Youth First Supports Adolescent Wellbeing at Scale: A Longitudinal Cluster Randomized Controlled Trial of a Teacher-Delivered School-Based Program in India

Peter Cooper,[1] Katherine Sachs Leventhal,[2,3] Gunjika Misra,[4,5] Nandita Bhatla,[4] Hari Shanker,[4] and Steve Leventhal[2]

[1] Independent Evaluation Consultant, Weed, CA, US;

[2] WorldBeing, Baltimore, MD, US;

[3] University of Edinburgh, School of Health in Social Sciences, Edinburgh, UK;

[4] WorldBeing India Foundation, New Delhi, India;

[5] International Institute of Health Management Research, New Delhi, India.

Author Note

Peter Cooper, https://orcid.org/0009-0009-1135-959X; Katherine Sachs Leventhal, katel@worldbeing.org, https://orcid.org/0000-0003-4521-2675; Gunjika Misra, gunjikam@worldbeing.org, https://orcid.org/0000-0002-1031-966X; Nandita Bhatla, nanditab@worldbeing.org, https://orcid.org/0000-0003-1001-9074; Hari Shanker, haris@worldbeing.org, https://orcid.org/0009-0009-0600-9077 ; Steve Leventhal, stevel@worldbeing.org.

Correspondence concerning this article should be addressed to Peter Cooper, Independent Evaluation Consultant, 3935 Stewart Springs Road, Weed, CA, 96094.. Email: peterlcooper@me.com .

Data Availability Statement: Data are not publicly available due to ethical safeguards in place for children and vulnerable populations, put in place during the ethical review process.

Funding Statement: The original data collection was supported by The David & Lucile Packard Foundation. Secondary analysis was supported by ICONIQ.

Ethics Approval Statement: This research was approved by the Sangath Institutional Review Board and conducted in accordance with the approved protocol.

Potential Conflict of Interest Disclosure: PC was an independent external evaluation consultant on this research. KSL, GM, NB, HS, and SL were employed by WorldBeing, the organization that developed and conducts Youth First (evaluated in this study).

Abstract

Introduction

Globally, adolescents are facing a “polycrisis” of intersecting challenges that place them at heightened risk for poor mental health outcomes, particularly in low- and middle-income countries (LMICs), where 90% of the world’s adolescents live. School-based wellbeing programs offer a promising, cost-effective prevention strategy, yet evidence from large-scale, government-delivered interventions with longer-term follow-ups remains limited. This study evaluates the effectiveness of Youth First, a teacher-delivered, in-school wellbeing program, as conducted in Bihar, India, via government systems, addressing key evidence gaps relevant to LMIC government adoption.

Methods

We conducted secondary analysis of data from a cluster-randomized controlled trial across 69 government schools in Bihar, India, enrolling 5,445 Grade 7–8 students. Schools were randomized to Youth First or control (standard education). Data were collected at baseline, 8 months, and 20 months. Outcomes measured included physical and mental wellbeing, psychosocial assets, gender attitudes, empowerment, and tobacco use. Analyses followed an intention-to-treat framework using mixed-effects difference-in-differences models with random intercepts for schools and students.

Results

Youth First led to statistically significant improvements in social and peer support across multiple time points and cohorts. Additional significant effects were observed for self-efficacy, empowerment, internalizing symptoms, psychological wellbeing, and tobacco use, with patterns varying by grade and exposure duration. Notably, several effects persisted or intensified one year post-intervention, including for psychological wellbeing and self-efficacy.

Conclusion

This longitudinal cluster-randomized controlled trial demonstrates that a scalable, government-delivered school wellbeing program can produce meaningful and sustained benefits across multiple priority outcomes for adolescents in an LMIC setting.

Key words: Schools; Wellbeing; Adolescents; Cluster-randomized controlled trial; Low- and middle-income countries; Promotion

Introduction

Globally, adolescents face challenges including climate change, social media pressures, rising violence and crime, poverty, discrimination, migration, and displacement (Kałwak et al., 2024; World Federation of Public Health Associations, 2025; World Health Organization, 2025). These crises are amplified for adolescents living in Low and Middle-Income Countries (LMICs; Alegría et al., 2018; Sánchez-Castro et al., 2024), which are also home to 90% of the world's adolescents (United Nations, Department of Economic and Social Affairs, Population Division, 2019). Mental health challenges exist in a vicious cycle with each of these other crises: mental health concerns are both exacerbated by and can worsen such challenges (Alegría et al., 2018; Mental Health and Poverty Project, 2008; Sánchez-Castro et al., 2024). Given that 75% of mental disorders emerge by age 24 (Kessler et al., 2005) and 1 in 7 adolescents worldwide experiences a mental disorder (World Health Organization, 2025), it is crucial to find better and more feasible ways to prevent mental health problems and promote mental wellbeing among adolescents.

One promising direction is to integrate universal wellbeing programs into schools. These programs could be particularly beneficial and cost-effective in lower-middle income countries, where a recent review showed that school-based mental health programs can avert a disability-adjusted life year (DALY) for $1.59 and return $346 per $1 invested (Stelmach et al., 2022, Supplemental Material File 4, p. 5). Such programs are gaining attention and commitment from LMIC governments (Barry et al., 2013; Harte & Barry, 2024), with multiple state and national governments integrating attention to wellbeing programming at policy levels (UNICEF, 2023; Ministry of Education, 2020) and multilateral agencies releasing guidelines in support (World Health Organization and United Nations Educational, Scientific and Cultural Organization, 2021). However, despite increasing interest, there are few scalable programs with adequate evidence that have been conducted in LMICs (Barry et al., 2013; Gimba et al., 2020; Bradshaw et al., 2021).

There are three areas in which more research is needed to support full-scale, sustainable government adoption of wellbeing programming among LMIC adolescents: (1) real-world, effectiveness evidence on scalable programs; (2) clarity surrounding high-priority government concerns; and (3) longer-term follow-ups. First, although there are multiple efficacy trials of such programs (e.g., interventions delivered by trained counselors or at smaller scale), evidence supporting effectiveness (effects under "real-world" conditions, including delivered by intended at-scale delivery personnel, such as government teachers, and at larger scale) is relatively rare.

Second, governments often prioritize youth wellbeing when evidence clearly shows that such programs lead to improvements on existing policy mandates. Yet, for several priority areas, including substance use (e.g., tobacco), mental health conditions (e.g., anxiety and depression), and concrete indicators of gender empowerment (e.g., freedom of movement), evidence from school-based wellbeing programs remains limited. Although prior studies have reported positive effects (Barry et al., 2013; Harte & Barry, 2024; Lekagme et al., 2025), findings have lacked the specificity and consistency required to support large-scale adoption. For example, substance use has often been examined in aggregate rather than for specific substances (e.g., Leventhal et al., 2016). Additionally, much of the strongest evidence for these programs has included positive psychological constructs like wellbeing, resilience, and coping, with less attention to mental health symptoms that feature more prominently in government priorities. Further, evaluations of wellbeing programs have more often examined gender attitudes and norms than observable behaviors (Beckwith et al., 2023; Melendez-Torres et al., 2024).

Finally, follow-ups are rare among wellbeing program evaluations in LMICs (Harte & Barry, 2024). This gap is primarily due to resource constraints; indeed, follow-ups are less common in LMICs than in HICs (Franzen et al., 2017; Grande et al., 2023; McGregor et al., 2014). However, in order to fully understand these programs' benefits, it is crucial to evaluate past post-test as far as possible. As adolescence is a key period of transition, and supporting

wellbeing among adolescents can support positive trajectories into adulthood (e.g., Schoon, 2021), longitudinal follow-up is essential.

Youth First is an in-school wellbeing program for adolescents that has served 3.3 million individuals to date in India, Kenya, and Rwanda since its inception in 2009. It is delivered by trained schoolteachers during the school day via weekly sessions across 1-2 years. Youth First is among the most extensively evaluated school-based wellbeing programs in any LMIC. A recent systematic review of school-based wellbeing interventions found that only 3 out of 34 programs had been evaluated more than once through controlled or quasi-experimental designs (Ahmed et al., 2023), whereas Youth First has undergone five controlled or quasi-experimental evaluations. Findings across studies indicate multi-domain improvements, including psychosocial assets and wellbeing (e.g., peer relationships, psychological wellbeing, resilience, social-emotional assets, and self-efficacy), gender attitudes, health-related behaviors and access (e.g., nutrition, handwashing practices, ability to obtain care when needed, and lifetime substance use), and education indicators (e.g., grade progression, time spent on homework, student-teacher relationships; S. Leventhal & Andrew, 2020; Cooper, 2025).

Youth First is evidence-based, culturally-specific, and designed for scale and government adoption: government schoolteachers receive training that is both rigorous yet feasible to integrate within government systems (K. S. Leventhal et al., 2018) and provided with easy-to-follow session manuals (WorldBeing, 2018). Youth First uses a “just right” dosage: enough to be effective but not overwhelming for teachers to complete along with other responsibilities (Goodwin et al., 2021). Youth First, and practices based on its learning and evidence, is in process of being integrated into multiple state and national government systems across India, Kenya, and Rwanda. It is therefore a good candidate for exploring the three evidence gaps described above.

To address these evidence gaps, we conducted a secondary analysis of data gathered during a randomized controlled trial (RCT) of Youth First, evaluating a scaled-up,

government-delivered implementation of the program, during implementation and over one-year post intervention. We hypothesized that students receiving Youth First would show greater improvements in physical, psychological, and social wellbeing (primary outcomes), as well as in other psychosocial outcomes, including self-efficacy, resilience, internalizing symptoms, gender equality attitudes, empowerment, and substance use (secondary outcomes), compared with students at control schools.

Methods

Design

This study involved secondary analysis of data gathered during a cluster-randomized controlled trial (cRCT) conducted in Bihar, India, from 2018-2020. Schools served as the unit of randomization. A total of 70 schools naïve to the Youth First program were randomly allocated to either the intervention arm (n = 40 schools), which implemented the Youth First curriculum, or the control arm (n = 30 schools), which continued with standard education practices.

Schools were randomly assigned to study arms using a computer-generated randomization sequence after eligibility was confirmed and baseline recruitment procedures were initiated. The allocation sequence was generated by a member of the research team who was not involved in program implementation.

Data were collected at three timepoints: baseline (prior to the start of the intervention); first follow-up (approximately 8 months after intervention initiation); and second follow-up (approximately 20 months after intervention initiation). These timepoints were selected to capture both short-term (8-month) and longer-term (20-month) intervention effects.

The duration of exposure to Youth First varied by student grade. Within intervention schools, students enrolled in Grade 8 at baseline received one year of intervention, while students enrolled in Grade 7 at baseline received two consecutive years of intervention.

Outcomes were assessed for both cohorts at all three timepoints, allowing for examination of intervention effects over time across differing exposure levels.

Schools assigned to the control arm continued with usual school activities throughout the study period. The trial followed an intention-to-treat framework, with students' outcomes analyzed according to their school's randomized assignment. Analyses accounted for the clustered structure of the data with students nested within schools.

Participants

Participants were students enrolled in Grades 7 and 8 at baseline in government schools in two districts of Bihar, India: Dharbanga and Patna. All students enrolled in eligible grades at participating schools were invited to participate. Students were excluded only if parental consent was not obtained, they were absent during data collection, or they were unable to provide assent.

At baseline, 5,445 students across 69 schools participated in the study (40 intervention schools and 29 control schools; one control school withdrew from participation prior to baseline). Schools were distributed evenly across districts within study arms. Follow-up data were collected at baseline, first follow-up (8 months) and second follow-up (20 months), with retention rates of 92% and 86%, respectively.

Written parental consent and student assent were obtained prior to participation. All study procedures were approved by Sangath Institutional Review Board.

Sample Size and Power Calculation

The sample size was determined a priori based on the primary outcomes of student physical, psychological, and social wellbeing over the follow-up period. Assuming a standardized effect size of 0.18, an intra-cluster correlation coefficient (ICC) of 0.05, a two-sided alpha of 0.05, and 80% power, we estimated that 70 schools (40 intervention and 30 control) with an average of 45 students per school would be required to detect a statistically significant difference between study arms. These calculations accounted for the cluster-randomized design and anticipated levels of attrition.

The achieved baseline sample included 69 schools and 5,445 students, corresponding to an average of 78.9 students per school, broadly consistent with power calculation assumptions. One school was excluded during baseline data collection by parent request.

Procedure

Baseline data were collected prior to initiating the intervention. Student surveys were administered at each timepoint during regular school hours using paper-based questionnaires in Hindi. Data collectors received training on study protocols, ethical research among adolescents, and procedures for responding to and referring students in distress as needed.

Measures

Primary Outcomes

Primary outcomes assessed adolescents' physical and mental health. These outcomes were measured using scales from the KIDSCREEN-52 questionnaire, a screening instrument for the quality of life of children and adolescents (Ravens-Sieberer, et al., 2014). To assess physical health, the 5-item KIDSCREEN Physical Wellbeing Subscale was used. To assess mental health, the 6-item KIDSCREEN Psychological Wellbeing Subscale and the 6-item KIDSCREEN Social and Peer Support Subscale were used. Across all three scales, items were scored from 1-5, with higher scores indicating higher physical, psychological, or social wellbeing. For each scale, scores were computed as the sum of item responses.

All three KIDSCREEN scales have demonstrated good reliability and validity among adolescents in prior studies and similar settings (Power et al., 2019; Ravens-Sieberer, et al., 2014; Tzavara et al., 2012). In the current sample, internal consistency was acceptable (Cronbach's α = .73-.81), and factor-analytic results were consistent with expected scale structure.

Secondary Outcomes

Secondary outcomes included psychosocial assets and skills, gender equality attitudes, empowerment, and substance use. Internal consistency estimates for secondary outcome measures ranged from Cronbach's α = .65-.83.

Assets and skills. Assets and skills were measured using three validated scales: Schwarzer's Generalized Self-Efficacy Scale to assess self-efficacy (Schwarzer & Jerusalem, 1995); the Connor-Davidson Resilience Scale-10 to assess emotional resilience (Campbell-Sills & Stein, 2007), and the Internalizing Factors Subscale of the Strength and Difficulties Questionnaire to assess internalizing symptoms, including anxiety, depression, and social withdrawal (Goodman, 1997).

Gender equality attitudes. Gender equality attitudes were measured using the Gender Roles and Gender-Based Violence Subscales of the International Center for Research on Women's adaptation of the Gender-Equitable Men Scale (Achyut et al., 2017).

Empowerment. Empowerment was measured using the Empowerment Scale from the Global Early Adolescent Study, measuring freedom of decision-making, movement, and voice (Zimmerman et al., 2019).

Substance use. Substance use was assessed using a single item asking respondents to report the number of times they had smoked tobacco in the previous 30 days. This measure was converted to a binary variable for analysis (smoked one or more times in the past 30 days vs. did not smoke).

Covariates

Pre-specified covariates included student age, sex, marital status, and continuous school enrollment at baseline. Covariates were selected a priori based on theoretical relevance and prior literature.

Measure Adaptation and Translation

All measures were translated and back-translated into Hindi by bilingual English-Hindi speakers. Cognitive interviews and pilot testing were conducted to ensure comprehension

and cultural relevance. Minor adaptations were made to item wording without altering underlying constructs.

Analysis

Quantitative data from student surveys were analyzed using a difference-in-differences (DiD) framework to estimate intervention effects by comparing changes in outcomes from baseline to each follow-up between intervention and control groups. Baseline values of the outcome were not included as covariates; instead, baseline differences between study arms were accounted for through DiD specification.

To account for the cluster-randomized design and the longitudinal structure of the data, analyses employed mixed-effects regression models with observations nested within students and students nested within schools. Models included random intercepts for schools to account for clustering of students within schools and random intercepts for students to account for repeated observations within students over time. Models were estimated using robust standard errors. ICCs for primary outcomes at baseline were approximately 0.08 for physical, psychological, and social wellbeing.

Time was modeled as a categorical variable representing baseline, first follow-up (8 months), and second follow-up (20 months), with baseline serving as the reference category. The interaction between study arm and time point provided the DiD estimate of intervention effects at each follow-up. The primary model can be expressed as:

$$Y_{ijt} = \beta_0 + \beta_1 \text{Treatment}_j + \beta_2 \text{Time}_t + \beta_3 (\text{Treatment}j \times \text{Time}t) + Xij\boldsymbol{\gamma} + u_j + v_i + \varepsilon ijt$$

where Y_{ijt} represents the outcome for student i in school j at time t; Treatment_j indicates assignment to the intervention arm; Time$t$ represents survey wave; and the interaction term estimates the DiD effect. Xij denotes baseline covariates. Random intercepts were included for schools (u_j) and students (v_i), accounting for clustering and repeated measurements, respectively.

For continuous outcomes, linear mixed-effects regression models were used, with results reported as adjusted mean differences and 95% confidence intervals (CIs). Tobacco use, measured as a binary outcome, was analyzed using mixed-effects logistic regression, with results reported as adjusted odds ratios (aORs) and 95% CIs.

Models adjusted for pre-specified covariates. For models analyzing baseline to first follow-up differences, respondents from Grades 7 and 8 were analyzed together. For models analyzing baseline to second follow-up differences, analysis was stratified by respondents' baseline grade to account for different durations of intervention exposure (2 years for Grade 7 and 1 year for Grade 8).

Missing data were addressed using multiple imputation by chained equations under the assumption that data were missing at random (Rubin, 1987). Twelve imputed datasets were generated using an imputation model that included 356 variables, including all outcome measures, baseline covariates, and auxiliary variables associated with missingness.

Statistical significance was assessed using a two-sided $\alpha = 0.05$. All analyses were conducted using R version 4.5.2.

Results

Results are presented regarding one-year effects for both grades, two-year intervention effects among Grade 7 students, and one-year post-intervention effects among Grade 8 students.

Sample Characteristics

At baseline, 5,445 students from 69 schools participated in the study, with 3,054 intervention students and 2,391 control students. Fifty-one percent of students were in Grade 7 and 49% in Grade 8 at baseline. The mean age of participants was 12.74 years (SD = 1.20), and 56% identified as female. The study flow diagram is presented in Figure 1.

Baseline characteristics were generally similar between intervention and control groups (Table 1). Continuous school enrollment at baseline was significantly higher in the control compared to the intervention arm ($p = .041$). This variable was included as a covariate

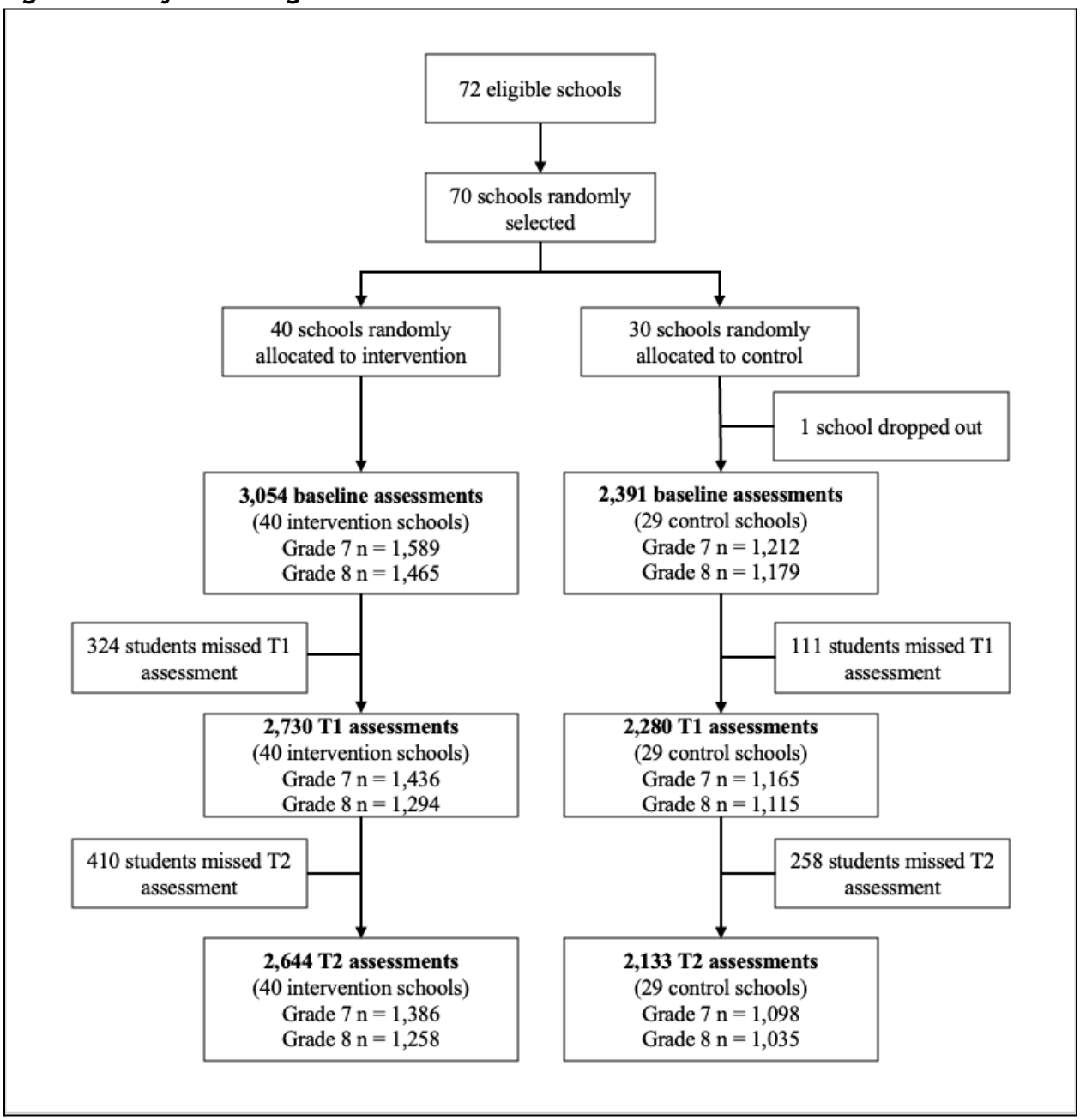
in subsequent analysis. No statistically significant differences were observed across arms for the remaining demographic characteristics.

Table 1. Baseline Background Characteristics of Study Participants, by Study Arm

Characteristic	Total Sample (N = 5,445)	Intervention (n = 3,054)	Control (n = 2,391)	p-value
Age, mean (SD)	12.74 (1.20)	12.70 (1.22)	12.80 (1.17)	.210
Female, n (%)	2,364 (56%)	1,709 (56%)	1,367 (57%)	.676
Grade, n (%)				
Grade 7	2,799 (51%)	1,588 (52%)	1,211 (51%)	.581
Grade 8	2,646 (49%)	1,466 (48%)	1,180 (49%)	.581
Never married at baseline, n (%)	5,276 (97%)	2,943 (96%)	2,333 (98%)	.165
Continuous school enrollment at baseline, n (%)	4,907 (90%)	2,701 (88%)	2,206 (92%)	.041

Notes. Values are means (standard deviations) unless otherwise indicated. Percentages may not sum to 100 due to rounding. *p*-values reflect comparisons between intervention and control groups at baseline and account for clustering at the school level. SD = standard deviation.

Figure 1. Study Flow Diagram.



Notes: T1 = First Follow-Up, T2 = Second Follow-Up

At first follow-up, data were available for 5,010 students (92% retention). Retention at first follow-up was significantly lower in the intervention arm (89%) than in the control arm (95%, p = .000). At second follow-up (approximately 20 months post-intervention initiation), data were available for 4,777 students (88% retention). Retention rates at second follow-up did not differ significantly by study arm.

Implementation Fidelity

Youth First was implemented in the 40 intervention schools using manualized training, mentorship, and session support procedures. Across 119 observed sessions, facilitators demonstrated high implementation fidelity and quality (means = 2.5 and 2.3, respectively, on a 1–3 scale), and 93% of schools completed all 25 sessions. In total, 167 teachers delivered Youth First to 3,581 students in Grades 7 and 8.

One-Year Intervention Effects

Combined Grade 7 and Grade 8 Sample

Across Grades 7 and 8, students in intervention schools demonstrated a statistically significant improvement in social and peer support relative to students in control schools at first follow-up (Table 2). The adjusted difference-in-differences estimate indicated a mean increase of .84 point on the social and peer support scale relative to controls (β = 0.84, 95% CI [0.14, 1.24], p = .000). In contrast, no statistically significant differences were observed between study arms for physical and psychological wellbeing at first follow-up.

Significant positive intervention effects were also observed for several secondary outcomes, including self-efficacy, internalizing symptoms, attitudes toward gender roles, empowerment, and tobacco use. No statistically significant differences were detected for emotional resilience, or attitudes towards gender-based violence. Observed effect sizes were small (see Supplementary Tables S1-S5 for effect sizes).

Table 2. One-Year Intervention Effects on Primary and Secondary Outcomes (Baseline to First Follow-Up)

Outcome	Study Arm	Baseline	First Follow-Up	Difference	Adjusted DiD	Adjusted Odds Ratio	95% CI	p-value
KIDSCREEN Physical Wellbeing Scale	Intervention	18.06	18.35	0.29	-0.24		-0.59, 0.11	.259
(Min-max: 5-25)	Control	18.48	19.01	0.54				
KIDSCREEN Psychological Wellbeing Scale	Intervention	16.69	17.13	0.44	0.09		-0.28, 0.46	.691
(Min-max: 5-25)	Control	17.47	17.82	0.35				
KIDSCREEN Social and Peer Support Scale	Intervention	12.99	13.58	0.59	0.84		0.44, 1.24	.000
(Min-max: 5-25)	Control	13.98	13.73	-0.25				
Schwarzer's Generalized Self-Efficacy Scale	Intervention	31.14	31.71	0.56	0.54		0.09, 1.00	.020
(Min-max: 10-40)	Control	32.35	32.38	0.02				
Connor-Davidson Resilience Scale - 10	Intervention	26.09	26.25	0.16	0.18		-0.42, 0.78	.548
(Min-max: 0-40)	Control	27.56	27.55	-0.01				
SDQ Internalizing Factors Subscale	Intervention	10.63	10.07	-0.56	-0.32		-0.60, -0.04	.026
(Min-max: 0-20)	Control	10.32	10.08	-0.24				
ICRW Gender Roles Subscale	Intervention	25.10	25.41	0.31	-0.91		-1.39, -0.43	.000
(Min-max: 10-40)	Control	24.61	25.38	1.22				
ICRW Gender-Based Violence Subscale	Intervention	13.14	13.55	0.41	0.09		-0.18, 0.36	.511
(Min-max: 5-20)	Control	13.39	13.70	0.32				
GEAS Empowerment Scale	Intervention	33.39	34.42	1.03	1.15		0.52, 1.78	.000
(Min-max: 0-54)	Control	34.69	34.58	-0.11				
Smoked tobacco in previous month	Intervention	7.13	3.31	-3.82		2.35	1.19, 2.68	.009
(percentage)	Control	4.25	3.44	-0.81				

Notes: Estimates represent adjusted difference-in-differences (DiD) effects comparing changes from baseline to first follow-up between intervention and control groups. Models account for clustering at the school level and adjust for baseline age, gender, marital status, and continuous school enrollment. For the SDQ Internalizing Factors subscale and the ICRW Gender Roles and Gender-Based Violence subscales, negative values indicate improvement. For all other measures, positive values indicate improvement. For tobacco use, estimates are adjusted odds ratios. CI = confidence interval.

Grade-Stratified One-Year Effects

When examined separately by grade, intervention effects were observed for both grades, with overlapping and grade-specific patterns across outcomes. Among Grade 7 students, Youth First participation was associated with statistically significant effects on social and peer support, internalizing symptoms, attitudes toward gender roles, and empowerment at first follow-up (Table 3). Effects for other outcomes were not statistically significant. Among Grade 8 students, statistically significant effects were observed for social and peer support, resilience, empowerment, and tobacco use at first follow-up, while effects for other outcomes were not statistically significant (Table 4).

Table 3. One-Year Intervention Effects on Primary and Secondary Outcomes for Grade 7 Students (Baseline to First Follow-Up)

Outcome	Study Arm	Baseline	First Follow-Up	Difference	Adjusted DiD	Adjusted Odds Ratio	95% CI	p-value
KIDSCREEN Physical Wellbeing Scale	Intervention	17.81	18.32	0.51	-0.25		-0.75, 0.25	.384
(Min-max: 5-25)	Control	18.07	18.83	0.76				
KIDSCREEN Psychological Wellbeing Scale	Intervention	16.24	17.09	0.85	0.13		-0.39, 0.66	.638
(Min-max: 5-25)	Control	17.11	17.83	0.72				
KIDSCREEN Social and Peer Support Scale	Intervention	12.83	13.68	0.85	0.75		0.20, 1.29	.017
(Min-max: 5-25)	Control	13.87	13.98	0.10				
Schwarzer's Generalized Self-Efficacy Scale	Intervention	30.52	31.32	0.80	0.58		-0.08, 1.24	.086
(Min-max: 10-40)	Control	31.79	32.01	0.22				
Connor-Davidson Resilience Scale - 10	Intervention	25.51	25.77	0.27	-0.52		-1.38, 0.34	.236
(Min-max: 0-40)	Control	26.75	27.54	0.79				
SDQ Internalizing Factors Subscale	Intervention	10.71	10.15	-0.55	-0.58		-0.98, -0.18	.010
(Min-max: 0-20)	Control	10.45	10.43	-0.03				
ICRW Gender Roles Subscale	Intervention	24.95	25.35	0.40	-1.22		-1.89, -0.55	.000
(Min-max: 10-40)	Control	24.47	26.09	1.62				
ICRW Gender-Based Violence Subscale	Intervention	13.13	13.42	0.28	-0.07		-0.45, 0.31	.705
(Min-max: 5-20)	Control	13.26	13.61	0.35				
GEAS Empowerment Scale	Intervention	32.68	33.99	1.31	1.30		0.39, 2.20	.005
(Min-max: 0-54)	Control	34.03	34.05	0.02				
Smoked tobacco in previous month	Intervention	7.86	3.25	-4.61		1.86	1.50, 2.59	.025
(percentage)	Control	5.56	4.04	-1.52				

Notes: Estimates represent adjusted difference-in-differences (DiD) effects comparing changes from baseline to first follow-up between intervention and control groups for Grade 7. Models account for clustering at the school level and adjust for baseline age, gender, marital status, and continuous school enrollment. For the SDQ Internalizing Factors subscale and the ICRW Gender Roles and Gender-Based Violence subscales, negative values indicate improvement. For all other measures, positive values indicate improvement. For tobacco use, estimates are adjusted odds ratios. CI = confidence interval.

Table 4. One-Year Intervention Effects on Primary and Secondary Outcomes for Grade 8 Students (Baseline to First Follow-Up)

Outcome	Study Arm	Baseline	First Follow-Up	Difference	Adjusted DiD	Adjusted Odds Ratio	95% CI	p-value
KIDSCREEN Physical Wellbeing Scale	Intervention	18.33	18.75	0.06	-0.24		-0.73, 0.25	.397
(Min-max: 5-25)	Control	18.90	19.10	0.30				
KIDSCREEN Psychological Wellbeing Scale	Intervention	17.18	23.86	0.00	0.02		-0.50, 0.54	.064
(Min-max: 5-25)	Control	17.84	23.86	-0.02				
KIDSCREEN Social and Peer Support Scale	Intervention	13.16	19.35	0.30	0.92		0.36, 1.48	.003
(Min-max: 5-25)	Control	14.09	19.16	-0.62				
Schwarzer's Generalized Self-Efficacy Scale	Intervention	31.82	30.93	0.31	0.50		-0.13, 1.12	.122
(Min-max: 10-40)	Control	23.94	31.29	-0.18				
Connor-Davidson Resilience Scale - 10	Intervention	26.72	26.55	0.05	0.90		0.09, 1.71	.030
(Min-max: 0-40)	Control	28.40	26.75	-0.85				
SDQ Internalizing Factors Subscale	Intervention	10.55	9.24	-0.57	-0.11		-0.52, 0.29	.581
(Min-max: 0-20)	Control	10.18	0.42	-0.45				
ICRW Gender Roles Subscale	Intervention	25.26	25.62	0.22	-0.60		-1.29, 0.10	.094
(Min-max: 10-40)	Control	24.75	25.76	0.81				
ICRW Gender-Based Violence Subscale	Intervention	13.15	13.64	0.55	0.27		-0.12, 0.65	.180
(Min-max: 5-20)	Control	13.52	13.81	0.28				
GEAS Empowerment Scale	Intervention	34.16	34.44	0.73	0.98		0.10, 1.87	.030
(Min-max: 0-54)	Control	35.37	34.72	-0.25				
Smoked tobacco in previous month	Intervention	5.65	3.43	-2.22		1.85	1.38, 3.22	.139
(percentage)	Control	2.75	2.63	-0.12				

Notes: Estimates represent adjusted difference-in-differences (DiD) effects comparing changes from baseline to first follow-up between intervention and control groups for Grade 8. Models account for clustering at the school level and adjust for baseline age, gender, marital status, and continuous school enrollment. For the SDQ Internalizing Factors subscale and the ICRW Gender Roles and Gender-Based Violence subscales, negative values indicate improvement. For all other measures, positive values indicate improvement. For tobacco use, estimates are adjusted odds ratios. CI = confidence interval.

Two-Year Intervention Effects Among Grade 7 Students

Two-year intervention effects were examined among students who were in Grade 7 at baseline, who received two consecutive years of intervention. Effects were estimated by comparing changes in outcomes from baseline to second follow-up (approximately 20 months after intervention initiation) between intervention and control schools.

After two years of intervention exposure, students in intervention schools demonstrated statistically significant improvements in social and peer support relative to students in control schools (Table 5). The adjusted difference-in-differences estimate indicated a mean increase of 0.83 points on the social and peer support scale among intervention students ($\beta = 0.83$, 95% CI [0.29, 1.37], $p = .006$). In contrast, no statistically significant effects were observed for physical and psychological wellbeing at second-year follow-up.

Significant two-year effects were also observed for several secondary outcomes, including self-efficacy and empowerment. In contrast, no statistically significant differences were observed for emotional resilience, internalizing symptoms, attitudes towards gender roles or gender-based violence, and tobacco use. Effect sizes across outcomes were small.

While internalizing symptoms, gender role attitudes, and tobacco use showed significant improvements at first follow-up (8 months) among Grade 7 students, these effects were not sustained at second follow-up one year later, whereas effects for self-efficacy emerged over the longer exposure period.

Table 5. Two-Year Intervention Effects on Primary and Secondary Outcomes for Grade 7 Students (Baseline to Second Follow-Up)

Outcome	Study Arm	Baseline	Second Follow-Up	Difference	Adjusted DiD	Adjusted Odds Ratio	95% CI	p-value
KIDSCREEN Physical Wellbeing Scale	Intervention	17.81	18.61	0.80	-0.49		-0.96, 0.02	.080
(Min-max: 5-25)	Control	18.07	19.36	1.29				
KIDSCREEN Psychological Wellbeing Scale	Intervention	16.24	23.78	7.54	0.38		-0.13, 0.89	.189
(Min-max: 5-25)	Control	17.11	24.27	7.16				
KIDSCREEN Social and Peer Support Scale	Intervention	12.83	20.00	7.18	0.83		0.29, 1.37	.006
(Min-max: 5-25)	Control	13.87	20.22	6.34				
Schwarzer's Generalized Self-Efficacy Scale	Intervention	30.52	30.31	-0.22	0.78		0.14, 1.42	.017
(Min-max: 10-40)	Control	31.79	30.79	-1.00				
Connor-Davidson Resilience Scale - 10	Intervention	25.51	26.02	0.51	0.68		-0.11, 1.47	.092
(Min-max: 0-40)	Control	26.75	26.58	-0.17				
SDQ Internalizing Factors Subscale	Intervention	10.71	9.66	-1.05	-0.27		-0.68, 0.13	.185
(Min-max: 0-20)	Control	10.45	9.68	-0.78				
ICRW Gender Roles Subscale	Intervention	24.95	25.75	0.81	-0.39		-1.06, 0.28	.255
(Min-max: 10-40)	Control	24.47	25.67	1.20				
ICRW Gender-Based Violence Subscale	Intervention	13.13	13.62	0.49	-0.19		-0.56, 0.18	.304
(Min-max: 5-20)	Control	13.26	13.94	0.68				
GEAS Empowerment Scale	Intervention	32.68	33.96	1.28	0.86		0.01, 1.70	.047
(Min-max: 0-54)	Control	34.03	34.46	0.42				
Smoked tobacco in previous month	Intervention	7.86	5.21	-2.65		1.73	1.38, 2.53	.907
(percentage)	Control	5.56	3.54	-2.02				

Notes: Estimates represent adjusted difference-in-differences (DiD) effects comparing changes from baseline to second follow-up between intervention and control groups for Grade 8. Models account for clustering at the school level and adjust for baseline age, gender, marital status, and continuous school enrollment. For the SDQ Internalizing Factors Subscale and the ICRW Gender Roles and Gender-Based Violence Subscales, negative values indicate improvement. For all other measures, positive values indicate improvement. For tobacco use, estimates are adjusted odds ratios. CI = confidence interval.

One-Year Post-Intervention Effects Among Grade 8 Students

At the second follow-up, which was one year post-intervention completion for Grade 8 students, Grade 8 students in intervention schools exhibited significant gains in both psychological wellbeing and social and peer support compared with students in control schools (Table 6). The adjusted difference-in-differences estimates indicated a mean increase of 0.66 points in psychological wellbeing (95% CI [0.14, 1.18], p = .022) and a mean increase of 1.11 points in social and peer support (95% CI [0.55, 1.67], p = .019) relative to controls. No statistically significant effects were observed for physical health at this timepoint.

Significant positive one-year post-intervention effects were also observed for several secondary outcomes, including self-efficacy, emotional resilience, internalizing symptoms, empowerment, and tobacco use. No statistically significant differences were detected for attitudes toward gender roles or gender-based violence at this time point.

Effects for social and peer support, emotional resilience, and empowerment, which were evident at the one-year follow-up, persisted one-year post-intervention, while improvements in psychological wellbeing, self-efficacy, internalizing symptoms, and tobacco use were observed only at the post-intervention assessment. Overall, effect sizes across outcomes were small.

Table 6. One-Year Post-Intervention Effects on Primary and Secondary Outcomes for Grade 8 Students (Baseline to Second Follow-Up)

Outcome	Study Arm	Baseline	Second Follow-Up	Difference	Adjusted DiD	Adjusted Odds Ratio	95% CI	p-value
KIDSCREEN Physical Wellbeing Scale	Intervention	18.33	18.75	0.42	0.22		-0.25, 0.68	.415
(Min-max: 5-25)	Control	18.90	19.10	0.20				
KIDSCREEN Psychological Wellbeing Scale	Intervention	17.18	23.86	6.67	0.66		0.14, 1.18	.022
(Min-max: 5-25)	Control	17.84	23.86	6.01				
KIDSCREEN Social and Peer Support Scale	Intervention	13.16	19.35	6.18	1.11		0.55, 1.67	.000
(Min-max: 5-25)	Control	14.09	19.16	5.08				
Schwarzer's Generalized Self-Efficacy Scale	Intervention	31.82	30.93	-0.88	0.76		0.14, 1.40	.019
(Min-max: 10-40)	Control	23.94	31.29	-1.65				
Connor-Davidson Resilience Scale - 10	Intervention	26.72	26.55	-0.17	1.49		0.70, 2.28	.000
(Min-max: 0-40)	Control	28.40	26.75	-1.65				
SDQ Internalizing Factors Subscale	Intervention	10.55	9.24	-1.31	-0.55		-0.96, -0.14	.008
(Min-max: 0-20)	Control	10.18	9.42	-0.76				
ICRW Gender Roles Subscale	Intervention	25.26	25.62	0.35	-0.66		-1.34, 0.03	.061
(Min-max: 10-40)	Control	24.75	25.76	1.01				
ICRW Gender-Based Violence Subscale	Intervention	13.15	13.64	0.49	0.19		-0.19, 0.58	.326
(Min-max: 5-20)	Control	13.52	13.81	0.30				
GEAS Empowerment Scale	Intervention	34.16	34.44	0.28	0.99		0.15, 1.82	.028
(Min-max: 0-54)	Control	35.37	34.72	-0.65				
Smoked tobacco in previous month	Intervention	5.65	5.26	-0.39		1.63	1.32, 2.36	.013
(percentage)	Control	2.75	5.07	2.32				

Notes: Estimates represent adjusted difference-in-differences (DiD) effects comparing changes from baseline to second follow-up between intervention and control groups for Grade 8. Models account for clustering at the school level and adjust for baseline age, gender, marital status, and continuous school enrollment. For the SDQ Internalizing Factors Subscale and the ICRW Gender Roles and Gender-Based Violence Subscales, negative values indicate improvement. For all other measures, positive values indicate improvement. For tobacco use, estimates are adjusted odds ratios. CI = confidence interval.

Discussion

This study, based on secondary data analysis from a longitudinal cRCT of Youth First among Bihar, India adolescents, represents one of the largest RCTs to date of an in-school wellbeing intervention for middle school students in an LMIC (e.g., compared to 34 LMIC studies identified in a recent systematic review of school-based wellbeing interventions, the current study is larger than 85% of the studies for all age groups and 100% of middle-school focused studies; Ahmed et al., 2023). Youth First improved a host of outcomes vs. control (social and peer support, psychological wellbeing, self-efficacy, emotional resilience, internalizing factors, gender attitudes, empowerment, and tobacco use) and one year post-intervention, the program continued to benefit students. In some cases (psychological wellbeing and self-efficacy), benefits for students intensified over time. In addition, this study contributes insights into three priority areas: (1) *real-world, scaled effectiveness;* (2) *key priority domains;* and (3) *longer-term-follow-up*.

Real-world, large-scale effectiveness

Previous studies of wellbeing programs in schools have often evaluated interventions delivered by those outside of the intended at-scale system (e.g., counselors) rather than by the intended at-scale providers (e.g., teachers; Barry et al., 2013; Bradshaw et al., 2021). Many studies have also been on a much smaller scale. This evidence gap exists in part because there are very few programs that are scalable and fully prepared for integration into government systems in LMICs (Gimba et al., 2020; Harte & Barry, 2024); in addition, resources required for conducting and evaluating larger, government-integrated interventions are scarce. In contrast, this evaluation focused on 5,445 students, examining an intervention that has been conducted among millions of students to date via government systems (WorldBeing, 2026). In addition, the program was conducted in its intended form, including scalable and adoptable teacher training and manuals, and within the school day.

In other studies of scaled-up health interventions, effects have attenuated vs. effects prior to scale-up by approximately 40% (Lane et al., 2021). Shifting delivery personnel can also

have unintended consequences: for example, a large-scale effectiveness trial in India of an adolescent health and wellbeing program, the SEHER trial, showed little effect from a teacher-delivered intervention in schools, yet showed effects when counselor-delivered (Shinde et al., 2020). Unfortunately, counselors are not yet reliably present across schools in India, impeding scalability of that model via existing systems. It is therefore highly encouraging that Youth First benefited students across a host of outcomes even when conducted at larger scale and utilizing existing government delivery personnel (teachers).

Although this study was not designed to tease out the reasons for Youth First's ability to achieve effects when teacher-delivered and at larger scale, it is likely that design and implementation elements are responsible, including Youth First's emphasis on transformative teacher training, its "just right" program intensity, and its "inside-out" concept progression. Training teachers to conduct wellbeing programming is crucial, particularly in LMICs, as the pedagogical and facilitation techniques required are often distinct from what teachers are used to and have been trained for previously (Schonert-Reichl, 2017; Murano et al., 2019; Mastercard Foundation, 2020; Bremner et al., 2023). As such, specific teacher training strategies are crucial. Youth First comprises not only student sessions but also rigorous, manualized, yet imminently feasible training for teachers that has been honed over the course of 15 years to be delivered over only a few days of in-person time (K. S. Leventhal et al., 2018). In other studies of Youth First, Youth First's specific teacher training approach has been identified as one of the main implementation drivers of success (Cooper, 2025). Having teachers trained on Youth First concepts also benefits students even if no student sessions are delivered: one study that adapted and provided Youth First teacher training in Punjab, India, showed that students whose teachers were trained improved 0.10 standard deviations on their end-of-year mathematics exams, even without student sessions (Kaur, 2024).

In addition, Youth First is designed to include evidence-based activities in an adequate duration for program effects, yet is not so intensive as to overwhelm teachers and students with content in an already crowded curriculum and timetable. As others have pointed out

(Goodwin et al., 2021), having enough time in the school day, particularly given the other demands on teachers' time in LMICs, is a major barrier to integrating wellbeing programming into schools. However, Youth First is designed to be delivered across approximately 15-25 sessions per school year; an intensity that allows for coverage of many key concepts, yet can fit into the school schedule with time to spare for holidays, exams, and other unpredictable time constraints.

Finally, Youth First adheres carefully to an "inside-out" concept progression, in which students focus first on intrapersonal assets and skills, such as self-concept and emotion regulation, and then move on to focus on interpersonal assets and skills, such as conflict resolution and assertive communication. This progression has been shown in higher-income country settings to support success of other wellbeing programs in schools (Cipriano et al., 2023). This study aligns with and furthers this previous research by showing that a program adhering to this practice is also effective in an LMIC setting.

Although this study suggests that these design and implementation elements may be important, future studies should investigate these ingredients as specific moderators across multiple school-based wellbeing programs in LMICs, such as through meta-analytic methods (e.g., Durlak et al., 2011, 2015; Sklad et al., 2012; Taylor et al., 2017). Identifying which design and implementation elements are necessary and sufficient for program effects will be crucial to guiding government adoption.

Key Priority Areas: Mental Health Symptoms, Substance Use, and Gender Empowerment

In India, some of the greatest priorities for state and national governments, including Ministries of Education and Health, include gender empowerment, mental disorders, and substance use (particularly tobacco; e.g., Ministry of Women and Child Development, 2015; Ministry of Health and Family Welfare, 2016; Ministry of Education, 2021). Unfortunately, there is limited evidence supporting scalable programs that are effective for these outcomes, leaving ministries with little clear direction. For example, one intervention among adolescents in Haryana, India, showed improvements in gender attitudes and behavior vs. control (Dhar et

al., 2022), but was an intensive, multi-year intervention, making government integration more difficult.

By contrast, the current study shows that it is possible to significantly improve adolescent mental disorder symptoms, gender empowerment behaviors, and tobacco use via an intervention that can be feasibly delivered by teachers within the school day, requiring only a few days of in-person training for teachers. The fact that this intervention improved all of these domains through a single intervention is even more efficient, such that multiple priority areas can be improved via one effort.

Longer-term evidence

This study showed that one year post-intervention, those who had received Youth First still experienced better outcomes than controls on a number of different indicators (psychological wellbeing, peer and social support, self-efficacy, emotional resilience, internalizing symptoms, empowerment, and tobacco use). Follow-up assessments or longitudinal studies like this remain uncommon in evaluations of school-based wellbeing programs. A large meta-analysis of 252 interventions across 53 countries found that only 11% included follow-up beyond four months (Cipriano et al., 2023). Longitudinal follow-up is particularly rare in studies involving adolescents in low- and middle-income countries due to resource and implementation constraints (Barry et al., 2013; Tol et al., 2011). Thus, the fact that Youth First produced improvements at one year post-intervention is highly encouraging.

This study contributes to an emerging body of research that suggests that, in some cases, wellbeing interventions' effects may not only persist but intensify over time (Brunwasser & Garber, 2016; Taylor et al., 2017; Cipriano et al., 2023). This possibility is also supported via theories, such as Fredrickson's "broaden-and-build" theory, in which supporting positive outcomes leads to more positive outcomes in an upward spiral (Fredrickson, 2004). If true, it is likely that efforts to quantify the true benefits of wellbeing interventions should integrate longer-term follow-up timepoints. In addition, this

conceptualization further supports conducting such interventions specifically during adolescence in order to set youth on a positive trajectory towards future success.

Limitations

The original cRCT was designed to include four data collection waves: baseline, first follow-up (~8 months), second follow-up (~20 months), and third follow-up (~32 months). The third follow-up was intended to assess longer-term post-intervention effects, including one-year post-intervention outcomes for the Grade 7 cohort, which received two years of the intervention, and two-year post-intervention outcomes for the Grade 8 cohort, which received one year of the intervention. However, prolonged school closures resulting from the COVID-19 pandemic necessitated cancellation of the third follow-up, limiting the ability to examine longer-term effects.

A second limitation relates to the interpretation of intervention effects across time points and cohorts. Although findings for outcomes such as social and peer support and empowerment were generally consistent, effects for other outcomes (including psychological wellbeing, self-efficacy, emotional resilience, internalizing symptoms, gender-related attitudes, and tobacco use) varied across follow-ups and between Grade 7 and 8 students. The study was not designed to identify mechanisms underlying these temporal and cohort-specific differences, and future research is needed to better understand factors that may contribute to variation in timing and durability of intervention effects.

Conclusion

In addressing mental health needs of adolescents worldwide, school-based interventions remain promising. This study provides evidence applicable to key knowledge gaps, showing that a scalable, universal school-based mental health promotion program shows real-world effectiveness on a host of mental health, empowerment, and substance use outcomes for adolescents in India, and that effects persist and in some cases increase over time after program completion. These points are of critical importance to governments with priority areas such as gender empowerment, substance use, and mental health among

adolescents, as well as to practitioners, school administrators, and academics seeking to find better, more scalable ways to support adolescent wellbeing in the longer term.

References

Achyut, P., Bhatla, N., Kumar, U., Verma, H., Bhattacharya, S., Singh, G., & Verma, R. (2017). *Changing course: Implementation and evaluation of the Gender Equity Movement in Schools (GEMS) program in specific sites – Vietnam, India, and Bangladesh*. International Center for Research on Women. https://www.ungei.org/sites/default/files/Changing-Course-Implementation-and-Evaluation-of-the-Gender-Equity-Movement-in-Schools-GEMS-program-in-specific-sites-Vietnam-India-and-Bangladesh-2017-eng.pdf

Ahmed, S. K., Dabrowski, A., Dix, K., & Carslake, T. (2023). *School-based interventions that support mental health and psychosocial wellbeing in low- and middle-income countries*. Australian Council for Educational Research. https://doi.org/10.37517/978-1-74286-714-4

Alegría, M., NeMoyer, A., Falgàs Bagué, I., Wang, Y., & Alvarez, K. (2018). Social Determinants of Mental Health: Where We Are and Where We Need to Go. *Current Psychiatry Reports*, *20*(11), 95. https://doi.org/10.1007/s11920-018-0969-9

Barry, M. M., Clarke, A. M., Jenkins, R., & Patel, V. (2013). A systematic review of the effectiveness of mental health promotion interventions for young people in low and middle income countries. *BMC Public Health*, *13*(1), 835. https://doi.org/10.1186/1471-2458-13-835

Beckwith, S., Mengmeng, L., Barker, K. M., Gayles, J., Kagesten, A. E., Lundgren, R., Dintrans, P. V., Wilopo, S. A., & Moreau, C. (2023). The impacts of two gender-transformative interventions on early adolescent gender norms perceptions: A difference-in-difference analysis. *The Journal of Adolescent Health: Official Publication of the Society for Adolescent Medicine*, *73*(1S), S55–S64. https://doi.org/10.1016/j.jadohealth.2022.09.033

Bradshaw, M., Gericke, H., Coetzee, B. J., Stallard, P., Human, S., & Loades, M. (2021). Universal school-based mental health programmes in low- and middle-income countries: A

systematic review and narrative synthesis. *Preventive Medicine*, *143*(106317). https://doi.org/10.1016/j.ypmed.2020.106317

Bremner, N., Sakata, N., & Cameron, L. (2023). Teacher education as an enabler or constraint of learner-centred pedagogy implementation in low-to middle-income countries. *Teaching and Teacher Education*, *126*, 104033. https://doi.org/10.1016/j.tate.2023.104033

Brunwasser, S. M., & Garber, J. (2016). Programs for the prevention of youth depression: Evaluation of efficacy, effectiveness, and readiness for dissemination. *Journal of Clinical Child & Adolescent Psychology*, *45*(6). https://doi.org/10.1080/15374416.2016.1236727

Campbell-Sills, L., & Stein, M. (2007). Psychometric analysis and refinement of the Connor-Davidson Resilience Scale (CD-RISC): Validation of a 10-item measure of resilience. *Journal of Traumatic Stress*, *20*(6), 1019–1028.

Cipriano, C., Strambler, M. J., Naples, L. H., Cheyeon, H., Kirk, M., Wood, M., Sehgal, K., Zieher, A. K., Eveleigh, A., McCarthy, M., Funaro, M., Annett, P., Chow, J. C., & Durlak, J. (2023). The state of evidence for social and emotional learning: A contemporary meta-analysis of universal school-based SEL interventions. *Child Development*, *94*(5), 1181–1204. https://doi.org/10.1111/cdev.13968

Cooper, P. L. (2025). *Youth First Rwanda Realist Evaluation Summary Report*. WorldBeing. https://cdn.prod.website-files.com/64e7405a819e87ea609d70ef/6939ae37c862f621d8390c25_YFR%20Realist%20Evaluation%20Summary%20Report_FINAL_2025.12.10.pdf

Dhar, D., Jain, T., & Jayachandran, S. (2022). Reshaping Adolescents' Gender Attitudes: Evidence from a School-Based Experiment in India. *American Economic Review*, *112*(3), 899–927. https://doi.org/10.1257/aer.20201112

Durlak, J. A., Domitrovich, C. E., Weissberg, R. P., & Gullotta, T. P. (2015). *Handbook of social and emotional learning: Research and practice*. The Guilford Press.

Durlak, J. A., Weissberg, R. P., Dymnicki, A. B., & Schellinger, K. B. (2011). The impact of enhancing students' social and emotional learning: A meta-analysis of school-based

universal interventions. *Child Development*, *82*(1), 405–432. https://doi.org/10.1111/j.1467-8624.2010.01564.x

Franzen, S., Chandler, C., & Lang, T. (2017). Health research capacity development in low and middle income countries: Reality or rhetoric? A systematic meta-narrative review of the qualitative literature. *BMJ Open*, *7*(1). https://doi.org/10.1136/bmjopen-2016-012332

Fredrickson, B. L. (2004). The broaden–and–build theory of positive emotions. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *359*(1449), 1367–1377. https://doi.org/10.1098/rstb.2004.1512

Gimba, S. M., Harris, P., Saito, A., Udah, H., Martin, A., & Wheeler, A. J. (2020). The modules of mental health programs implemented in schools in low- and middle-income countries: Findings from a systematic literature review. *BMC Public Health*, *20*(1), 1581. https://doi.org/10.1186/s12889-020-09713-2

Goodman, R. (1997). The Strengths and Difficulties Questionnaire: A research note. *Journal of Child Psychology and Psychiatry*, *38*(5), 581–586.

Goodwin, J., Behan, L., & O'Brien, N. (2021). Teachers' views and experiences of student mental health and well-being programmes: A systematic review. *Journal of Child & Adolescent Mental Health*, *33*(1–3), 55–74. https://doi.org/10.2989/17280583.2023.2229876

Grande, A. J., Hoffmann, M. S., Evans-Lacko, S., Ziebold, C., de Miranda, C. T., Mcdaid, D., Tomasi, C., & Ribeiro, W. S. (2023). Efficacy of school-based interventions for mental health problems in children and adolescents in low and middle-income countries: A systematic review and meta-analysis. *Frontiers in Psychiatry*, *13*. https://doi.org/10.3389/fpsyt.2022.1012257

Harte, P., & Barry, M. M. (2024). A scoping review of the implementation and cultural adaptation of school-based mental health promotion and prevention interventions in low-and middle-income countries. *Cambridge Prisms: Global Mental Health*, *12*(11:e55). https://doi.org/10.1017/gmh.2024.48

Kałwak, W., Weziak-Bialowolska, D., Wendołowska, A., Bonarska, K., Sitnik-Warchulska, K., Bańbura, A., Czyżowska, D., Gruszka, A., Opoczyńska-Morasiewicz, M., & Izydorczyk, B. (2024). Young adults from disadvantaged groups experience more stress and deterioration in mental health associated with polycrisis. *Scientific Reports*, *14*(8757). https://doi.org/10.1038/s41598-024-59325-8

Kaur, J. (2024). *How Much Do I Matter?: Teacher Self-Beliefs, Effort, and Student Learning* [Working Paper]. https://www.jalnidhkaur.com/wp-content/uploads/2023/11/Kaur_JMP_Nov21.pdf

Kessler, R. C., Berglund, P., Demler, O., Jin, R., Merikangas, K. R., & Walters, E. E. (2005). Lifetime prevalence and age-of-onset distribution of DSM-IV disorders in the National Comorbidity Survey Replication. *Arch Gen Psychiatry*, *62*(6), 593–602. https://doi.org/10.1001/archpsyc.62.6.593

Lane, C., McCrabb, S., Nathan, N., Naylor, P.-J., Bauman, A., Milat, A., Lum, M., Sutherland, R., Byaruhanga, J., & Wolfenden, L. (2021). How effective are physical activity interventions when they are scaled-up: A systematic review. *International Journal of Behavioral Nutrition and Physical Activity*, *18*(1), 16. https://doi.org/10.1186/s12966-021-01080-4

Lekagme, R. B., Jain, R., Sheen, J., Solanki, P., Zhou, Y., Romero, L., Barry, M. M., Chen, L., Karim, M. N., & Ilic, D. (2025). Systematic review and meta-analysis of the effectiveness of whole-school interventions promoting mental health and preventing risk behaviours in adolescence. *Journal of Youth and Adolescence*, *54*(2), 271–289. https://doi.org/10.1007/s10964-025-02135-6

Leventhal, K. S., Andrew, G., Collins, C. S., DeMaria, L., Shanker, H., & Leventhal, S. (2018). Training School Teachers to promote Mental and Social Well-being in Low and Middle Income Countries: Lessons to facilitate scale-up from a Participatory Action Research Trial of Youth First in India. *International Journal of Emotional Education*, *10*(2), 42–58.

Leventhal, K. S., DeMaria, L. M., Gillham, J. E., Andrew, G., Peabody, J., & Leventhal, S. (2016). A psychosocial resilience curriculum provides the “missing piece” to boost adolescent

physical health: A randomized controlled trial of Girls First in India. *Social Science & Medicine*, *161*, 37–46. https://doi.org/10.1016/j.socscimed.2016.05.004

Leventhal, S., & Andrew, G. (2020). *Girls First: Summary Key Research Findings, 2014-2020*. https://cdn.prod.website-files.com/64e7405a819e87ea609d70ef/64f86d7a0c82165ae8f58d12_CorStone_Girls-First_Key-Findings-Research_2014-20.pdf

Mastercard Foundation. (2020). *Secondary Education in Africa*. Mastercard Foundation. https://doi.org/10.15868/socialsector.35972

McGregor, S., Henderson, K. J., & Kaldor, J. M. (2014). How are health research priorities set in low and middle income countries? A systematic review of published reports. *PLoS ONE*, *9*(10). https://doi.org/10.1371/journal.pone.0108787

Melendez-Torres, G., Orr, N., Farmer, C., Shaw, N., Chollet, A., Rizzo, A., Kiff, F., Rigby, E., Hagell, A., Priolo Filho, S. R., Taylor, B., Young, H., Bonell, C., & Berry, V. (2024). School-based interventions to prevent dating and relationship violence and gender-based violence: STOP-DRV-GBV systematic review. *Public Health Research*, *12*(3). https://doi.org/10.3310/KTWR6997

Mental Health and Poverty Project. (2008). *Policy brief 1: Breaking the vicious cycle of mental ill-health and poverty.* DFID. https://assets.publishing.service.gov.uk/media/57a08bb8ed915d622c000e5d/MHPB1.pdf

Ministry of Education. (2020). *National Education Policy 2020*. Government of India.

Ministry of Education. (2021). *Tobacco-Free Educational Institutions (ToFEI) implementation manual and guidelines*. Government of India. https://tobaccofreeindia.in/Dashboard/ToFEIGuidelines

Ministry of Health and Family Welfare. (2016). *National Mental Health Survey of India, 2015–16: Prevalence, patterns and outcomes*. Government of India & National Institute of Mental Health and Neuro Sciences. https://main.mohfw.gov.in / https://nimhans.ac.in

Ministry of Women and Child Development. (2015). *Beti Bachao Beti Padhao: Operational manual*. Government of India. https://spniwcd.wcd.gov.in/uploads/pdf/1710061539_9xKA1jy6eB.pdf

Murano, D., Way, J. D., Martin, J. E., Walton, K. E., Anguiano-Carrasco, C., & Burrus, J. (2019). The need for high-quality pre-service and inservice teacher training in social and emotional learning. *Journal of Research in Innovative Teaching & Learning*, *12*(2), 111–113. https://doi.org/10.1108/JRIT-02-2019-0028

Power, R., Akhter, R., Muhit, M., Wadud, S., Heanoy, E., Karim, T., Badawi, N., & Khandaker, G. (2019). Cross-cultural validation of the Bengali version KIDSCREEN-27 quality of life questionnaire. *BMC Pediatrics*, *19*(1), 19. https://doi.org/10.1186/s12887-018-1373-7

Ravens-Sieberer, U., Herdman, M., Devine, J., Otto, C., Bullinger, M., Rose, M., & Klasen, F. (2014). The European KIDSCREEN approach to measure quality of life and well-being in children: Development, current application, and future advances. *Quality of Life Researc*, *23*(3), 791–803. https://doi.org/10.1007/s11136-013-0428-3

Rubin, D. B. (1987). *Multiple imputation for nonresponse in surveys*. John Wiley & Sons, Inc. https://doi.org/10.1002/9780470316696

Sánchez-Castro, J. C., Pilz González, L., Arias-Murcia, S. E., Mahecha-Bermeo, V. A., Stock, C., & Heinrichs, K. (2024). Mental health among adolescents exposed to social inequality in Latin America and the Caribbean: A scoping review. *Frontiers in Public Health*, *12*, 1342361. https://doi.org/10.3389/fpubh.2024.1342361

Schonert-Reichl, K. A. (2017). Social and Emotional Learning and Teachers. *The Future of Children*, *27*(1), 137–155.

Schoon, I. (2021). A Socioecological Developmental Systems Approach for the Study of Human Resilience. In *Multisystemic Resilience*. Oxford University Press. https://doi.org/10.1093/oso/9780190095888.003.0019

Schwarzer, R., & Jerusalem, M. (1995). Generalized Self-Efficacy scale. In J. Weinman & S. Wright (Eds.), *Measures in health psychology: A user's portfolio. Causal and control beliefs* (pp. 35–37). NFER-NELSON.

Shinde, S., Weiss, H. A., Khandeparkar, P., Pereira, B., Sharma, A., Gupta, R., Ross, D. A., Patton, G., & Patel, V. (2020). A multicomponent secondary school health promotion intervention and adolescent health: An extension of the SEHER cluster randomised controlled trial in Bihar, India. *PLoS Medicine*, *17*(2), e1003021. https://doi.org/10.1371/journal.pmed.1003021

Sklad, M., Diekstra, R., De Ritter, M., Ben, J., & Gravesteijn, C. (2012). Effectiveness of school-based universal social, emotional, and behavioral programs: Do they enhance students' development in the area of skill, behavior, and adjustment? *Psychology in the Schools*, *49*(9), 892–909. https://doi.org/10.1002/pits.21641

Stelmach, R., Kocher, E. L., Kataria, I., Jackson-Morris, A. M., Saxena, S., & Nugent, R. (2022). The global return on investment from preventing and treating adolescent mental disorders and suicide: A modelling study. *BMJ Global Health*, *7*(6), e007759. https://doi.org/10.1136/bmjgh-2021-007759

Taylor, R. D., Oberle, E., Durlak, J. A., & Weissberg, R. P. (2017). Promoting Positive Youth Development Through School-Based Social and Emotional Learning Interventions: A Meta-Analysis of Follow-Up Effects. *Child Development*, *88*(4), 1156–1171. https://doi.org/10.1111/cdev.12864

Tol, W. A., Barbui, C., Galappatti, A., Silove, D., Betancourt, T., Souza, R., Golaz, A., & van Ommeren, M. (2011). Mental health and psychosocial support in humanitarian settings: Linking practice and research. *The Lancet*, *378*(9802), 1581–1591. https://doi.org/10.1016/S0140-6736(11)61094-5

Tzavara, C., Tzonou, A., Zervas, I., Ravens-Sieberer, U., Dimitrakaki, C., & Tountas, Y. (2012). Reliability and validity of the KIDSCREEN-52 health-related quality of life questionnaire

in a Greek adolescent population. *Annals of General Psychiatry*, *11*(1), 3. https://doi.org/10.1186/1744-859X-11-3

UNICEF. (2023). *Global Annual Results Report 2022: Goal Area 2—Every child, including adolescents, learns and acquires skills for the future (UNICEF Strategic Plan 2022-2025)*. United Nations Children's Fund. https://www.unicef.org/media/142921/file/Global%20annual%20results%20report%202022%3A%20Goal%20area%202.pdf?utm_source=chatgpt.com

United Nations, Department of Economic and Social Affairs, Population Division. (2019). *World population prospects 2019: Highlights* (ST/ESA/SER.A/423)). United Nations. https://api.pageplace.de/preview/DT0400.9789210042352_A37800241/preview-9789210042352_A37800241.pdf

World Federation of Public Health Associations. (2025). *Youth mental health crisis: Causes, solutions & global call to action*. https://www.wfpha.org/youth-mental-health-crisis/#:~:text=Some%20solutions%20to%20the%20youth%20mental%20health,care**%20*%20Teletherapy%20and%20mental%20health%20apps

World Health Organization. (2025). *Mental health of adolescents: Fact sheet*. World Health Organization. https://www.who.int/news-room/fact-sheets/detail/adolescent-mental-health

World Health Organization and United Nations Educational, Scientific and Cultural Organization. (2021). *Making every school a health-promoting school: Global standards and indicators*. WHO. https://iris.who.int/handle/10665/341907

WorldBeing. (2018). *Youth First India Facilitator Manual* [Facilitator Manual]. WorldBeing.

WorldBeing. (2026). *Scaling to reach millions*. WorldBeing. https://www.worldbeing.org/

Zimmerman, L. A., Moreau, C., Wilopo, S., & Blum, R. (2019). Measuring agency as a dimension of empowerment among young adolescents globally: Findings from the Global Early

Adolescent Study. *SSM-Population Health*, *8*(100454).

https://doi.org/10.1016/j.ssmph.2019.100454

Supplementary Tables

Supplementary Table S1. Standardized Effect Sizes (Cohen's d) for Intervention Effects at One Year

Outcome Domain	Outcome Measure	Grade	Comparison Period	DiD Estimate (β)	SE	95% CI	Cohen's *d*	Interpretation
Primary Outcomes								
Wellbeing	Physical Health	7+8	Baseline → 8 mo	-0.24	0.18	-0.59, 0.11	-0.05	Negligible
Wellbeing	Psychological Wellbeing	7+8	Baseline → 8 mo	0.09	0.19	-0.28, 0.46	0.02	Negligible
Wellbeing	Social & Peer Support	7+8	Baseline → 8 mo	0.84	0.20	0.44, 1.24	0.16	Small
Secondary Outcomes								
Assets	Self-Efficacy	7+8	Baseline → 8 mo	0.54	0.23	0.09, 1.00	0.09	Small
Assets	Emotional Resilience	7+8	Baseline → 8 mo	0.18	0.30	-0.42, 0.78	0.02	Negligible
Mental health	Internalizing Symptoms	7+8	Baseline → 8 mo	-0.32	0.15	-0.60, -0.04	-0.09	Small
Gender	Gender Roles Attitudes	7+8	Baseline → 8 mo	-0.91	0.25	-1.39, -0.43	-0.14	Small
Gender	Gender-Based Violence Attitudes	7+8	Baseline → 8 mo	0.09	0.14	-0.18, 0.36	0.03	Negligible
Empowerment	Freedom of Voice and Movement	7+8	Baseline → 8 mo	1.15	0.32	0.52, 1.78	0.14	Small

Notes. Effect sizes (Cohen's *d*) were calculated by dividing the adjusted difference-in-differences estimate by the pooled baseline standard deviation of the outcome. Negative values indicate improvements for outcomes where lower scores reflect better functioning (e.g., internalizing symptoms, gender roles, and gender-based violence). Estimates are from mixed-effects models adjusted for baseline covariates and clustering at the school level.

Supplementary Table S2. Standardized Effect Sizes (Cohen's d) for Intervention Effects at One Year - Grade 7

Outcome Domain	Outcome Measure	Grade	Comparison Period	DiD Estimate (β)	SE	95% CI	Cohen's *d*	Interpretation
Primary Outcomes								
Wellbeing	Physical Health	7	Baseline → 8 mo	-0.25	0.25	-0.75, 0.25	-0.05	Negligible
Wellbeing	Psychological Wellbeing	7	Baseline → 8 mo	0.13	0.27	-0.39, 0.66	0.03	Negligible
Wellbeing	Social & Peer Support	7	Baseline → 8 mo	0.75	0.28	0.20, 1.29	0.14	Small
Secondary Outcomes								
Assets	Self-Efficacy	7	Baseline → 8 mo	0.58	0.34	-0.08, 1.24	0.09	Negligible
Assets	Emotional Resilience	7	Baseline → 8 mo	-0.52	0.44	-1.38, 0.34	-0.07	Negligible
Mental health	Internalizing Symptoms	7	Baseline → 8 mo	-0.58	0.21	-0.98, -0.18	-0.16	Small
Gender	Gender Roles Attitudes	7	Baseline → 8 mo	-1.22	0.34	-1.89, -0.55	-0.19	Small
Gender	Gender-Based Violence Attitudes	7	Baseline → 8 mo	-0.07	0.19	-0.45, 0.31	-0.02	Negligible
Empowerment	Freedom of Voice and Movement	7	Baseline → 8 mo	1.30	0.46	0.39, 2.20	0.15	Small

Notes. Effect sizes (Cohen's *d*) were calculated by dividing the adjusted difference-in-differences estimate by the pooled baseline standard deviation of the outcome. Negative values indicate improvements for outcomes where lower scores reflect better functioning (e.g., internalizing symptoms, gender roles, and gender-based violence). Estimates are from mixed-effects models adjusted for baseline covariates and clustering at the school level.

Supplementary Table S3. Standardized Effect Sizes (Cohen's d) for Intervention Effects at One Year - Grade 8

Outcome Domain	Outcome Measure	Grade	Comparison Period	DiD Estimate (β)	SE	95% CI	Cohen's *d*	Interpretation
Primary Outcomes								
Wellbeing	Physical Health	8	Baseline → 8 mo	-0.24	0.25	-0.73, 0.25	-0.05	Negligible
Wellbeing	Psychological Wellbeing	8	Baseline → 8 mo	0.02	0.27	-0.50, 0.54	0.00	Negligible
Wellbeing	Social & Peer Support	8	Baseline → 8 mo	0.92	0.29	0.36, 1.48	0.18	Small
Secondary Outcomes								
Assets	Self-Efficacy	8	Baseline → 8 mo	0.50	0.32	-0.13, 1.12	0.08	Negligible
Assets	Emotional Resilience	8	Baseline → 8 mo	0.90	0.42	0.09, 1.71	0.12	Small
Mental health	Internalizing Symptoms	8	Baseline → 8 mo	-0.11	0.21	-0.52, 0.29	-0.03	Negligible
Gender	Gender Roles Attitudes	8	Baseline → 8 mo	-0.60	0.36	-1.29, 0.10	-0.09	Negligible
Gender	Gender-Based Violence Attitudes	8	Baseline → 8 mo	0.27	0.20	-0.12, 0.65	0.07	Negligible
Empowerment	Freedom of Voice and Movement	8	Baseline → 8 mo	0.98	0.45	0.10, 1.87	0.12	Small

Notes. Effect sizes (Cohen's *d*) were calculated by dividing the adjusted difference-in-differences estimate by the pooled baseline standard deviation of the outcome. Negative values indicate improvements for outcomes where lower scores reflect better functioning (e.g., internalizing symptoms, gender roles, and gender-based violence). Estimates are from mixed-effects models adjusted for baseline covariates and clustering at the school level.

Supplementary Table S4. Standardized Effect Sizes (Cohen's d) for Intervention Effects at Two Years

Outcome Domain	Outcome Measure	Grade	Comparison Period	DiD Estimate (β)	SE	95% CI	Cohen's *d*	Interpretation
Primary Outcomes								
Wellbeing	Physical Health	7	Baseline → 20 mo	-0.49	0.24	-0.96, 0.02	-0.10	Negligible
Wellbeing	Psychological Wellbeing	7	Baseline → 20 mo	0.38	0.26	-0.13, 0.89	0.08	Negligible
Wellbeing	Social & Peer Support	7	Baseline → 20 mo	0.83	0.27	0.29, 1.37	0.16	Small
Secondary Outcomes								
Assets	Self-Efficacy	7	Baseline → 20 mo	0.78	0.33	0.14, 1.42	0.12	Small
Assets	Emotional Resilience	7	Baseline → 20 mo	0.68	0.40	-0.11, 1.47	0.09	Negligible
Mental health	Internalizing Symptoms	7	Baseline → 20 mo	-0.27	0.21	-0.68, 0.13	-0.07	Negligible
Gender	Gender Roles Attitudes	7	Baseline → 20 mo	-0.39	0.34	-1.06, 0.28	-0.06	Negligible
Gender	Gender-Based Violence Attitudes	7	Baseline → 20 mo	-0.19	0.19	-0.56, 0.18	-0.06	Negligible
Empowerment	Freedom of Voice and Movement	7	Baseline → 20 mo	0.86	0.43	0.01, 1.70	0.10	Small

Notes. Effect sizes (Cohen's *d*) were calculated by dividing the adjusted difference-in-differences estimate by the pooled baseline standard deviation of the outcome. Negative values indicate improvements for outcomes where lower scores reflect better functioning (e.g., internalizing symptoms, gender roles, and gender-based violence). Estimates are from mixed-effects models adjusted for baseline covariates and clustering at the school level.

Supplementary Table S5. Standardized Effect Sizes (Cohen's d) for Intervention Effects at One Year Post-Intervention

Outcome Domain	Outcome Measure	Grade	Comparison Period	DiD Estimate (β)	SE	95% CI	Cohen's *d*	Interpretation
Primary Outcomes								
Wellbeing	Physical Health	8	Baseline → 20 mo	0.22	0.24	-0.25, 0.68	0.05	Negligible
Wellbeing	Psychological Wellbeing	8	Baseline → 20 mo	0.66	0.26	0.14, 1.18	0.14	Small
Wellbeing	Social & Peer Support	8	Baseline → 20 mo	1.11	0.29	0.55, 1.67	0.21	Small
Secondary Outcomes								
Assets	Self-Efficacy	8	Baseline → 20 mo	0.76	0.32	0.13, 1.40	0.13	Small
Assets	Emotional Resilience	8	Baseline → 20 mo	1.49	0.40	0.70, 2.28	0.20	Small
Mental health	Internalizing Symptoms	8	Baseline → 20 mo	-0.55	0.21	-0.96, -0.14	-0.15	Small
Gender	Gender Roles Attitudes	8	Baseline → 20 mo	-0.66	0.35	-1.34, 0.03	-0.10	Negligible
Gender	Gender-Based Violence Attitudes	8	Baseline → 20 mo	0.19	0.20	-0.19, 0.58	0.05	Negligible
Empowerment	Freedom of Voice and Movement	8	Baseline → 20 mo	0.99	0.43	0.15, 1.82	0.12	Small

Notes. Effect sizes (Cohen's *d*) were calculated by dividing the adjusted difference-in-differences estimate by the pooled baseline standard deviation of the outcome. Negative values indicate improvements for outcomes where lower scores reflect better functioning (e.g., internalizing symptoms, gender roles, and gender-based violence). Estimates are from mixed-effects models adjusted for baseline covariates and clustering at the school level.