, allowing us to verify adherence to the intervention protocol.

The student sessions were conducted during regular school hours and on school premises. As shown in Figure A.5, attendance averaged around 85 percent and was similar across study arms.

[4]GCR is a leading Mozambican NGO working on youth and female empowerment and children's rights. Operating since 2008, it has extensive experience engaging communities and schools to raise awareness about GBV and child protection. For more information, see `https://gcr.org.mz`. GCR also manages Mozambique's national toll-free children's helpline, *Linha Fala Criança* (116), which provides a platform for reporting violence and offering victim support.

3 Experimental Design

To evaluate the intervention, we implemented a clustered randomized controlled trial in the province of Sofala, covering seven of its thirteen districts: Beira, Dondo, Buzi, Nhamatanda, Chibabava, Cheringoma, and Maringue. These districts were selected based on accessibility and security considerations. We aimed to include the universe of primary schools within these districts; however, due to remoteness and exposure to climatic disruptions such as cyclones and floods, 326 out of 340 schools (96 percent) were ultimately included in the study.

Randomization was stratified by district and by the baseline prevalence of school violence. This measure was constructed using students' self-reported experiences of school-related violence in the month preceding the baseline survey. Within each district, schools with a prevalence at or above the district median were classified as "high-violence." Within each stratum, schools were randomly assigned to one of four groups: Treatment 1 (T1), Treatment 2 (T2), Treatment 3 (T3), or Control.

All treatment schools (T1-T3) received the GFP (teacher) training, while the student training was cross-randomized across treatment arms. Specifically, in T1, only girls participated in the student sessions; in T2, only boys participated; and in T3, both girls and boys were trained together in mixed-gender sessions.[5] Schools in the control group received neither the GFP training nor the student training. All students participating in the training were 6th- and 7th-graders interviewed at baseline. Figure 1 and 2 summarize the experimental design and show the geographic distribution of schools and their treatment assignments.

The logic behind the cross-randomization was to target genders differently, reflecting their distinct roles in GBV—as potential victims, bystanders, and peer-level perpetrators. By targeting girls only (T1), we aimed to assess whether reductions in violence and improvements in schooling outcomes were driven by victims' proactive behaviors—that is, their ability to recognize abuse, seek support, and report incidents when they occur. By targeting boys only (T2), we aimed to capture changes in bystander engagement and attitudes toward GBV among potential future perpetrators. In this context, boys play a dual role: they can act as bystanders who protect girls from violence, but they can also be potential future perpetrators whose gen-

[5] Attendance patterns demonstrate high compliance with the intervention protocol, with no significant gender differences in attendance in T3 (Table A.3).

der norms can be shaped to reduce the risk of violence in their future relationships. Targeting both girls and boys together (T3) enables us to examine whether joint exposure to the same messages can promote collective norm change, not only by changing what constitutes acceptable behavior and strengthening mutual accountability among peers, but also by fostering greater accountability from below toward teacher-perpetrators.

Taken together, by combining the GFP training with different forms of student training, our design allows us to explore potential complementarities between actions at the top and bottom of the school hierarchy. This enables us to test whether empowering victims and/or bystanders amplifies the effects of institutional strengthening, thereby achieving larger reductions in GBV and improvements in girls' educational outcomes, or whether broader institutional and collective norm changes suffice to restore victims' ability to safely invest in their education.

3.1 Conceptual framework

We conceptualize GBV in schools as an abuse of authority within a hierarchical organization characterized by asymmetric power and weak enforcement. In this structure, teachers and students occupy distinct positions in the hierarchy, yet both can act as perpetrators of violence when sanctions are weak. The persistence of GBV thus reflects a failure of accountability mechanisms to deter misconduct. Our experimental design activates accountability at multiple levels of the hierarchy, allowing us to study how distinct channels interact.

From the perspective of teacher-perpetrators, accountability operates primarily through vertical and horizontal mechanisms that ensure both institutional oversight and credible deterrence. The GFP training enhances horizontal accountability among teachers by improving their ability to identify and respond to misconduct among peers. It also links schools to formal channels of vertical accountability through the school council and the national helpline, thereby increasing the perceived probability that abuse will be sanctioned. Greater reporting by victims and bystanders further strengthens bottom-up accountability, raising the likelihood that teacher misconduct is detected and amplifying deterrence through stronger social sanctioning.

From the perspective of student-perpetrators, accountability operates primarily through horizontal accountability and social sanctions. The student training seeks to

alter the expected costs of engaging in violence. Among girls, the training enhances awareness of what constitutes abuse and strengthens victims' willingness to report it, increasing the perceived likelihood of detection (bottom-up accountability). Among boys, the intervention promotes bystander engagement and shifts social norms to discourage peer or dating violence (horizontal accountability). The mixed-gender training fosters collective norm change, raising the social cost of abusive behavior within the student body (cross-level accountability).

4 Data

Our analysis draws on multiple complementary data sources collected before and after the implementation of the intervention, which took place during the 2021–2022 academic year amid COVID-19 disruptions. We combine detailed student survey data on GBV experiences, perceptions, and attitudes; teacher surveys capturing their knowledge of GBV laws and perceived sanctions; Gender Focal Point (GFP) surveys measuring proactive behaviors and knowledge in addressing GBV cases; administrative school records that we digitized to track enrollment and transitions to secondary education; and case-level administrative data from the national helpline *Linha Fala Criança* (LFC), which documents formal reports and social work investigations of GBV. Together, these sources allow us to map changes in violence, reporting, and deterrence to girls' educational trajectories.

4.1 Student Survey Data

We begin our analysis with panel survey data from students, which we later complement with administrative data. The baseline sample includes 9,107 students —4,605 boys and 4,502 girls— enrolled in upper-primary schools in the Sofala province. The baseline survey was conducted to student in 6-7th grades between May and September 2021. In treated schools, these very same students received the student training.

We followed respondents roughly one year after the intervention. To maximize tracking, the endline survey was implemented in two waves. The first wave, conducted between June and November 2022, reinterviewed students in all baseline primary schools and in the full universe of 71 secondary schools across the seven

study districts. This strategy ensured coverage of both students who remained in primary school and those who had transitioned to secondary education. In total, wave one followed 6,401 students—approximately 70 percent of the baseline sample. The second wave, conducted between September and October 2023, followed up with students who could not be reached earlier, primarily through household visits. Across both waves, we successfully tracked 83 percent of the baseline sample (82 percent among girls and 84 percent among boys). Attrition rates were balanced across treatment and control groups (see Table A.2).[6] Figure 3 summarizes the project timeline.

Interviews were conducted on school premises during regular school hours at baseline and either at school or at respondents' homes at endline. All interviews were administered face-to-face on tablets by enumerators of the same sex as the respondent. Because many students had limited literacy, interviews were conducted by trained enumerators rather than self-completed. Enumerators were trained in interviewing minors on sensitive topics and in crisis response management following World Health Organization (WHO) guidelines for research on Violence Against Women. All interviews were conducted under conditions that ensured privacy. Before each interview, enumerators obtained written parental consent and child assent. At the end of every interview, respondents were informed about available support services. All participants received information about the national child helpline *Linha Fala Criança*, regardless of whether they disclosed violence, and those who reported an incident were provided with the contact details of the school's Gender Focal Point (GFP) for follow-up.[7]

The student questionnaire lasted about one hour and gathered detailed information on students' backgrounds, their own experiences with gender-based violence (GBV), their observations of violence against female peers, and their views and attitudes toward such behaviors. Students' experiences of GBV were measured using a

[6]While all treatment arms had similar attrition relative to control, there was slightly more attrition among girls in T2 schools relative to T1 and T3 schools. We construct attrition bounds to assess if this differential attrition may be driving our findings and find that it does not – see Appendix Tables A.8 and A.9.

[7]The study was approved by the National Bioethics Committee of Mozambique, received IRB approval from Princeton University, and ethical clearance from Trinity College Dublin. Appendix A. provides a detailed account of the mitigation and safeguarding strategies implemented during data collection.

module adapted from the World Health Organization's Violence Against Women instrument (Ellsberg et al., 2005). The questionnaire listed behavioral characterizations of violent episodes across emotional, physical, and sexual forms of GBV, allowing respondents to disclose multiple perpetrators and specify when each episode occurred (see Appendix Section E.I). To complement these self-reports, students were also asked about violence experienced by other girls in their class, as respondents may find it easier to report indirectly than to disclose personal experiences. Students' ability to identify acts of violence was assessed through hypothetical vignettes depicting GBV in school settings, while a separate set of questions measured the norms and acceptability of GBV. Motivated by the importance of sexual abuse by teachers in the local context, the survey also included a module on transactional sex, capturing coercive or exchange-based interactions that might not emerge under standard GBV measures. At endline, students additionally completed standardized reading and mathematics assessments based on the Early Grade Reading and Mathematics Assessments (EGRA and EGMA), adapted for the Mozambican context by Chimbutane et al. (2022).

Given the sensitive nature of GBV, we adopted several measures to minimize reporting bias. First, we deliberately incorporated third-party reporting questions that elicit indirect accounts of violence while reducing the pressure of self-disclosure.[8] Next, we anchored students' self-reports of GBV to a point in time prior to the intervention by asking whether the incidents occurred before 2021. This allows us to detect changes in reporting behavior over time and across treatment arms. Finally, following Dhar et al. (2022), we administered a short social desirability scale based on the Crowne and Marlowe (1960) module to identify respondents more likely to provide socially acceptable answers, which we use to test for heterogeneous effects.

Table 1 presents descriptive statistics for girls in our sample at baseline, by treatment group. The average respondent was 13.5 years old. About 40 percent of mothers and 25 percent of fathers had no formal education. Roughly one in twenty girls reported ever having a romantic partner, indicating that dating was not a prevalent phenomenon among our sample at the time of the intervention. At baseline,

[8]Third-party reporting (TPR) methods are widely used to measure sensitive behaviors such as abortion or intimate partner violence when direct self-reports may be biased. Our TPR measure follows the "anonymous third-party reporting" approach described in Giorgio et al. (2021) and Owolabi et al. (2023), adapted to the school context.

29 percent of girls reported experiencing violence by another student in the past month, and 6 percent reported violence perpetrated by teachers or school staff. Peer-reported violence against other girls shows similar magnitudes. Emotional violence was the most common form (36 percent), followed by physical (25 percent) and sexual violence (16 percent). Across all variables, baseline means are balanced between treatment and control schools with less than 10 percent of pairwise comparisons rejecting equality, indicating that randomization achieved balance.[9]

4.2 Student Administrative Data

We complement student survey data with administrative school records from both primary and secondary schools in the study districts. Because students often transfer when progressing to secondary education, our endline tracking extended beyond baseline schools. Moreover, schools in our study area do not consistently assign unique student identifiers. To link survey respondents to school records, we digitized enrollment lists from all baseline primary schools and from the universe of secondary schools across the seven study districts, matching students by name and expected grade (6, 7, or 8). Some students were absent from official rosters yet located during field visits. We therefore built a harmonized student roster, treating enrollment as confirmed either through digitized records or verified presence during school visits. This hybrid dataset allows us to track students' educational trajectories across schools and to link experiences of GBV and reporting to school enrollment outcomes.

4.3 GFP Survey Data

To assess changes in awareness and proactive behaviors among school-level Gender Focal Points (GFPs), we conducted an endline survey between July and October 2023.[10] The questionnaire captured GFPs' engagement with students and teachers,

[9]Table A.1 shows additional baseline descriptive statistics at the school level. Overall, school characteristics seem well-balanced, with fewer than 10% of the tests showing significant differences.

[10]We were not able to collect baseline information on the GFPs. In Table A.4, we use GFPs' predetermined characteristics that should not have been affected by the intervention (e.g., age, gender, place of birth, schooling level) to test for balance in covariates. Fewer than 10 percent of differences are statistically significant, confirming that GFPs in control schools constitute a valid counterfactual for those in treated schools.

their actions to report GBV cases, and their coordination with the school council and the national helpline *Linha Fala Criança* (LFC). These indicators provide a comprehensive view of GFPs' capacity to respond to GBV within schools and their dual role in strengthening horizontal accountability and activating channels that enforce vertical accountability and external oversight.

4.4 Teacher Survey Data

To capture teachers' perceptions of the probability and severity of sanctions for GBV, we administered a short teacher survey between June and September 2022, concurrently with the first wave of the student endline. We aimed to interview four teachers per school, prioritizing those who taught grades 6 or 7 during the 2021–2022 school year.[11] Teachers were asked whether they knew what penalty the Penal Code prescribes for GBV-related offenses and, if so, how long they believed the corresponding prison sentence to be. These questions provide a direct measure of perceived penalties from the perspective of potential teacher-perpetrators.

4.5 *Linha Fala Criança* Administrative Data

To capture formal reporting and institutional responses to GBV beyond the school level, we obtained case-level administrative data from Mozambique's national child helpline, covering the period 2020–2023. The dataset records all child protection calls, including those reporting violence in or around schools. For each call, we observe the date, district of origin, the caller's age, and the nature of the issue reported. We also obtained information on all social work investigation cases initiated by LFC during the same period.[12] [13] These data provide a measure of external oversight in the GBV reporting chain.

Together, these data sources provide a comprehensive view of gender-based violence against girls within schools. The student survey captures changes in aware-

[11] When fewer were available, teachers from other grades or subjects were substituted following a standardized protocol.

[12] Due to data confidentiality restrictions, we only observe the caller's district and no other identifying information. Therefore, we aggregate reports to the district level.

[13] Interviews with LFC staff indicate that most calls concern serious GBV cases, particularly sexual abuse by teachers or older students.

ness, attitudes, and experiences of GBV at the lower levels of the school hierarchy; while the teacher survey reflects deterrence among potential perpetrators at higher levels. The GFP survey measures increased awareness and proactive engagement among key school personnel. The LFC data captures external oversight beyond the school. Combined with verified school enrollment records, these sources allow us to map changes in violence, deterrence, and reporting to girls' educational outcomes — a key outcome motivating this study.

5 Results

5.1 Estimation

To assess the effects of the intervention, we estimate:

$$Y_{icd} = \alpha + \sum_{k=1}^{3} \beta_k T_{cd}^k + Y_{icd}^0 + \theta_d + \gamma X_{icd} + \epsilon_{icd}, \tag{1}$$

where Y_{icd} is the outcome of interest at endline for respondent i who was attending school c in strata d at baseline, T_{cd}^k is an indicator equal to 1 if school c was assigned to treatment group k (where $k = 1, 2, 3$) and 0 otherwise, θ_d are randomization strata fixed effects (i.e., district $\times$ high-GBV dummies), Y_{icd}^0 is the baseline level of the outcome variable. Whenever the outcome variable is self-reported, we control for X_{icd} – respondent's social desirability score as measured at the baseline survey. Under the identifying assumption that the control group forms a valid counterfactual for the treatment groups, β_k provides the causal effect of the intervention when only girls ($k = 1$), only boys ($k = 2$), or both genders ($k = 3$) were included in the student training.

Since the randomization was conducted at the school level, we cluster the standard errors by school. To take into account multiple hypothesis testing, we group outcomes that test the same hypothesis in families and correct the p-values using the procedure proposed by Benjamini et al. (2006). This allows us to control the *false discovery rate* within families of outcome variables. We correct the p-values by treatment arm and group the outcomes into families as specified in the table notes.[14]

[14]In Appendix C. we describe the three minor deviations from the pre-analysis plan.

5.2 Students' Awareness and Proactive Behaviors

We begin by examining whether the intervention activated the awareness and proactive behaviors it was designed to foster among students – key conditions for accountability to emerge within schools. In hierarchical systems, accountability depends on the ability and willingness of actors at lower levels to recognize and contest abuses of authority. When these actors lack the cognitive tools to identify violence, or when prevailing norms justify abuse, no restorative action is likely to occur. Students who cannot classify episodes of gender-based violence (GBV) as violence, or who, even when recognizing the harm, perceive such acts as acceptable, are unlikely to seek help, report abuse, or contribute to restoring a safe school environment.

To assess whether the intervention improved students' recognition of GBV, Table 2 reports treatment effects on students' ability to recognize GBV. We measure this skill through correct classification of short vignettes depicting GBV situations (column 1) and the correct identification of seven statements—four describing GBV acts and three non-GBV acts (columns 2–3).[15] Students in schools receiving gender-segregated training (T1, girls-only, and T2, boys-only) improved their ability to identify GBV by roughly 15 percent relative to the control mean of 22.5 percent (3.9 percentage points for T1 and 3.0 percentage points for T2). No statistically significant improvement is observed in mixed-gender schools (T3). This pattern is consistent with our qualitative evidence indicating that discussions of violence were more constrained by gender dynamics in mixed sessions.[16]

In Table 3, we next examine whether the intervention influenced students' attitudes toward GBV. We find no measurable changes in the acceptability of violence across treatment arms, suggesting that the intervention did not foster a collective norm change, at least not at the lower levels of the school hierarchy.

While the student training increased awareness, recognizing violence alone might be insufficient to stop it. A central component of the intervention was therefore to encourage students to seek support and report incidents. In particular, harmful behaviors can only be addressed when recognition is accompanied by the willingness of actors at lower tiers of the hierarchy to seek help and report abuse. In the school

[15] Vignettes and items are described in detail in Appendix E.II.

[16] We also examined treatment heterogeneity by gender and find no systematic differences across groups, although point estimates are slightly larger for boys in some specifications.

context, students who act upon what they observe exert bottom-up accountability, enabling higher levels of the hierarchy to respond. With this in mind, we examine whether students' gains in awareness were accompanied by greater reporting of GBV cases to GFPs. Table 4, column (1), shows that the probability of a GFP declaring that at least one student reported a case of GBV is roughly four times higher in schools where only girls were treated. In contrast, training boys did not increase the reporting of GBV to the GFP. Our interpretation of these findings is twofold. First, the intervention was not successful in activating boys' bystander behavior.[17] Second, when girls themselves receive the training, they are empowered to report.

5.3 GFPs: Accountability Agents within Schools

What did the GFPs do with these reports? During the GFP training, GFPs were assigned multiple tasks reflecting different dimensions of accountability. They were instructed to exert vertical accountability by addressing potential student-perpetrators. They were instructed to engage in horizontal accountability by discussing GBV issues with other teachers. They were instructed to trigger vertical accountability by escalating teacher-perpetrated cases to the school council and school authorities who could impose sanctions. Finally, GFPs were trained to activate external oversight by reporting cases to the *Linha Fala Criança* (LFC) helpline. Table 4 presents the effects of the intervention on these GFP activities.

Vertical Accountability: Column (2) shows that across all treatment arms, GFPs were significantly more likely to discuss GBV-related topics with students, indicating that the training translated into concrete engagement at the student level. Such engagement serves a dual function: it signals that GBV concerns are taken seriously to potential student-perpetrators and provides a clear pathway for victims to report abuse. In this setting, accountability mainly operates through visibility, information, and support—both by raising the perceived social cost of misconduct and by lowering the barriers for victims to come forward.

Horizontal Accountability: Column (3) of Table 4 shows that GFPs in T1 schools were 15 percentage points more likely to have discussed GBV or shared related materials with teachers. This suggests that, when girls received the student training,

[17]We also examine whether the increase in reporting is driven by victims or by bystanders and find that most of the effects are driven by victims. Detailed results are available upon request.

GFPs intensified peer-to-peer accountability within the teaching staff. By increasing the salience of GBV norms and associated penalties, GFPs helped reinforce expectations of appropriate conduct within the teaching body. The fact that peer scrutiny is stronger when girls are trained underscores the role of victims' reporting in amplifying the perceived costs of misconduct at higher levels of the school hierarchy.

Escalating to school authorities: Column (4) shows that GFPs in all treatment arms were significantly more likely to report GBV cases to school authorities, with increases ranging from 9 to 12 percentage points relative to the control mean of 8 percent. Thus, the intervention successfully triggered vertical accountability within schools.

In column (5) we also show that GFPs in schools where only girls were trained (T1) were more likely to have discussed GBV cases with the school council – a government body which includes parents and community leaders. The effect corresponds to a 16 percentage point increase relative to the control group, where only 38 percent of GFPs reported such discussions. Once again, this indicates that when girls are trained, GFPs are not only more likely to escalate cases to school authorities but also to extend accountability outward to the community. According to our interviews with GFPs, reporting to the school council generally occurs when a victim report is available, since council meetings involve parents and community leaders. When GFPs only suspected a case, they typically informed school authorities or sought guidance from the LFC helpline.

External Oversight: Finally, we assess GFPs' knowledge of the national child-protection helpline (LFC). We focus on two indicators: whether the GFP reported knowing about the helpline and whether they could correctly recall its exact phone number (116). Column (6) shows that GFPs in all treatment schools were more likely to declare knowing about the helpline. However, this measure may reflect experimental-demand effects, as the existence of the helpline was repeatedly discussed during the GFP training. By contrast, the ability to recall the number – reported in column (7)–provides a more credible indicator of effective knowledge and engagement with the LFC.

In the absence of the intervention, only 15 percent of GFPs could recall the correct number, indicating a striking gap in basic knowledge among untrained staff. In T1 schools, GFPs were 24 percentage points more likely to recall the number (a 160

percent increase relative to the control mean). The effects in T2 and T3 schools were smaller, at 12 and 10 percentage points and less precisely estimated. Taken together, these findings show that the intervention strengthened GFPs' capacity to activate external oversight and support victims.[18]

5.4 Verification Through Administrative Helpline Records

To validate the survey-based evidence on GFP engagement with the 116 helpline and test for any increase in reporting to authorities, we draw on administrative data from the *Linha Fala Criança*. We examine whether call volumes and subsequent case investigations increased in districts exposed to the intervention relative to neighboring districts before and after program implementation. This is because *Linha Fala Criança* does not record the school where the call is coming from, but only the district of the call.

Our sample covers calls from 20 districts (7 treated plus 13 control), observed over the course of 20 bi-months.[19] Using this data, we estimate a generalized difference-in-difference model, comparing the number of calls across districts served by the intervention and neighboring districts in the period before and after the start of the intervention. In particular, we estimate:

$$Y_{dt} = \alpha + \beta_1 T_d \times Post_t + \beta_2 T_d + \beta_3 Post_t + \theta_d + \gamma_t + \mu_d t + \epsilon_{bt}, \tag{2}$$

where $Post_t$ is a dummy that takes the value 1 for the bi-months after the month of August minus the start of the intervention in every district; T_d is a dummy variable that takes the value 1 if a call comes from a district where the intervention took place, and 0 for calls coming from neighboring districts. The specification includes district fixed-effects, bi-month dummies, district-bimonthly linear trends, and bootstrapped standard errors are clustered by district.

Figure 4 presents the estimates for β_1 for the pre and post periods in specification (2) and the 95% confidence intervals. Panel (b) and (c) show that following the implementation of the intervention, treated districts experienced an increase in calls

[18]Qualitative interviews with teachers and GFPs corroborate this interpretation, showing that stronger student engagement—particularly among girls—elevated the salience and use of formal reporting channels (Appendix F.).

[19]We aggregate the data at district for every two months to minimize noise in the time-series.

seeking for assistance and in calls that translated into formal investigation cases. Such investigations typically require that social workers engage with schools, family, victims, and relevant legal authorities to provide adequate psychosocial, health, and legal support to the victim. Looking at LFC annual reports, we find that 80 percent of the calls to the helpline become cases where there is a formal investigation involving the school and the community.[20] As a result, communities – and perpetrators – are also more likely to become aware of the prevailing sanctions associated with violence against minors.

5.5 Deterrence: Teachers' Expected Sanctions

Consistent with the documented increase in formal oversight, we examine whether the intervention shifted teachers' expectations regarding legal sanctions around GBV. Drawing on our teacher survey, we find that teachers in treated schools —particularly where girls received the student training— report greater awareness of GBV laws and associate stricter penalties with sexual acts involving minors (Table 5, columns 1–2). As a falsification test, columns 3–5 of Table 5 show no effects on teachers' knowledge of laws unrelated to GBV, supporting interpretation of these results as a targeted shift in expected sanctions rather than a generalized survey-response pattern or broad increase in legal awareness.

5.6 Violence Against Girls: Prevalence Effects

Having documented changes in students' awareness, GFP proactivity, external oversight by the LFC, and teachers' expectations of sanctions, we next assess whether these shifts were accompanied by reductions in gender-based violence (GBV) against girls in schools, the central outcome of interest. We measure GBV using both girls' self-reported experiences and third-party reports from classmates on violence against girls in their classroom.

Table 6 examines incidents of GBV occurring in the month before the endline survey, approximately one year after the intervention. We distinguish violence per-

[20]This mechanism is in line with evidence from Uganda in the context of intimate-partner violence among adults. Green et al. (2020) show that awareness of the costs of IPV, not only changes the reporting behavior of victims but also increases the perceived social sanctions of others in the community.

petrated by students (columns 1–2) from violence perpetrated by teachers or school staff (columns 3–4). Columns 1 and 3 show girls' self-reported incidents, while columns 2 and 4 capture bystander reports, asking all respondents about their observations of violence against the girls in their class.

Estimates in Table 6 show that the intervention did not significantly affect violence against girls perpetrated by students, but did reduce violence committed by teachers or school staff. In column 3, girls in T1 schools are 0.8 percentage points less likely to report experiencing such violence in the past month. Given that 1.2 percent of girls in the control group reported violence by teachers or staff, this represents a 67 percent reduction in prevalence. Point estimates for T2 and T3 schools are also negative, though less precisely estimated, and not statistically different from T1.

The similarity of the effects in teacher-perpetrated GBV across treatment arms suggests that strengthening institutional capacity through GFP training was central to deterring teacher-perpetrated violence. Yet, the intervention did not reduce peer-to-peer GBV. This result aligns with our null effects on students' attitudes toward GBV and with the documented type of vertical accountability that GFPs can exert over students (mainly through visibility and information rather than sanctions).[21] Moreover, peer-to-peer GBV in this setting often involves milder forms of misconduct relative to abuse by teachers, which is typically of sexual nature – consistent with this, we find reductions in forced sexual violence by teachers across the treatment arms (see Table A.7). Finally, adolescents in this context typically begin dating at ages 16–17, while students in our sample were 13–14 at the time of the intervention. As a result, even if we do not observe immediate effects on dating-related violence, impacts may emerge as in the future.

Together, these results indicate that while the intervention activated the accountability chain for adult perpetrators, the absence of formal enforcement mechanisms for student misconduct limited its ability to reduce peer-to-peer violence in the short run. While we do not have data on the type of incidents reported to the helpline, qualitative evidence suggests that most reports concerned violence against girls perpetrated by adults. Combined with the lack of changes in GBV attitudes, this may explain why peer-perpetrated violence remained unaffected. In our final qualitative

[21] Qualitative interviews indicate that GFPs lack formal mechanisms to address student-perpetrated GBV beyond attempting to change attitudes and raising visibility within schools.

reports, GFPs also noted the absence of sanctions to address mild forms of student violence beyond informational and attitudinal interventions. In particular, the principal in one school reported "*Teachers nowadays respect children more because now the community is involved when these cases happen. Students now have the information to report. As soon as a case happens there are already GFPs at the schools, there is immediate intervention. So teacher abuses towards girls have decreased a lot indeed. But now, among students not that much... When I go home, I still see students harassing each other...*"

5.6.1 Addressing reporting bias

We have several pieces of evidence that suggest our findings on GBV prevalence are not driven by differential reporting bias across the treatment and control groups.

First, third-party reports closely mirror self-reports. Columns 2 and 4 in Table 6 show no effect on violence perpetrated by students and a large and statistically significant reduction in violence perpetrated by teachers or staff. Respondents were 0.8–0.9 percentage points less likely to report that girls in their class experienced teacher-perpetrated violence in the last month, relative to a control mean of 1.7 percent. While we cannot reject equality of treatment effects across arms, only the T1 and T3 estimates are statistically significant and remain so after multiple-hypothesis correction.[22]

Second, we examine heterogeneity in treatment effects by baseline social desirability, following Dhar et al. (2022). Using the Crowne–Marlowe scale, we identify respondents with higher social desirability at baseline and test whether they differentially respond to treatment. Appendix Table A.5 shows no systematic heterogeneity and interaction terms are statistically insignificant, suggesting that social desirability does not drive our findings. The one-year lag between intervention and endline further reduces concerns about experimenter demand.

Third, to assess whether the intervention affected respondents' willingness to disclose violence, we compare reports of pre-2021 violence collected at baseline and endline. If treatment altered reporting propensities, treated students should revise their reports of violence prior to the intervention period. Appendix Table A.6 shows

[22]Moreover, boys in T1 schools also report lower teacher-perpetrated violence against girls. Because boys did not receive the student component in these schools, this pattern is unlikely to reflect experimenter demand.

no differential change across treatment and control schools, providing additional reassurance against differential reporting.

We also test the robustness of our results with respect to selective attrition. Following Kling et al. (2007), we calculate upper- and lower-bound estimates accounting for differential non-response (Appendix Table A.8). We also compute Lee bounds (Appendix Table A.9). In both exercises, effects remain negative and economically meaningful, supporting the robustness of our findings to alternative assumptions about attritors.

5.7 Educational Outcomes

We now examine whether reductions in teacher- and staff-perpetrated violence translated into improvements in girls' schooling. As detailed in Section 4, we measure school enrollment using a harmonized roster that allows us to capture enrollment even if girls were attending to school but not formally registered or transferred to secondary schools within the study districts.[23] Column 1 of Table 7 reports effects on this enrollment indicator, while Columns 2 and 3 report effects on standardized math and Portuguese scores collected at endline.

Column 1 shows that in schools where only girls received the student training (T1), the likelihood of being enrolled at endline increased by 5.7 percentage points relative to the control mean of 61 percent, representing a nearly 10 percent increase in enrollment. In schools where both genders received training (T3), enrollment rose by 4.4 percentage points, although the estimate is less precise at conventional levels.In contrast, the effect in T2 schools, where only boys received the student-training, is smaller (1.6 percentage points) and imprecisely estimated, with the test of equal treatment effects between T1 and T2 being marginally insignificant (p-value=0.16). These results are consistent with the increase in reporting discussed in previous sections, where the increase was concentrated in schools with treated girls.

In terms of their test scores, we do not find any significant effects on girls' performance in math or Portuguese. Although point estimates are generally positive, they are imprecisely estimated. Girls in T1 schools score slightly higher in Portuguese

[23]Because administrative data are only available for schools in the study districts, estimates could be biased if the intervention affected migration. However, we do not find differential attrition between treatment and control schools (Table A.2), suggesting this channel is unlikely.

relative to girls in T2 schools; this difference being marginally significant (p-value = 0.069).

The results in Table 7 show that when girls received the student training, their school enrollment increased by 10 percent, at a critical point of transition from upper-primary to secondary school. Together with the reductions in teacher-perpetrated violence, these findings suggest that strengthening schools' capacity to address GBV can deter abuse by teachers. However, translating safety gains into improved educational outcomes requires that potential victims are also empowered to act. We interpret the results as evidence that introducing multiple layers of accountability at the top of the school hierarchy reduces GBV, but sustained improvements in girls' schooling depend on complementary, student-focused interventions that foster agency and reporting.

5.8 Discussion

Our findings highlight that in settings characterized by asymmetrical power dynamics, reductions in gender-based violence can translate into better economic outcomes for victims when two conditions are met: (i) potential perpetrators are deterred, and (ii) victims are proactive in reporting.

In the school context, teacher-perpetrated GBV arises as an abuse of authority and limited oversight that allows it to persist. Our intervention enhanced institutional capacity to address such abuse through GFPs acting as accountability agents, satisfying condition (i). This mechanism reduced teacher-perpetrated sexual violence across all treated schools. Yet, improvements in girls' school enrollment materialized only where girls themselves received the student training and became more likely to report, satisfying condition (ii).

These results speak to a broader question: Is strengthening institutional capacity enough restore the conditions under which victims safely invest in their education? Our evidence suggests the answer is No. Institutional accountability mechanisms can deter abuse, but in environments characterized by sharp power asymmetries, victims' engagement and proactivity appears critical for removing perpetrators and enabling a safer learning environment.

6 Cost–Effectiveness and Scalability

Está na Hora de Agir was a low cost intervetion. It cost USD 19.42 per student in total and had a marginal cost of USD 10.39 per student once curriculum development and content creation costs are excluded.[24] Given a roughly 10 percent increase in girls' enrollment when girls were trained, these results compare favorably to benchmark education programs in low-income settings.

To provide scale, at baseline 252 girls reported teacher perpetrated abuse in the past month. In control schools, 63 percent of these girls remained enrolled in school one year later. On this basis, we would expect about 159 abused girls to remain enrolled at endline. Instead, in schools where girls received the training (T1), we observe 76% of girls (who had been abused at baseline) remained in school on year later. This implies, if T1 had been scaled across all schools in our sample, 33 additional abused girls would have remained in school. Of course, girls who were not abused at baseline will also benefit from the reduced risk of teacher abuse and the safer school environment, possibly making them more likely to remain in school as a result.

Although precise monetary values for the economic cost of teacher-perpetrated GBV are difficult to obtain, global evidence suggests large losses. UNICEF estimates that violence against children costs roughly 2 percent of regional GDP in East Asia and the Pacific (UNICEF, 2015) while, in the United States, the CDC estimates place the lifetime cost of rape at USD 122,000 per victim and nearly USD 3.1 trillion to the economy over the lifetime of victims (Peterson et al., 2017). Our findings indicate that preventing GBV in schools through targeted empowerment and institutional strengthening can be highly cost-effective, particularly in contexts where formal systems are under-resourced and abuse carries large social and economic consequences.

Based on these results, the Ministry of Education and UNICEF plan to scale up the intervention to other provinces and are considering complementary institutional reforms, including formalizing a code of conduct for teachers. Our work

[24]The largest program expense was the training of Gender Focal Points (GFPs), which accounted for roughly two-thirds of total costs. Each GFP training cost about USD 100 per participant, covering facilitation by NGO staff, venue rental, per diems (food and accommodation), and transport allowances, following the official scales used for professional development activities in Mozambique. Fixed costs included the production of training videos and the purchase of projectors, which were rotated across districts. No additional materials were required for the student training, keeping marginal costs low.

thus demonstrates not only the effectiveness of layered accountability structures in reducing GBV, but also their feasibility and scalability.

7 Conclusion

This paper provides experimental evidence on whether reducing GBV in schools improves girls' educational outcomes, and through which mechanisms. We partnered with the Ministry of Education in Mozambique to design and evaluate a scalable intervention that strengthened schools' capacity to address GBV through GFP training and fostered student awareness and proactivity through student-training.

One year after implementation, the intervention led to a 67 percent reduction in teacher-perpetrated sexual violence against girls. These gains occurred across all treatment arms, underscoring the effectiveness of building school-level capacity to detecter and respond to abuse. Improvements in girls' schooling, however, emerged only when girls received the student training: in those schools, girls' enrollment increased by roughly 10 percent relative to control schools, while no significant gains were observed when only boys were trained.

We interpret these findings as highlighting the necessity of a dual approach to mitigating gender-based violence and improving girls' schooling. Such an approach should involve deterring potential perpetrators and fostering proactive behaviors among victims, particularly through increased formal reporting. We posit that within contexts characterized by asymmetrical power dynamics, such as schools, reductions in GBV can enhance economic outcomes by combining mutiple levels of accountability from the top with victim proactivity to sanction abusers and restore power balance.

References

Adams-Prassl, A., K. Huttunen, E. Nix, and N. Zhang (2023, 09). Violence Against Women at Work*. *The Quarterly Journal of Economics*, qjad045.

Adams-Prassl, A., K. Huttunen, E. Nix, and N. Zhang (2024). The Economic Costs of Rape.

Amaral, S., G. Borker, N. Fiala, A. Kumar, N. Prakash, and M. M. Sviatschi (2023). Sexual harassment in public spaces and police patrols: Experimental evidence from urban india. Technical report, National Bureau of Economic Research.

Andrew, A., S. Krutikova, G. Smarelli, and H. Verma (2022). Gender Norms, Violence and Adolescent Girls' Trajectories. Evidence from a Field Experiment in India. Working paper.

Asiedu, E., D. Karlan, M. Lambon-Quayefio, and C. Udry (2021). A call for structured ethics appendices in social science papers. *Proceedings of the National Academy of Sciences 118*(29), e2024570118.

Banerjee, A., E. La Ferrara, and V. H. Orozco (2019a, July). The entertaining way to behavioral change: Fighting hiv with mtv. Working Paper 26096, National Bureau of Economic Research.

Banerjee, A., E. La Ferrara, and V. H. Orozco (2019b, May). Entertainment, Education, and Attitudes Toward Domestic Violence. *AEA Papers and Proceedings 109*, 133–37.

Benjamini, Y., A. M. Krieger, and D. Yekutieli (2006, 09). Adaptive linear step-up procedures that control the false discovery rate. *Biometrika 93*(3), 491507.

Borker, G. (2022). Constraints to women's use of public transport in developing countries, part ii: Safety. *Global Indicators Briefs 2*(10).

Boudreau, L. E., S. Chassang, A. Gonzalez-Torres, R. Heath, and N. B. of Economic Research (2023). Monitoring harassment in organizations. Technical report, National Bureau of Economic Research Cambridge, MA.

Bursztyn, L., I. K. Haaland, N. Röver, and C. Roth (2025). The social desirability atlas. Technical report, National Bureau of Economic Research.

Cassidy, R., A. Dam, W. Janssens, U. Kiani, and K. Morsink (2023). Targeting Men,

Women or Both to Reduce Child Marriage. Working paper.

Chakraborty, T., A. Mukherjee, S. R. Rachapalli, and S. Saha (2018). Stigma of sexual violence and women's decision to work. *World Development 103*, 226–238.

Chimbutane, F., C. Lauchande, C. Herreraalmanza, N. Karachiwalla, J. Leight, and D. Maggio (2022). Educating Children Together Phase 2 (ECT2) Final Evaluation Report. USDA Evaluation Report: `https://pdf.usaid.gov/pdf_docs/PA00Z94F.pdf`.

Crowne, D. P. and D. Marlowe (1960). A new scale of social desirability independent of psychopathology. *Journal of Consulting Psychology 24*(4), 349–354.

Devries, K. M., L. Knight, J. C. Child, A. Mirembe, J. Nakuti, R. Jones, J. Sturgess, E. Allen, N. Kyegombe, J. Parkes, et al. (2015). The Good School Toolkit for reducing physical violence from school staff to primary school students: a cluster-randomised controlled trial in Uganda. *The Lancet Global Health 3*(7), e378–e386.

Dhar, D., T. Jain, and S. Jayachandran (2022, March). Reshaping Adolescents' Gender Attitudes: Evidence from a School-Based Experiment in India. *American Economic Review 112*(3), 899–927.

Edmonds, E., B. Feigenberg, and J. Leight (2023). Advancing the agency of adolescent girls. *Review of Economics and Statistics 105*(4), 852–866.

Ellsberg, M., L. Heise, WHO, et al. (2005). Researching violence against women: a practical guide for researchers and activists.

Ferraz, C. and F. Finan (2025). Malfunctioning democracies: Understanding accountability failures in developing countries. Technical report, National Bureau of Economic Research.

Garlick, R., E. Field, and K. Vyborny (2025). Women's mobility and labor supply: Experimental evidence from pakistan.

Giorgio, M., E. Sully, and D. W. Chiu (2021). An assessment of third-party reporting of close ties to measure sensitive behaviors: The confidante method to measure abortion incidence in ethiopia and uganda. *Studies in Family Planning 52*(4), 513–538.

Glaeser, E. L., B. Sacerdote, and J. A. Scheinkman (1996). Crime and social interac-

tions. *The Quarterly journal of economics 111*(2), 507–548.

Green, D. P., A. M. Wilke, and J. Cooper (2020). Countering violence against women by encouraging disclosure: A mass media experiment in rural uganda. *Comparative Political Studies 53*(14), 2283–2320.

Gutierrez, I. A., O. Molina, and H. R. Nopo (2018). Stand Against Bullying: An Experimental School Intervention. IZA Discussion Paper No. 11623.

Hares, S. and G. Smarrelli (2023). *What We Know (and Don't Know) about Violence against Girls and Boys in School*. Center for Global Development.

Karmaliani, R., J. McFarlane, H. M. A. Khuwaja, Y. Somani, S. S. Bhamani, T. Saeed Ali, N. Asad, E. D. Chirwa, and R. Jewkes (2020). Right To Play's intervention to reduce peer violence among children in public schools in Pakistan: a cluster-randomized controlled trial. *Global health action 13*(1), 1836604.

Kling, J. R., J. B. Liebman, and L. F. Katz (2007). Experimental Analysis of Neighborhood Effects. *Econometrica 75*(1), 83–119.

La Ferrara, E., A. Chong, and S. Duryea (2012, July). Soap Operas and Fertility: Evidence from Brazil. *American Economic Journal: Applied Economics 4*(4), 1–31.

Lee, D. S. (2009). Training, Wages, and Sample Selection: Estimating Sharp Bounds on Treatment Effects. *The Review of Economic Studies 76*(3), 1071–1102.

Owolabi, O. O., M. Giorgio, E. Leong, and E. Sully (2023). The confidante method to measure abortion: implementing a standardized comparative analysis approach across seven contexts. *Population health metrics 21*(1), 9.

Peterson, C., S. DeGue, C. Florence, and C. N. Lokey (2017). Lifetime economic burden of rape among us adults. *American journal of preventive medicine 52*(6), 691–701.

Reyes, H., S. Kuttner, L. Morrison, I. Santagostino Recavarren, D. Kirkwood, and T. Bah (2023). Mozambique gender assessment: Leveraging women and girls' potential. *World Bank*.

Sanin, D. (2021). Paid work for women and domestic violence: Evidence from the rwandan coffee mills. Technical report, Technical report, Working Paper.

Shah, M., J. Seager, J. Montalvao, and M. Goldstein (2023, August). Sex, power,

and adolescence: Intimate partner violence and sexual behaviors. Working Paper 31624, National Bureau of Economic Research.

Sharma, K. (2022). Tackling sexual harassment: Experimental evidence from india. *Working Paper*.

Siddique, Z. (2022). Media-reported violence and female labor supply. *Economic Development and Cultural Change 70*(4), 1337–1365.

Smarelli, G. (2023). Improving School Management of Violence. Evidence from a Nation-wide Policy in Peru. Working paper.

Smarrelli, G., D. Wu, L. Baago-Rasmussen, S. Hares, and D. Naker (2024). *Violence in Schools: Prevalence, Impact, and Interventions*. Center for Global Development.

UN Women and UNESCO (2016). *Global Guidance on Addressing School-Related Gender-Based Violence*. UNESCO Publishing.

UNICEF (2015). *Estimating the Economic Burden of Violence against Children in East Asia and the Pacific*. UNICEF.

UNICEF (2018). Violence Against Childen in Schools in Mozambique: Towards an Institutionalized Reporting and Referral Mechanism.

Figure 1: Experimental design

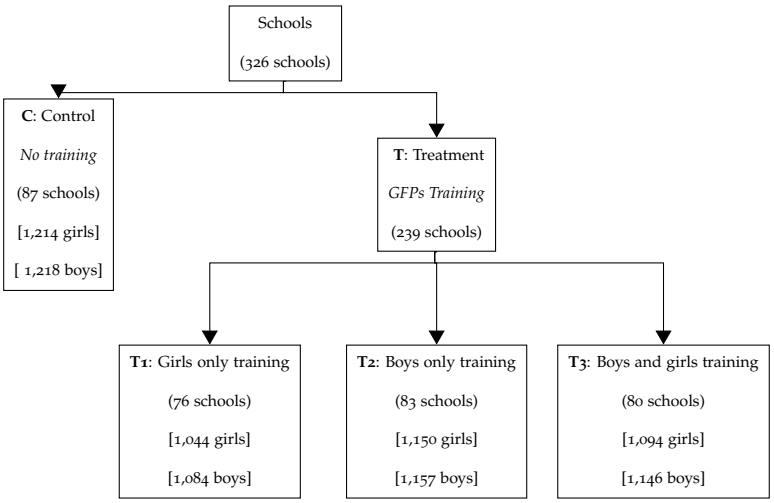


Figure 2: Map of study area and schools, by treatment status

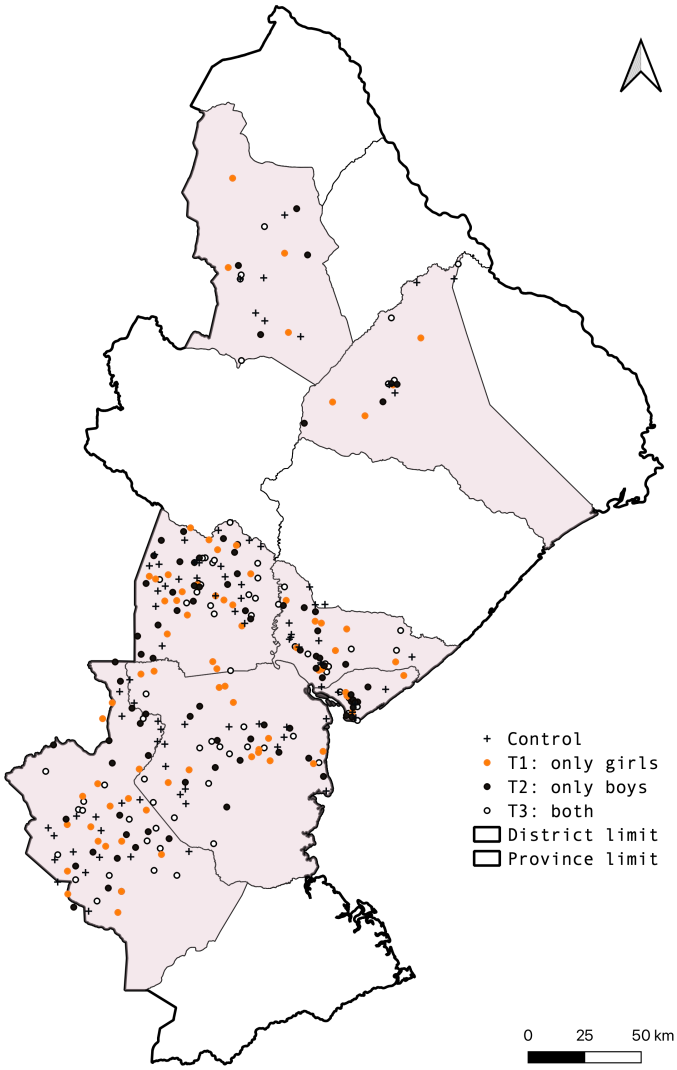


Figure 3: Project timeline

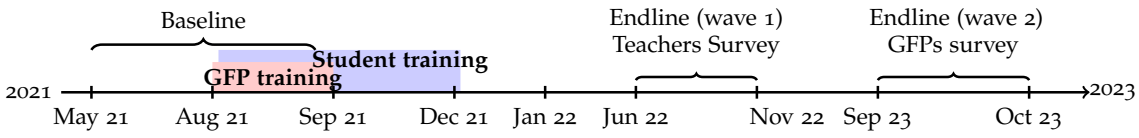
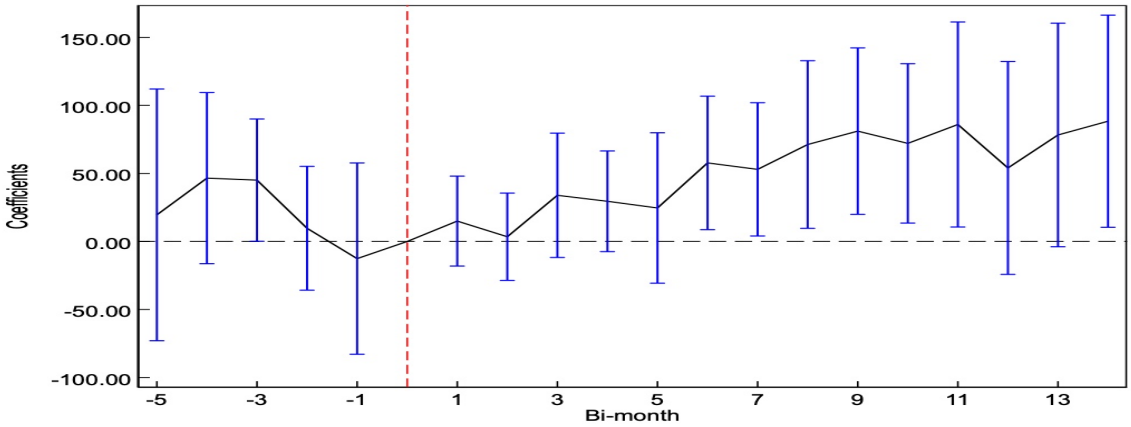
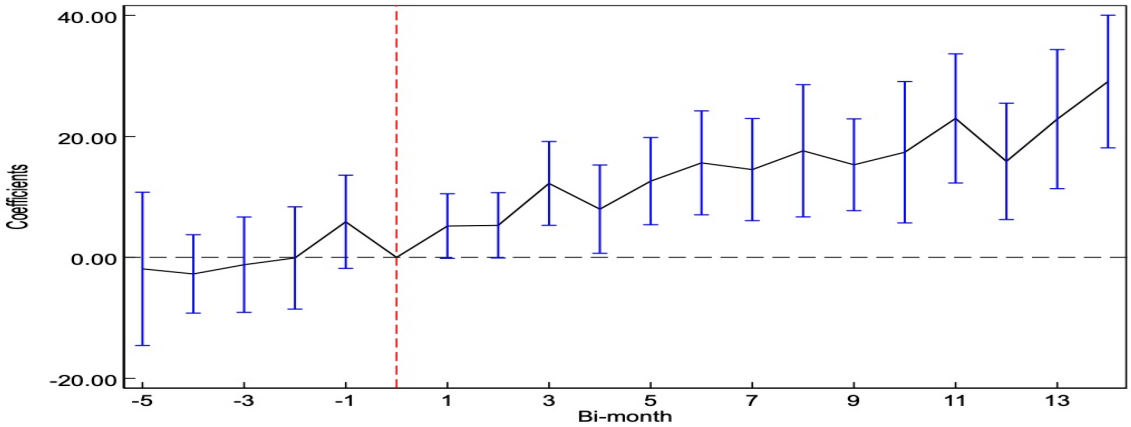


Figure 4: Event Study - Effects on calls to *Linha Fala Criança (LFC) and cases*

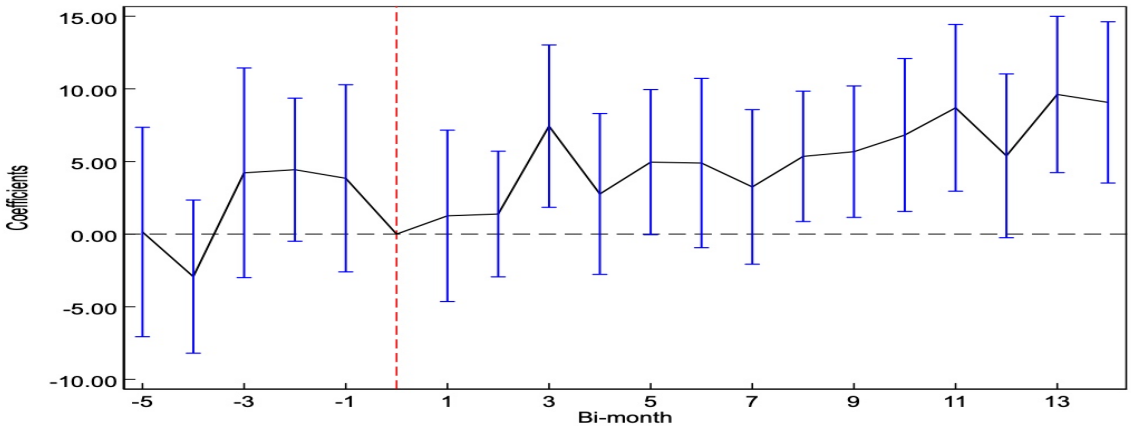
(a) Total calls



(b) Calls for assistance



(c) Cases



Notes: These graphs plot the coefficient obtained from a regression of the outcomes on the interaction between the presence of the intervention *Está na Hora de Agir* in the district and dummies for the bi-months leading up to the start of the intervention and bi-months after its introduction. Control districts are those geographically bordering the treated districts. Each bar represents the estimated coefficients and the capped vertical line shows the estimated 95% bootstrapping confidence interval. Covariates include district fixed effects, bi-month fixed effects and district linear trends.

Table 1: Baseline Characteristics and Balance Across Treatment Groups (Girls)

	(1) **Control** Mean (SD)	(2) **T1** Mean (SD)	(3) **T2** Mean (SD)	(4) **T3** Mean (SD)	(5) **T1-C**	(6) **T1-T2**	(7) **T1-T3**	(8) **T2-C**	(9) **T2-T3**	(10) **T3-C**
Number of students	*927*	*820*	*860*	*876*	-	-	-	-	-	-
Panel A: Violence in the last month										
Violence by a student (self-rep.)	0.290	0.287	0.284	0.271	0.971	0.846	0.378	0.831	0.477	0.403
	(0.454)	(0.452)	(0.451)	(0.444)						
Violence by teachers/staff (self-rep.)	0.094	0.059	0.051	0.051	0.034	0.678	0.727	0.005	0.974	0.009
	(0.292)	(0.235)	(0.221)	(0.221)						
Emotional violence	0.363	0.386	0.361	0.348	0.211	0.264	0.027	0.890	0.316	0.392
	(0.481)	(0.487)	(0.481)	(0.477)						
Physical violence	0.254	0.263	0.249	0.261	0.566	0.493	0.684	0.950	0.753	0.824
	(0.436)	(0.440)	(0.433)	(0.439)						
Sexual violence	0.171	0.182	0.169	0.149	0.444	0.583	0.086	0.893	0.242	0.227
	(0.376)	(0.386)	(0.375)	(0.356)						
Violence against girls by a student	0.184	0.176	0.208	0.169	0.801	0.171	0.577	0.290	0.052	0.433
	(0.388)	(0.381)	(0.406)	(0.375)						
Violence against girls by teachers/staff	0.042	0.027	0.022	0.031	0.153	0.588	0.653	0.042	0.302	0.305
	(0.201)	(0.162)	(0.147)	(0.173)						
Panel B: Other outcomes and socio-demographic characteristics										
Age	13.497	13.457	13.555	13.336	0.549	0.198	0.310	0.458	0.015	0.083
	(1.434)	(1.504)	(1.448)	(1.420)						
No education, mother	0.430	0.420	0.409	0.369	0.750	0.659	0.146	0.433	0.337	0.061
	(0.495)	(0.494)	(0.492)	(0.483)						
Secondary+ education, mother	0.089	0.099	0.100	0.113	0.565	0.960	0.725	0.625	0.702	0.350
	(0.285)	(0.299)	(0.300)	(0.317)						
No education, father	0.234	0.220	0.223	0.211	0.941	0.986	0.683	0.950	0.640	0.603
	(0.424)	(0.415)	(0.417)	(0.408)						
Secondary+ education, father	0.165	0.204	0.213	0.202	0.259	0.822	0.878	0.153	0.692	0.294
	(0.371)	(0.403)	(0.410)	(0.402)						
Ever had a partner	0.046	0.055	0.048	0.042	0.446	0.600	0.251	0.817	0.495	0.621
	(0.210)	(0.228)	(0.213)	(0.201)						
Has a partner	0.031	0.043	0.035	0.033	0.299	0.542	0.357	0.632	0.709	0.937
	(0.174)	(0.202)	(0.184)	(0.179)						
Initiation Rituals	0.285	0.321	0.308	0.290	0.203	0.759	0.234	0.309	0.359	0.906
	(0.452)	(0.467)	(0.462)	(0.454)						
Social desirability score	-0.010	0.043	-0.012	0.075	0.366	0.503	0.838	0.772	0.347	0.238
	(1.001)	(0.962)	(0.995)	(0.982)						

Note. All information refers to the baseline survey. The sample is restricted to girls who were tracked and resurveyed at endline. Columns 1-4 display the mean and standard deviation of the variable of interest among girls in control, T1, T2, and T3 schools respectively. Columns 5-10 display p-values based on a regression of the variable of interest on treatment dummies, controlling for randomization strata with standard errors clustered at the school (unit of randomization) level.

Table 2: Effects on Students' Ability to Identify GBV

	(1) Both vignettes as violent	(2) All 7 items correct	(3) Proportion of correct items
Girls (T1)	0.039$^{**}_{\star}$	0.005$^{**}_{\star}$	0.007
	(0.018)	(0.002)	(0.008)
Boys (T2)	0.030*	0.004*	-0.003
	(0.016)	(0.002)	(0.007)
Both (T3)	-0.002	0.002	-0.002
	(0.017)	(0.002)	(0.008)
H_0 : T1 = T2	0.617	0.521	0.181
H_0 : T1 = T3	0.057	0.533	0.915
H_0 : T2 = T3	0.025	0.255	0.228
Mean Control	.225	.002	.454
Obs.	7128	7061	7061

Notes: Regression coefficients are based on ANCOVA models with randomization strata (district $\times$ high school-level violence) fixed effects. Standard errors are clustered at the school level in parentheses. All specifications control for the social desirability score at baseline. The dependent variable in column 1 is an indicator variable of whether the respondent correctly identifies two vignettes depicting GBV as very violent (see Appendix E.II.1). The dependent variables in columns 2 and 3 are based on the correct identification of four GBV items and three non-GBV items in Appendix E.II.2. The outcome in column 2 is an indicator variable of whether the respondent correctly identifies all seven items. The outcome variable in column 3 is the proportion of items correctly identified. Statistical significance is indicated by $*$ $p < 0.1$, $**$ $p < 0.05$, $***$ $p < 0.01$ for unadjusted p-values and by $\star$ $p < 0.1$, $\star\star$ $p < 0.05$, $\star\star\star$ $p < 0.01$ for p-values adjusted for multiple hypothesis testing. When correcting for multiple hypothesis testing, we group all 3 outcomes into one family.

Table 3: Effects on Students' Attitudes Toward GBV

	(1) Acceptability violence	(2) Acceptability GBV	(3) Norms on dating violence
Girls (T1)	-0.008	0.005	0.017
	(0.021)	(0.021)	(0.020)
Boys (T2)	-0.009	-0.011	-0.011
	(0.021)	(0.022)	(0.020)
Both (T3)	-0.029	-0.006	-0.011
	(0.021)	(0.021)	(0.021)
H_0 : T1 = T2	0.972	0.467	0.177
H_0 : T1 = T3	0.324	0.801	0.990
H_0 : T2 = T3	0.307	0.624	0.183
Mean Control	.518	.347	.512
Obs.	7102	7081	7112

Note: Regression coefficients are based on ANCOVA models with randomization strata (district × high school-level violence) fixed effects. All specifications control for the social desirability score at baseline. Standard errors are clustered at the school level in parentheses. The dependent variable is an indicator of whether the respondent considers any of the four statements in Appendix E.III reflecting violence to be acceptable (column 1), considers any of the eight statements reflecting GBV to be acceptable (column 2), or whether they agree with any of the seven statements reflecting gender norms around dating violence (column 3). All specifications control for the baseline value of the dependent variable. Statistical significance is indicated by $* \ p < 0.1$, $** \ p < 0.05$, $*** \ p < 0.01$ for unadjusted p-values and by $\star \ p < 0.1$, $\star\star \ p < 0.05$, $\star\star\star \ p < 0.01$ for p-values that are adjusted for multiple hypothesis testing. When correcting the p-values for multiple hypothesis testing, we group all 3 outcomes in one family.

Table 4: Effects on GFP Actions and Reporting Mechanisms

	Students' reporting of GBV	Talks to students	Engage with teachers	Report to school authorities	Report to school council	Knows how to contact the helpline	Knows number of the helpline
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Girls (T1)	$0.098^{**}_{\star\star}$	$0.342^{***}_{\star\star\star}$	$0.151^{*}_{\star\star}$	$0.122^{**}_{\star\star}$	$0.162^{**}_{\star\star}$	$0.529^{***}_{\star\star\star}$	$0.236^{***}_{\star\star\star}$
	(0.043)	(0.074)	(0.081)	(0.058)	(0.077)	(0.068)	(0.069)
Boys (T2)	0.050	$0.312^{***}_{\star\star\star}$	0.023	$0.124^{**}_{\star\star\star}$	0.082	$0.325^{***}_{\star\star\star}$	$0.117^{*}_{\star}$
	(0.034)	(0.071)	(0.076)	(0.051)	(0.074)	(0.074)	(0.062)
Both (T3)	0.003	$0.278^{***}_{\star\star\star}$	0.113	$0.094^{*}_{\star\star}$	0.117	$0.392^{***}_{\star\star\star}$	0.100
	(0.025)	(0.071)	(0.077)	(0.051)	(0.078)	(0.075)	(0.063)
H_0 : T1 = T2	0.314	0.711	0.109	0.975	0.287	0.004	0.113
H_0 : T1 = T3	0.169	0.661	0.237	0.611	0.645	0.380	0.811
H_0 : T2 = T3	0.026	0.429	0.638	0.665	0.572	0.054	0.077
Mean Control	.024	.214	.417	.083	.381	.274	.155
Obs.	318	318	318	318	318	318	318

Notes: Regression coefficients are based on ANOVA models with randomization strata (district × high school-level violence) fixed effects. Standard errors are clustered at the school level in parentheses. The dependent variable in column 1 is an indicator equal to 1 if the respondent indicated receiving a GBV report from a student. In column 2 it is an indicator equal to 1 if the respondent talked to students at school about GBV topics. In column 3 it is an indicator equal to 1 if the respondent shared material or talked to teachers about GBV issues. In column 4 it is an indicator equal to 1 if the respondent reported GBV cases to the school authorities. In column 5 it is an indicator equal to 1 if the respondent reported GBV cases to the school council. In columns 6 and 7 it is an indicator equal to 1 if the respondent knew how to contact the helpline and could state the correct number for the *Linha Fala Criança* helpline. All the activities occurred in the past academic year. Statistical significance is indicated by $* \; p < 0.1$, $** \; p < 0.05$, $*** \; p < 0.01$ for unadjusted p-values and by $\star \; p < 0.1$, $\star\star \; p < 0.05$, $\star\star\star \; p < 0.01$ for p-values that are adjusted for multiple hypothesis testing. When correcting the p-values for multiple hypothesis testing, we group the outcomes in three families: outlined activities (1–2), students' reporting (3), activities upon reporting (4–7).

Table 5: Effects on Teachers' Knowledge of GBV Laws and Penalties

	GBV Laws and Sentences		Laws not concerning GBV		
	(1)	(2)	(3)	(4)	(5)
T1: only girls	0.211***	2.639**	0.057	0.014	-0.087
	(0.063)	(1.171)	(0.058)	(0.016)	(0.066)
T2: only boys	0.123**	0.629	0.037	-0.005	-0.046
	(0.055)	(1.201)	(0.053)	(0.011)	(0.065)
T3: both	0.066	3.633**	-0.025	-0.003	-0.059
	(0.064)	(1.674)	(0.063)	(0.012)	(0.071)
H_0 : T1 = T2	0.150	0.075	0.717	0.192	0.510
H_0 : T1 = T3	0.346	0.037	0.289	0.817	0.850
H_0 : T2 = T3	0.036	0.495	0.195	0.265	0.685
Mean Control	.285	10.559	.153	.021	.021
Obs.	551	168	551	551	551

Notes: (1) is a dummy variable = 1 if the person declares to know the Law on the Promotion and Protection of Children's Rights (Law No. 7/2008) or declares to know the sentence (in years) for sexual acts with children under 16, with or without consent; (2) Years of sentence for sexual acts with children under 16, with or without consent, according to the person; (3) knows about the law of Domestic Violence Perpetrated Against Women Act (2009); (4) knows about the Labour Law (Law No. 23/2007); (5) knows the Civil Registration Code (2004). * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table 6: Effects on prevalence of violence against girls

	Perpetrated by Students		Perpetrated by Teachers or Staff	
	Self reported	Reported by others	Self reported	Reported by others
	(1)	(2)	(3)	(4)
Girls (T1)	0.005	-0.002	-0.008$^{**}_{\star\star}$	-0.009$^{**}_{\star\star}$
	(0.022)	(0.011)	(0.004)	(0.004)
Boys (T2)	-0.005	0.002	-0.004	-0.008*
	(0.020)	(0.010)	(0.004)	(0.004)
Both (T3)	0.018	0.005	-0.006	-0.009$^{**}_{\star}$
	(0.021)	(0.011)	(0.004)	(0.004)
H_0 : T1 = T2	0.613	0.678	0.216	0.586
H_0 : T1 = T3	0.253	0.779	0.706	0.557
H_0 : T2 = T3	0.567	0.511	0.406	0.930
Mean Control	.184	.088	.012	.017
Obs.	3471	7096	3471	7096

Note: Regression coefficients are based on ANCOVA models with randomization strata (district $\times$ high school-level violence) fixed effects. Standard errors are clustered at the school level in parentheses. The dependent variables in columns 1 and 3 are indicators of whether the respondent reported experiencing any type of violence in the past month from other students in the school (column 1) or from teacher or school staff (column 3). The dependent variables in columns 2 and 4 are an indicator of whether the respondent reported witnessing any type of violence against girls in the past month from any other student in the school (column 2) or any teacher or school staff (column 4). All specifications control for the baseline value of the dependent variable and social desirability score at baseline. Statistical significance is indicated by $* \; p < 0.1$, $** \; p < 0.05$, $*** \; p < 0.01$ for unadjusted p-values and by $\star \; p < 0.1$, $\star\star \; p < 0.05$, $\star\star\star \; p < 0.01$ for p-values that are adjusted for multiple hypothesis testing. When correcting the p-values for multiple hypothesis testing, we group the outcomes in two families: by other students (1–2) and by teachers & staff (3–4).

Table 7: Effects on girls' schooling

		Test scores	
	Enrollment	Math	Portuguese
	(1)	(2)	(3)
Girls (T1)	0.057$^{**}_{\star}$	0.004	0.094
	(0.029)	(0.068)	(0.061)
Boys (T2)	0.016	0.020	-0.013
	(0.027)	(0.065)	(0.064)
Both (T3)	0.044	0.033	0.009
	(0.028)	(0.063)	(0.065)
H_0 : T1 = T2	0.160	0.812	0.069
H_0 : T1 = T3	0.326	0.829	0.731
H_0 : T2 = T3	0.653	0.655	0.156
Mean Control	.607	-.008	-.015
Obs.	4258	3483	3483

Note: The dependent variable in column 1 is an indicator of whether the respondent was enrolled at school at the moment of the interview and was interviewed. In column 2 the dependent variable is the math standardized test score based on the proportion of correct answers, and in column 3 it is the Portuguese standardized test score based on the proportion of correct answers. The construction of the dependent variables in columns 1 and 2 differs only for secondary schools. All specifications control for randomization strata (district $\times$ high school-level violence) fixed effects. In addition, specifications in columns 2 and 3 control for the baseline value of the dependent variable. Statistical significance is indicated by $*$ $p < 0.1$, $**$ $p < 0.05$, $***$ $p < 0.01$ for unadjusted p-values and by $\star$ $p < 0.1$, $\star\star$ $p < 0.05$, $\star\star\star$ $p < 0.01$ for p-values that are adjusted for multiple hypothesis testing. When correcting the p-values for multiple hypothesis testing, we group the outcomes in two families: school enrolment (1) and test scores (2–3).

Appendix

A. Ethical Considerations

The research had two study branches that required interactions with human subjects: (i) training and interviewing Gender Focal Points, and (ii) surveying students. Following Asiedu et al. (2021), below we describe how we addressed various ethical concerns to ensure the safety, privacy, and referral of all study participants. All activities were developed by the researchers jointly with UNICEF Mozambique and GCR. The implementation of the study was also completed by the MINEDH as part of a pilot initiative on the professional development of teachers. The research team operated as advisor of the implementation process, and monitored some of the activities.

A.I Policy Equipoise and Scarcity

Asiedu et al. (2021) argue that ethical randomization of participants to different treatment conditions requires either (1) policy equipoise (i.e., that there is "uncertainty regarding participants' net benefits from each of the study relative to the other arms and [relative] to the best possible policy which participants could have access") or (2) scarcity (i.e., that "no participant can be predicted to be worse off in any arm of the study than under counterfactual policy," there is "scarcity of the resources required for the arms in which participants are better off," and "all ex-ante unidentifiable participants have equal moral or legal claims to the scarce programs"). We consider that in our case, both conditions are met for the reasons outlined below.

Policy equipoise. In this paper we concluded that there is significant uncertainty and policy interest with respect to the net benefits for teachers and students from being enrolled in the study relative to the others. Previous evidence - discussed in the main paper - did not identify studies with any indication of potential negative effects to either teachers or students. In addition, teachers in our context are already embedded in a system where they are responsible to manage gender and GBV issues in schools. In addition, the school curriculum of students enrolled in grades 6 and 7 currently cover aspects related to GBV. As a result, ex-ante we do not identify that there could be adverse effects from attending a training sponsored by the MINEDH. In addition, the added benefits of the policy evaluation are reasonably justified given the potential for scale-up of the intervention.

Scarcity. As discussed in the paper there are two sources of uncertainty that justify the method of choice of the study. First, existing evidence is unclear about the potential net benefits of school-based interventions to address GBV. Second, the most suitable audience to target - girls only, boys only or both - in the intervention is

unclear. In other words, there was policy equipoise, and ex ante, our study design did not favor one specific policy intervention over others.

In the context of this study, we sampled all primary schools in the districts involved in the study. As a result, teachers and students were randomly allocated to either the treatment arms or the control arm. Therefore, all participants had ex ante equal claims to the intervention.

A.II Research Team's Role

The researchers were involved in the design of the curriculum of the intervention, overseeing the implementation of the teachers' training, and the experimental design. The research team was also responsible for securing funding for the study and hiring and training key personnel working on the study. The implementing partner for our study was the Direcção Provincial da Educação da Provincia (DPEP) de Sofala, the local representation of MINEDH at the provincial level, and GCR. These partners have extensive experience in the topic and in implementing interventions on sensitive topics with children. The MINEDH is also the public institution responsible for the management of well-being of students in schools.

A.III Potential Harms to Participants and Non-Participants

We undertook this study against a backdrop of a policy of equipoise, and hence, ex ante, there was uncertainty regarding the potential benefits or unintended consequences across experimental arms. Given the prior expertise of the implementing agencies and the underlying school curriculum, we hypothesize that GBV would decrease as a result of the treatments.

We also deliberated carefully about the potential risks from our interventions to participants, and enumerators. Below we we enumerate the steps taken to mitigate the risks.

Risks to GFP's:

- Emotional and psychological stress from the consumption and reflection of sensitive concepts of the intervention.
- Potential emotional and psychological stress associated with a greater awareness of students experiences of GBV, and from activating the referral process.

Mitigation Strategies:

- The curriculum was developed jointly with the partners. It was piloted and was subject to IRB revisions. Through this work the curriculum achieved a balance between exposure of sensitive concepts, and reflection over these in groups and with their peers.

- Training: Comprehensive training was provided on the different concepts and teachers were guided through the safety protocols in use in Mozambique.
- Support services: Teachers had frequent access to GCR paralegals - trained staff to work on GBV issues - to address any issues they faced with the implementation, referral of GBV cases by students, and to refer or use of legal aid, medical care, counseling, and temporary shelters. Prior to each session, teachers also had a discussion session with the appointed GCR staff to resolve any issues they could be facing. The district, provincial and national GFP were part of the implementation process, and there was an institutional commitment to support GFPs in schools throughout the course of the implementation and after.
- Peer support: Within a district, we created WhatsApp groups for all GFPs to exchange their experiences. This was done in order to provide peer-to-peer support during the implementation.
- Identification and cooperation: Identification cards and letters from the DPEP, MINEDH, and the local IRB were provided upon contact with schools to ensure cooperation and legitimacy of the study.

Risks to Students:

- Discomfort or distress from participating in surveys.
- Concerns about confidentiality and data security.
- Emotional and psychological stress from the consumption and reflection of sensitive concepts of the intervention.

Mitigation Strategies:

- Consent and withdrawal: All survey participants were informed about their right to consent, withdraw, and skip questions without repercussions. We also gathered parental consent to conduct the surveys.
- Confidentiality: Surveys were conducted in schools and in a private space to ensure confidentiality and data security.
- Field protocol: Enumerators were trained in gender norms, GBV concepts over the course of a week, and were trained on the World Health Organization protocol to collecting GBV survey data from children. Male students were interviewed by male enumerators and female students by female enumerators.

- Support services following the completion of surveys: Students were informed about the possibility to seek support, in case they needed, upon the completion of surveys. This information entailed sharing information about the LFC and GFP to those who disclose instances of violence.

- **Referral and response services during the intervention:** We followed a simplified version of the multi-sectoral mechanism of support to victims of GBV. This tool was developed by UNICEF – see Figure A.2 – and aims to connect victims to institutional, legal and health services trained to provide GBV support. Under this system, GFPs were guided through the different tools, their role in it, and how they must refer students who are victims of GBV. GFPs were instructed to seek victim consent prior to reporting. They could also call LFC helpline to seek guidance on how to proceed with respect to situations they were struggling to address within the school. The system was covered in the training and it was part of the manual shared with teachers. In practice, GFPs were instructed to use the hotline service since this was in many cases the only closely available option, and it was also the most simple tool to discuss and practice during sessions. With students seeking support on GBV we shared information on the first and most easily accessible tool to seek support i.e., the GFP, and information on the helpline LFC.

  When it comes to response systems, once a case is reported to the hotline, trained LFC case-workers review the information. When a call is verified to be a true report of GBV, case-workers initiate a process of investigation where further information is collected on the victim, situation and perpetrator. Victims are then referred to support systems that are close to where they reside. Case-workers that are part of - and vetted by - LFC provide support to victims in person. If the victim consents to pursue legal action, the perpetrators are then dealt by the legal system. The MINEDH is not an integral part of the referral system, but after discussions with case-workers and LFC teams, it was mentioned that during the course of an investigation, MINEDH officials and other school teachers may become aware that there is a GBV investigation taking place.

  It is important to emphasize that LFC is a free service available in the province and is targeted to responses to GBV issues. LFC operates with trained GBV teams, is linked with a network of NGOs and other institutional support systems to refer and respond to GBV in the province. LFC case-workers are trained by UNICEF on the GBV referral system and other GBV response techniques.

Risks to enumerators and other support staff:

- Emotional distress from recalling or discussing experiences of GBV.

- Safety risks during the survey process.

- Breach of confidentiality and privacy.

Mitigation Strategies:

- Informed consent: There was clear communication about the study's purpose, risks, and the right to withdraw or skip questions.
- Privacy measures: We ensured that surveys were conducted in private settings, and stopped them if privacy was compromised.
- Anonymity and data security: We ensured that all the data were anonymized and encrypted, with separate storage of anonymization and decryption keys.
- Safety protocols: The enumerators and respondents could stop the interview at any point if their safety was compromised.
- Field protocol: Enumerators were trained in gender norms, GBV concepts over the course of two weeks, and were trained on the World Health Organization protocol for collecting GBV survey data from children.
- Ethical Oversight: There was continuous monitoring by the research team to ensure adherence to ethical standards and protocols.

A.IV Conflicts of Interest and Intellectual Freedom

The researchers involved in this study have no financial or non-financial conflicts of interest to disclose. This study was funded by the J-PAL Post-Primary Education Initiative, Overdeck Fund at Princeton University, the Leibniz Association, UNICEF, GIZ, and Weiss Foundation. The researchers did not receive any personal compensation tied to the outcomes of the research.

The researchers had full autonomy in designing the study, collecting and analyzing data, and reporting the findings. There were no restrictions imposed by the funding agency, institutions, government bodies, or any other external parties on the intellectual freedom and academic independence of the researchers to conduct this study and disseminate the results. The study was approved by the National Education Council presided by Conceita Sortane, Minister of Education and Human Development at the time.

A.V Feedback to Participants

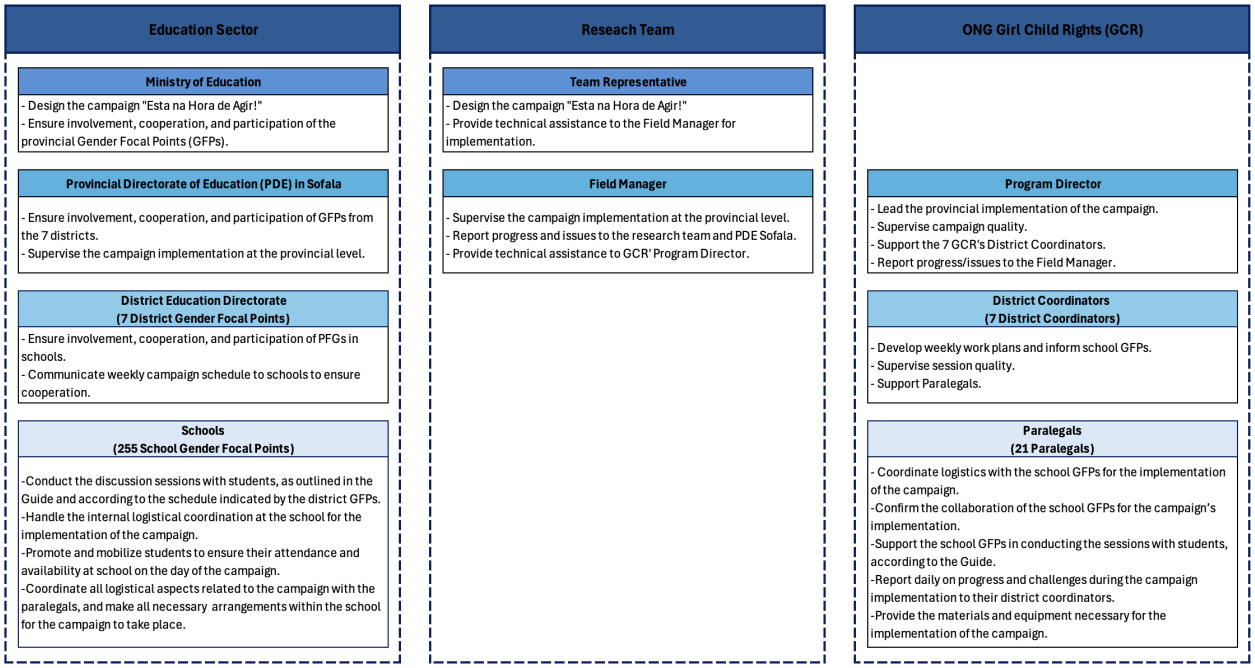
Our study's primary objective is to better understand the prevalence of gender-based violence in schools and to inform policies about the most effective policies to address it. As a result, the findings from the study have been shared with the MINEDH, UNICEF, GCR, and through regular meetings with the main stakeholders in the Education sector in the country.

A.VI Foreseeable Misuse of Research Results

The study's results have strong internal validity, and owing to our study design, the effect sizes across all our specifications are fairly reliable to inform policies in the districts of the study. While the findings from our research can be useful in other settings, we would like to draw caution against generalizing our results in other contexts. Beyond this, we do not foresee any plausible risks of the results being misused.

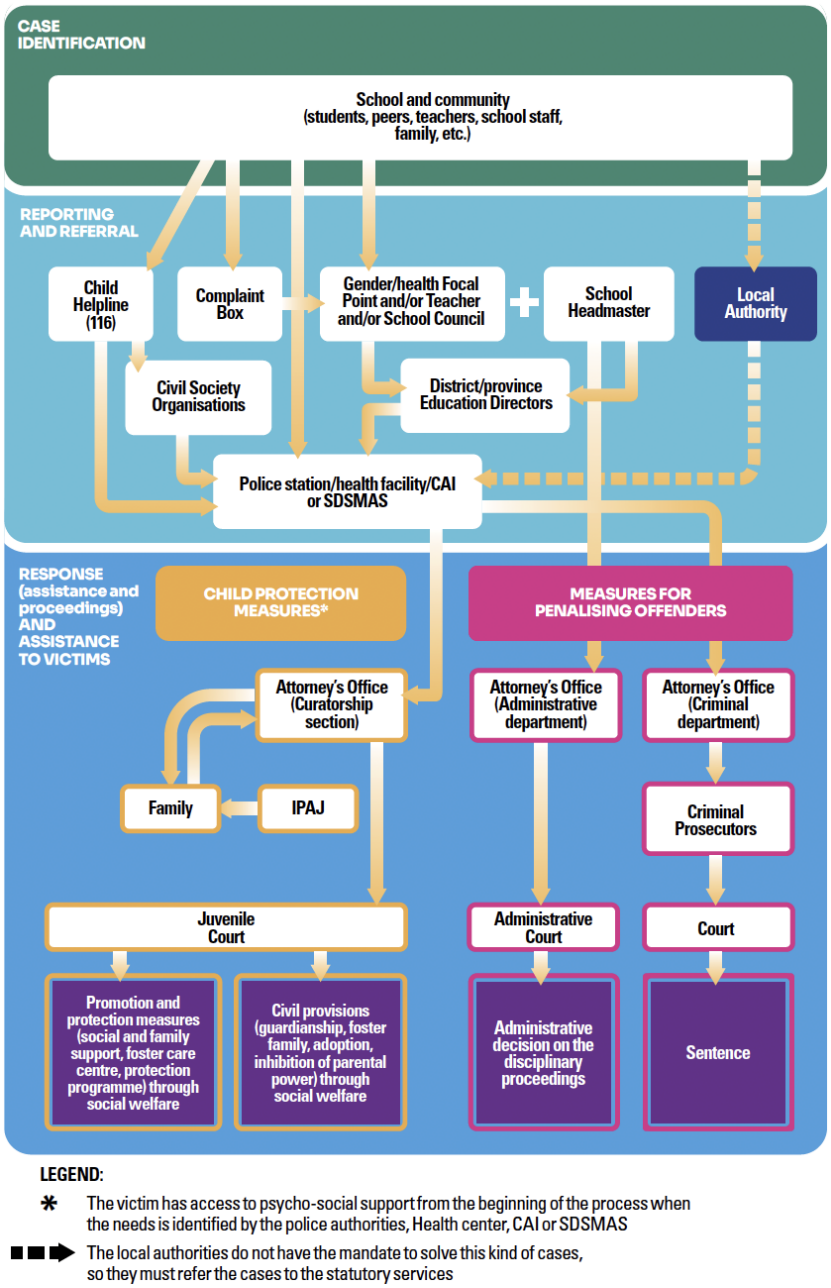
B. Appendix – Intervention details

Figure A.1: Role of GFPs in *Está na Hora de Agir*



Notes: The figure displays the organizational chart provided to GFPs during the GFP training. The bottom left panel specifies the GFP engagement role with students and their provision of logistic support to the intervention's implementation.

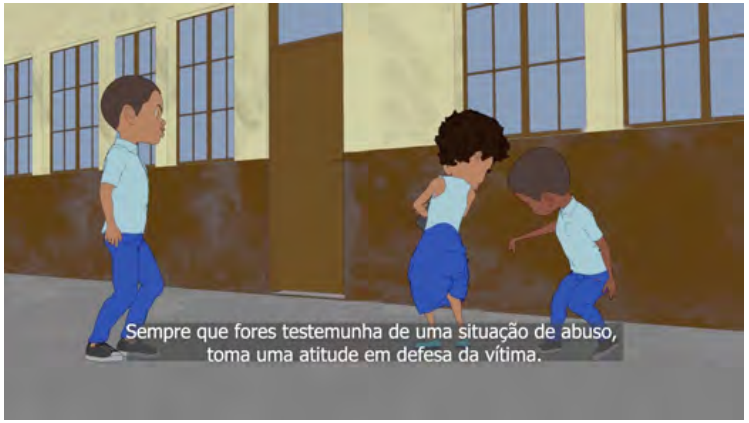
Figure A.2: UNICEF's Referral Protocol



Notes: The figure displays UNICEF's victim support mechanism used for the GFP training of the Está na Hora de Agir intervention.

Figure A.3: Sample screenshots from first animated video

(A) Screenshot 1



(B) Screenshot 2



(C) Screenshot 3



(D) Screenshot 4



Notes: The figure displays screenshots from the first animated video that was developed and used as part of the intervention. The video starts with an example of a situation where two male students lift the skirt of a female student (screenshot 1) followed by a male teacher who inappropriately rubs a female student's shoulders during class (screenshot 2) and then touches her private parts after class (screenshot 3). The video then shows the female student discussing the situation with the GFP (screenshot 4) and shows that the teacher perpetrating the GBV loses his job as a result. The video concludes with the slogan *Está na Hora de Agir* ("It's Time to Act") and the number for the *Linha Fala Criança* helpline.

Figure A.4: Sample screenshots from second animated video

(A) Screenshot 1



(B) Screenshot 2



(C) Screenshot 3



(D) Screenshot 4



Notes: The figure displays screenshots from the second animated video that was developed and used as part of the intervention. The video starts with an example of a situation where a male student pushes a female student because he is frustrated after losing a football game in school (screenshot 1). This is followed by another male student who is emotionally abusive to a female student because he jealous of her (screenshots 2 and 3). The video concludes with the slogans "*Tu tens o poder*" ("You have the power") and *Está na Hora de Agir* ("It's Time to Act") and the number for the *Linha Fala Criança* helpline.

C. Appendix – Deviations from the Pre-Analysis Plan

The study was pre-registered with AEA (ID: AEARCTR-0008361) under the title "Gender-based Violence and School Achievement" before the intervention was completed. We identify three minor deviations from the pre-analysis plan:

1. List experiments: Initially, we planned to use list experiments to address potential measurement error concerns associated with self-reports of GBV. However, the list experiments were focused on dating violence and not on violence perpetrated by teachers. As mentioned in the paper, we do not detect changes in dating violence, and we also observe a very low percentage of dating within this population. As a result, we refrained from using this information.

2. Perceived costs to perpetrators of GBV: We had planned to test the effect of the intervention on measures of the perceived costs to perpetrators of GBV. This measure was constructed using a vignette on dating violence. As mentioned in the paper, we do not detect changes in dating violence, and we also observe a very low percentage of dating within this population. As a result, we refrained from using this information, as the hypothetical scenario described in the vignette did not seem suitable to capture the outcome we had initially envisioned.

3. Calls to the hotline: We intended to test the effect of the intervention on calls made by students and teachers to the LFC. This objective was listed under the first primary outcome. Unfortunately, the marker for school location was not filled during the screening process of calls by GCR call handlers. Due to this issue, we are unable to match calls to the treatment assignment.

D. Appendix – Survey Instruments

E.I Prevalence of violence

To measure the prevalence of violence in the past month, respondents were first asked whether they had ever experienced any of the specified forms of violence. For affirmative responses, we followed up by asking when the most recent incident occurred. We also asked who the perpetrator was, using a predefined list that included a boyfriend, other students, teachers, school employees, family members, and strangers. Respondents could name multiple perpetrators for a single event, and each reported perpetrator was coded as having committed the corresponding behavior. We coded the indicator for any violence as 1 if the respondent reported experiencing any form of violence within the past month, and 0 otherwise.

1. Insulted you or made you feel bad about yourself?
2. Belittled or humiliated you in front of other people?
3. Did things to scare or intimidate you?
4. Threatened to hurt you or a friend of yours?
5. Hit you or threw something that could hurt you?
6. Pushed you or pulled your hair?
7. Punched you or hit you with something else that hurt you?
8. Kicked you, dragged you or spanked you?
9. Showed you his/her private parts or pretend to show himself to you?
10. Made nasty comments/expressions/looks/whistles at you?
11. Stalked you in a way that made you uncomfortable?
12. Groped/touched you in a way that made you uncomfortable?
13. Looked at you in a way that made you uncomfortable?
14. Made sexual comments to you in a way that made you uncomfortable?
15. Pulled your skirt/pants/shorts?

E.II Identification of violence

We use three indicators to measure adolescents' ability to identify violent acts.

Vignettes

The first indicator is based on the following two vignettes:

> **Vignette 1**:
> Ana and José are two students at the school. They are on their first date and are going to a community social gathering together. José spends the whole time talking to an old girlfriend. After José and Ana leave, Ana

gets angry and gives José a shove. He is sore but does not need medical attention.

Vignette 2:
Maria and Pedro are two students from the same school. They have been together for a month and are getting to know each other. They both went to a community meeting where they both drank a beer that someone offered them. After they leave, Maria gets very angry and hits Pedro. He is hurt and needs a bandage. This kind of thing has happened several times before.

After the vignettes were read out to the respondents, they were asked to report to what extent they thought this situation was violent or abusive. They could answer "Very violent," "More or less violent," "Slightly violent," or "Not violent at all." The indicator was coded as 1 if the respondent found both situations "Very violent" and 0 otherwise.

Situation Classification

The second and third indicators are based on a series of questions describing various situations. The respondent was asked to report if they thought the situation constituted GBV or not. The specific situations are the following:

a. A group of students is tugging on a girl's skirt.
b. A female student yells at a male student because he has different ideas.
c. A teacher uses a female student's pen without asking.
d. A male student insults a female student because he thinks her outfit is racy.
e. A teacher slaps a student if he doesn't pay attention.
f. A teacher pretends to show his private parts (parts that are normally not shown) to a female student.
g. A girlfriend pushes her boyfriend because she is jealous.

The correct answer should be that items a, d, f, and g constitute GBV and items b, c, and e do not. We then generate one indicator that is equal to 1 if the respondent identified all of these items correctly as being GBV or not, and a second indicator that is the proportion of items they identified correctly.

E.III Attitudes

We use three indicators to measure adolescents' attitudes toward gender-based violence.

1. Acceptability of violence: An indicator equal to 1 if the respondent considers the use of violence as being "acceptable" in any of the following situations:

- A student insults/humiliates another student if they argue or have a disagreement.
- A teacher threats/hurts/scares a student if they argue or have a disagreement.
- A student pushes/slaps/hits/kicks another student if they argue or have a disagreement.
- A teacher pushes/slaps/hits/kicks a student if he/she is not behaving as the teacher expects (e.g., not paying attention, disrupting the class).

2. Acceptability of GBV: An indicator equal to 1 if the respondent considers the use of violence as being "acceptable" in any of the following situations:

- A student shows or simulates showing his/her intimate parts to another student if he/she likes her/him.
- A teacher shows or simulates showing his/her intimate parts to a student if he/she likes her/him.
- A student stalks/stares at another student if he/she likes her/him.
- A teacher stalks/stares at a student if he/she likes her/him.
- A student touches/gropes another student if he/she likes her/him.
- A teacher touches/gropes a student if he/she likes her/him.
- A student makes comments or sexually harass another student if he/she likes her/him.
- A teacher makes comments or sexually harass a student if he/she likes her/him.

3. Gender Norms and Dating Violence: An indicator equal to 1 if the respondent agrees with any of the following statements:

- What boys want should take priority over what girls want when there is no money in the house.
- If a boy likes a girl, he should be able to kiss her even if she doesn't want to.
- If a boy likes a girl, he should be able to touch her even if she doesn't want to.
- A boyfriend takes care of his girlfriend by controlling where she goes and who she sees.
- A boyfriend has the right to have sex with his girlfriend even when she says no.
- In a relationship, the guy should always have the last word, even if it is not right.
- If girls wear short skirts or drink alcohol at a party, they are asking to be mistreated or abused.

E.IV Social desirability index

Our social desirability index is identical to the index used by Dhar et al. (2022), which is based on Crowne and Marlowe (1960). In particular, each respondent was

asked to state if the following statements are true or false for themselves:

1. It is sometimes hard for me to go on with my work if I am not encouraged.
2. I sometimes feel resentful when I don't get my way.
3. On a few occasions, I have given up doing something because I thought too little of my ability.
4. There have been times when I felt like rebelling against people in authority even though I knew they were right.
5. No matter who I'm talking to, I'm always a good listener.
6. There have been occasions when I took advantage of someone.
7. I'm always willing to admit it when I make a mistake.
8. I sometimes try to get even rather than forgive and forget.
9. I am always courteous, even to people who are disagreeable.
10. I have never been upset when people expressed ideas very different than my own.
11. There have been times when I was quite jealous of the good fortune of others.
12. I am sometimes irritated by people who ask favors of me.
13. I have deliberately said something that hurt someone's feelings.

The social desirability index sums how many of the responses are the socially desirable ones. More specifically, we create dummy variables equal to 1 if the respondent responded "False" to items 1–4, 6, 8, and 11–13; and "True" to items 5, 7, 9, and 10. The social desirability index is the mean of these 13 dummy variables, standardized with respect to the control group.

E. Appendix –Additional Tables and Figures

Table A.1: Baseline Characteristics and Balance Across Treatment Groups (Schools)

	(1) **Control** Mean (SD)	(2) **T1** Mean (SD)	(3) **T2** Mean (SD)	(4) **T3** Mean (SD)	(5) **T1-C**	(6) **T1-T2**	(7) **T1-T3**	(8) **T2-C**	(9) **T2-T3**	(10) **T3-C**
Number of schools	*87*	*76*	*83*	*80*	-	-	-	-	-	-
Has access to drinking water	0.609	0.645	0.675	0.637	0.632	0.690	0.823	0.378	0.540	0.809
	(0.491)	(0.482)	(0.471)	(0.484)						
Has access to electricity	0.161	0.184	0.145	0.188	0.434	0.205	0.754	0.613	0.333	0.637
	(0.370)	(0.390)	(0.354)	(0.393)						
Has bathroom	0.161	0.184	0.241	0.262	0.682	0.431	0.293	0.205	0.786	0.122
	(0.370)	(0.390)	(0.430)	(0.443)						
Total number of students	783.460	665.303	741.771	785.912	0.377	0.636	0.381	0.634	0.645	0.965
	(1,010.497)	(781.541)	(790.905)	(813.334)						
Female student ratio at school	0.450	0.455	0.455	0.450	0.566	0.993	0.521	0.590	0.542	0.909
	(0.067)	(0.052)	(0.064)	(0.072)						
Number of primary school students	161.885	150.613	147.494	157.875	0.803	0.675	0.972	0.540	0.641	0.828
	(269.914)	(211.056)	(179.636)	(191.973)						
Primary school students per turma	36.849	39.407	38.282	38.492	0.314	0.636	0.656	0.580	0.993	0.625
	(18.727)	(20.689)	(17.695)	(22.285)						
Number of primary school teachers	13.540	12.592	13.253	13.750	0.751	0.939	0.681	0.801	0.723	0.923
	(15.526)	(14.127)	(13.224)	(13.980)						
Female teacher ratio	0.374	0.383	0.452	0.387	0.676	0.098	0.941	0.018	0.060	0.709
	(0.271)	(0.298)	(0.248)	(0.265)						
Teachers with higher education	0.913	0.870	0.934	0.904	0.135	0.013	0.258	0.284	0.172	0.742
	(0.154)	(0.189)	(0.108)	(0.163)						

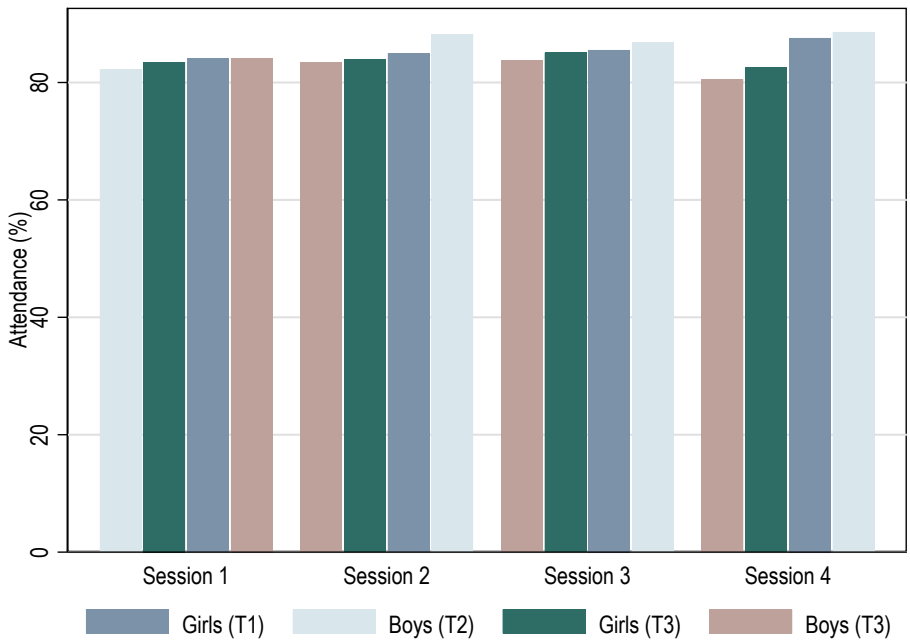
Note. Variables on school infrastructure and student composition are obtained from the 2019 School Census, whereas variables pertaining to teachers is drawn from the 2021 School Census. Columns 1-4 display the mean and standard deviation of the variable of interest in control, T1, T2, and T3 schools respectively. Columns 5-10 display p-values based on a regression of the variable of interest on treatment dummies, controlling for randomization strata with standard errors clustered at the school (unit of randomization) level.

Table A.2: Attrition

	(1) All	(2) Girls	(3) Boys
Girls (T1)	-0.020	-0.024	-0.016
	(0.022)	(0.025)	(0.024)
Boys (T2)	0.010	0.024	-0.004
	(0.023)	(0.027)	(0.024)
Both (T3)	-0.033	-0.037	-0.028
	(0.022)	(0.025)	(0.025)
Observations	8558	4258	4300
Control mean	0.178	0.192	0.165
P-value T1=T2	0.168	0.061	0.599
P-value T1=T3	0.516	0.582	0.633
P-value T2=T3	0.054	0.017	0.330

Note: The dependent variables are indicators =1 if the respondent could not be re-surveyed at endline. Regression coefficients are based on OLS models with randomization strata (district × high school-level violence) fixed effects. Standard errors are clustered at the school level in parentheses. Statistical significance is indicated by $* \ p < 0.1$, $** \ p < 0.05$, $*** \ p < 0.01$.

Figure A.5: Attendance to Student Training by Session



Notes: The figure displays the attendance rate to the intervention sessions, according to students' gender and treatment group.

Table A.3: Attendance to Student Training

	(1) Session 1	(2) Session 2	(3) Session 3	(4) Session 4
Girls (T1)	0.838***	0.854***	0.850***	0.873***
	(0.017)	(0.016)	(0.016)	(0.015)
Girls (T1)=1 × Boy	-0.839***	-0.855***	-0.851***	-0.875***
	(0.017)	(0.016)	(0.017)	(0.016)
Boys (T2)	0.000	-0.001	-0.000	-0.000
	(0.002)	(0.004)	(0.005)	(0.005)
Boys (T2)=1 × Boy	0.811***	0.871***	0.860***	0.874***
	(0.019)	(0.013)	(0.013)	(0.014)
Both (T3)	0.840***	0.849***	0.863***	0.840***
	(0.017)	(0.014)	(0.015)	(0.020)
Both (T3)=1 × Boy	0.004	-0.003	-0.014	-0.025*
	(0.015)	(0.015)	(0.016)	(0.015)
N. Clusters	326	326	326	326
Observations	8558	8558	8558	8558
Control mean	0.000	0.000	0.000	0.000
T1 + T1 × Boy	-0.001	-0.001	-0.001	-0.001
	(0.002)	(0.004)	(0.004)	(0.005)
p-value	0.816	0.768	0.838	0.781
T2 + T2 × Boy	0.812	0.870	0.860	0.874
	(0.019)	(0.013)	(0.012)	(0.013)
p-value	0.000	0.000	0.000	0.000
T3 + T3 × Boy	0.844	0.846	0.849	0.816
	(0.015)	(0.014)	(0.014)	(0.020)
p-value	0.000	0.000	0.000	0.000

Note: Dependent variable in each column indicates whether the student attended the corresponding training session. Regression coefficients are based on OLS models with randomization strata (district × high school-level violence) fixed effects. Standard errors are clustered at the school level in parentheses. Statistical significance is indicated by $* p < 0.1$, $** p < 0.05$, $*** p < 0.01$.

Table A.4: Balance test for GFPs

	Age	Gender (male=1)	Born in Sofala	Born in Beira	University graduate	Years in school	GFP in 2021
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Girls (T_1)	-0.173	0.027	0.041	0.058	-0.063	0.407	0.062
	(0.983)	(0.070)	(0.056)	(0.077)	(0.055)	(0.697)	(0.045)
Boys (T_2)	-0.383	-0.135**	0.019	0.071	0.038	0.267	0.058
	(0.931)	(0.061)	(0.057)	(0.077)	(0.059)	(0.677)	(0.044)
Both (T_3)	-0.083	0.071	-0.015	0.017	-0.034	0.311	0.080*
	(0.971)	(0.070)	(0.060)	(0.075)	(0.058)	(0.700)	(0.043)
Observations	318	318	318	318	318	318	318
Control mean	35.05	0.27	0.83	0.38	0.25	6.77	0.88
P-value T1=T2	0.82	0.01	0.69	0.87	0.07	0.84	0.91
P-value T1=T3	0.93	0.55	0.33	0.60	0.59	0.89	0.61
P-value T2=T3	0.74	0.00	0.56	0.49	0.23	0.95	0.52

Note: Regression coefficients are based on OLS models with randomization strata (district $\times$ high school-level violence) fixed effects. Standard errors are clustered at the school level in parentheses. Statistical significance is indicated by $*$ $p < 0.1$, $**$ $p < 0.05$, $***$ $p < 0.01$.

Table A.5: Effects on prevalence of violence against girls, by social desirability score

	Perpetrated by Students		Perpetrated by Teachers or Staff	
	Self reported	Reported by others	Self reported	Reported by others
	(1)	(2)	(3)	(4)
Girls (T1)	0.005	-0.002	-0.008**	-0.009**
	(0.022)	(0.011)	(0.004)	(0.004)
Girls (T1) × SDB	0.013	0.016*	0.001	-0.004
	(0.019)	(0.009)	(0.003)	(0.004)
Boys (T2)	-0.005	0.002	-0.004	-0.008*
	(0.020)	(0.010)	(0.005)	(0.004)
Boys (T2) × SDB	0.020	0.004	0.006	-0.003
	(0.019)	(0.010)	(0.005)	(0.004)
Both (T3)	0.018	0.006	-0.005	-0.010**
	(0.021)	(0.011)	(0.004)	(0.004)
Both (T3) × SDB	0.014	-0.010	0.000	-0.003
	(0.017)	(0.010)	(0.004)	(0.004)
Observations	3471	7098	3471	7098
Control mean	0.184	0.088	0.012	0.017
T1 + T1 × SDB	0.018	0.014	-0.007	-0.014
	(0.030)	(0.015)	(0.005)	(0.005)
p-value	0.553	0.370	0.144	0.013
T2 + T2 × SDB	0.015	0.006	0.002	-0.011
	(0.030)	(0.015)	(0.007)	(0.006)
p-value	0.615	0.667	0.769	0.064
T3 + T3 × SDB	0.032	-0.004	-0.005	-0.013
	(0.027)	(0.014)	(0.005)	(0.005)
p-value	0.238	0.757	0.332	0.021

Note: Regression coefficients are based on ANCOVA models with randomization strata (district × high school-level violence) fixed effects. Standard errors are clustered at the school level in parentheses. The dependent variables in columns 1 and 3 are indicators of whether a girl reported experiencing any type of violence in the past month from any other student in the school (column 1) or any teacher or school staff (column 3). The dependent variables in columns 2 and 4 are indicators of whether the student reported witnessing any type of violence against girls in the past month from any other student in the school (column 2) or any teacher or school staff (column 4). SDB is the respondent's social desirability score at baseline. All specifications control for the baseline value of the dependent variable. * p<0.10, ** p<0.05, *** p<0.01.

Table A.6: Changes in Reporting of Violence Against Girls Prior to the Intervention

	Perpetrated by students	Perpetrated by teachers or staff
	(1)	(2)
Girls (T1)	0.000	0.001
	(0.022)	(0.008)
Boys (T2)	-0.021	0.012
	(0.022)	(0.008)
Both (T3)	-0.014	0.001
	(0.021)	(0.009)
Observations	3470	3470
Control mean	-0.023	-0.017
P-value T1=T2	0.327	0.156
P-value T1=T3	0.498	0.950
P-value T2=T3	0.740	0.151

Note: The dependent variables are based on retrospective questions about violence the respondent experienced before 2021. We generate indicators of whether the respondent reported experiencing any type of violence from other students in the school or from teacher or school staff based on her responses at baseline and at endline surveys. We then take the difference between the endline and baseline indicators so the outcome is the *change* in reported prevalence of violence pre-2021. All specifications control for respondent's social desirability score at baseline. Regression coefficients are based on ANCOVA models with randomization strata (district $\times$ high school-level violence) fixed effects. Standard errors are clustered at the school level in parentheses. Statistical significance is indicated by $* \ p < 0.1$, $** \ p < 0.05$, $*** \ p < 0.01$.

Table A.7: Sexual Violence by Teachers or School Staff

	Forced sexual act	Unwanted touching	Forced kissing	Forced undress	Teacher self-exposed	Forced self-touching	Forced touching perpetrator
Girls (T1)	-0.008*	-0.002	-0.002	-0.002	-0.004	-0.002	-0.002
	(0.005)	(0.005)	(0.003)	(0.004)	(0.004)	(0.004)	(0.004)
Boys (T2)	-0.008*	-0.008**	-0.002	-0.004	-0.004	-0.003	-0.004
	(0.005)	(0.004)	(0.003)	(0.004)	(0.004)	(0.003)	(0.003)
Both (T3)	-0.008*	-0.006	0.004	-0.005	-0.006	-0.004*	-0.004
	(0.005)	(0.004)	(0.004)	(0.004)	(0.004)	(0.003)	(0.003)
H_0 : T1 = T2	0.946	0.049	0.923	0.432	0.938	0.721	0.461
H_0 : T1 = T3	0.990	0.447	0.078	0.979	0.666	0.689	0.952
H_0 : T2 = T3	0.953	0.323	0.096	0.418	0.629	0.418	0.442
Mean Control	.011	.009	.004	.007	.008	.006	.007
Obs.	3305	3284	3333	3201	3333	3284	3323

Notes: Regression coefficients are from ANCOVA models with randomization-strata fixed effects (district × high school-level violence). Standard errors clustered at the school level in parentheses. Dependent variables equal 1 if the respondent reported that a teacher or school staff member engaged in the specified act in the past month. All specifications control for the baseline value of the dependent variable and baseline social-desirability score. Statistical significance: $*p < 0.1$, $**p < 0.05$, $***p < 0.01$ (unadjusted); $\star p < 0.1$, $\star\star p < 0.05$, $\star\star\star p < 0.01$ (MHT-adjusted). All outcomes belong to a single MHT family.

Table A.8: Effects on Violence Against Girls – Attrition bounds à la Kling et al. (2007)

	Perpetrated by Students		Perpetrated by Teachers or Staff	
	Self reported	Reported by others	Self reported	Reported by others
	(1)	(2)	(3)	(4)
	Lower Kling-Liebman-Katz bounds			
Girls (T1)	0.027	-0.002	0.002	-0.009$^{**}_{\star}$
	(0.018)	(0.011)	(0.003)	(0.004)
Boys (T2)	-0.016	0.002	-0.004	-0.008*
	(0.016)	(0.010)	(0.004)	(0.004)
Both (T3)	0.017	0.005	-0.004	-0.009$^{**}_{\star}$
	(0.017)	(0.011)	(0.004)	(0.004)
H_0 : T1 = T2	0.015	0.678	0.036	0.586
H_0 : T1 = T3	0.050	0.779	0.970	0.557
H_0 : T2 = T3	0.605	0.511	0.036	0.930
Mean Control	.17	.088	.008	.017
Obs.	4258	7096	4258	7096
	Upper Kling-Liebman-Katz bounds			
Girls (T1)	-0.030	-0.002	-0.018$^{***}_{\star\star\star}$	-0.009$^{**}_{\star\star}$
	(0.018)	(0.011)	(0.003)	(0.004)
Boys (T2)	-0.015	0.002	-0.005	-0.008*
	(0.017)	(0.010)	(0.004)	(0.004)
Both (T3)	0.011	0.005	-0.008$^{**}_{\star\star}$	-0.009$^{**}_{\star\star}$
	(0.018)	(0.011)	(0.004)	(0.004)
H_0 : T1 = T2	0.398	0.678	0.000	0.586
H_0 : T1 = T3	0.136	0.779	0.427	0.557
H_0 : T2 = T3	0.030	0.511	0.000	0.930
Mean Control	.201	.088	.02	.017
Obs.	4258	7096	4258	7096

Note: Regression coefficients are based on ANCOVA models with randomization strata (district × high school-level violence) fixed effects. Standard errors are clustered at the school level in parentheses. The dependent variables in columns 1 and 3 are indicators of whether the respondent reported experiencing any type of violence in the past month from other students in the school (column 1) or from teacher or school staff (column 3). The dependent variables in columns 2 and 4 are an indicator of whether the respondent reported witnessing any type of violence against girls in the past month from any other student in the school (column 2) or any teacher or school staff (column 4). All specifications control for the baseline value of the dependent variable and social desirability score at baseline. Statistical significance is indicated by $* \; p < 0.1$, $** \; p < 0.05$, $*** \; p < 0.01$ for unadjusted p-values and by $\star \; p < 0.1$, $\star\star \; p < 0.05$, $\star\star\star \; p < 0.01$ for p-values that are adjusted for multiple hypothesis testing. When correcting the p-values for multiple hypothesis testing, we group the outcomes in two families: by other students (1–2) and by teachers & staff (3–4). Bounds calculated following Kling et al. (2007).

Table A.9: Effects on Violence Against Girls – Attrition bounds à la Lee (2009)

	Perpetrated by Students		Perpetrated by Teachers or Staff	
	Self reported	Reported by others	Self reported	Reported by others
	(1)	(2)	(3)	(4)
	Lower Lee bounds			
Girls (T1)	-0.021	-0.002	$-0.008^{***}_{\star\star}$	$-0.009^{**}_{\star\star}$
	(0.021)	(0.011)	(0.003)	(0.004)
Boys (T2)	0.000	0.002	0.001	-0.008^{*}
	(0.020)	(0.010)	(0.004)	(0.004)
Both (T3)	-0.027	0.005	$-0.008^{***}_{\star\star\star}$	$-0.009^{**}_{\star\star}$
	(0.019)	(0.011)	(0.003)	(0.004)
H_0 : T1 = T2	0.302	0.678	0.004	0.586
H_0 : T1 = T3	0.149	0.779	0.005	0.557
H_0 : T2 = T3	0.747	0.511	0.817	0.930
Mean Control	.178	.08	.008	.008
Obs.	3375	7096	3411	7096
	Upper Lee bounds			
Girls (T1)	0.011	-0.002	$-0.008^{*}_{\star}$	$-0.009^{**}_{\star}$
	(0.023)	(0.011)	(0.004)	(0.004)
Boys (T2)	-0.007	0.002	-0.004	-0.008^{*}
	(0.020)	(0.010)	(0.005)	(0.004)
Both (T3)	0.027	0.005	-0.005	$-0.009^{**}_{\star}$
	(0.022)	(0.011)	(0.005)	(0.004)
H_0 : T1 = T2	0.408	0.678	0.250	0.586
H_0 : T1 = T3	0.105	0.779	0.796	0.557
H_0 : T2 = T3	0.483	0.511	0.396	0.930
Mean Control	.185	.089	.012	.017
Obs.	3384	7096	3380	7096

Note: Regression coefficients are based on ANCOVA models with randomization strata (district × high school-level violence) fixed effects. Standard errors are clustered at the school level in parentheses. The dependent variables in columns 1 and 3 are indicators of whether the respondent reported experiencing any type of violence in the past month from other students in the school (column 1) or from teacher or school staff (column 3). The dependent variables in columns 2 and 4 are an indicator of whether the respondent reported witnessing any type of violence against girls in the past month from any other student in the school (column 2) or any teacher or school staff (column 4). All specifications control for the baseline value of the dependent variable and social desirability score at baseline. Statistical significance is indicated by $* \; p < 0.1$, $** \; p < 0.05$, $*** \; p < 0.01$ for unadjusted p-values and by $\star \; p < 0.1$, $\star\star \; p < 0.05$, $\star\star\star \; p < 0.01$ for p-values that are adjusted for multiple hypothesis testing. When correcting the p-values for multiple hypothesis testing, we group the outcomes in two families: by other students (1–2) and by teachers & staff (3–4). Bounds calculated following Lee (2009).

F. Appendix – GFPs' Interviews

In May 2024, we conducted a qualitative study to complement and contextualize the quantitative results. The study consisted of 14 semi-structured, in-depth interviews with Gender Focal Points (GFPs), equally distributed between Beira (7) and Dondo (7) districts. The sample included 5 GFPs from T1 schools, 4 from T2 schools, and 5 from T3 schools. The interviews examined GFPs' implementation of the *Está na Hora de Agir!* campaign, including their recollection of the training, the preventive and reporting actions they adopted, and their perceptions of shifts in student behavior, wellbeing, and reporting patterns. GFPs were also asked to reflect on the main quantitative findings to assess the extent to which these patterns aligned with their experience on the ground.

Our findings show that the reporting of victims and GFPs' familiarity with the helpline's exact number significantly improved only in schools where the training was exclusive to girls (Table 4). Our conversation with the GFPs indicate that GFPs in T1 were more aware of the LFC number because girls reported more cases of GBV:

"Teachers had more GBV cases to consult or report. They always called the Linha Fala Criança."

"Teachers remember [the LFC number] more because it was something that was usually used, so they ended up being more connected to the number."

We also learned how GFPs were constantly approaching students and teachers to talk about GBV and involving the community:

"And if the teacher grabbed the child's butt and hugged the child, we explained that type of behavior was not acceptable [...] I told the students that if anything happened, they should come and talk to me[...] Yes, and I will also forward it to the bosses (school council). [By involving the community], the teachers already feared punishment. "