Youth First Impact Evaluation – 2018-21 (Bihar, India) Summary Report

November 2025

Peter L. Cooper, Independent Evaluation Consultant[1]

Project Background

WorldBeing's **Youth First** program is a school-based resilience and health intervention designed to strengthen emotional wellbeing and improve educational outcomes among adolescents in government middle schools. The program uses a training-of-trainers model and a structured curriculum that adapts global evidence and best practices from emotional resilience, positive psychology, social-emotional learning, and conflict resolution into activities and materials tailored for Indian youth.

Since 2010, WorldBeing has implemented Youth First and the girls-only version (Girls First) in Bihar, India. The program expanded to Jharkhand in 2024, with further scale-up planned in Assam and Meghalaya in 2026 and beyond.

Impact Summary

This evaluation demonstrates that WorldBeing's Youth First program generated meaningful and sustained improvements in adolescents' mental health, social wellbeing, self-efficacy, empowerment, gender attitudes, and substance-use behaviors in government middle schools in Bihar, with benefits evident after one and two years of exposure and sustained one year post-intervention. Drawing on data from 5,445 students across 70 schools, the study provides rigorous causal evidence from one of the world's largest longitudinal cluster-randomized controlled trials of a school-based wellbeing intervention in a low- and middle-income country. It is also among the few such evaluations to include a longer-term post-intervention follow-up, offering rare insight into the durability of program impacts over time.

Study Purpose

From 2018-2021, a mixed-methods, experimental evaluation was conducted to assess the impact of Youth First on adolescents in Bihar. The primary objective was to examine the intervention's effects on students' physical and mental wellbeing. Secondary objectives included exploring effects on assets and skills linked to positive mental health (self-efficacy, emotional resilience) as well as internalizing symptoms, gender attitudes, empowerment, and substance use (tobacco). The main research questions were:

1. What are the effects of a one-year and a two-year program?
2. Do program effects persist one year after a one-year program?

This research was designed to inform program refinement and provide evidence to guide replication and scale-up of Youth First in India and other low- and middle-income countries (LMICs).

Research Design

The main study within this evaluation was a cluster-randomized controlled trial (cRCT) with two arms: 40 schools implementing Youth First (treatment) and 30 schools not implementing Youth First (control). Across these 70 schools, a total of 5,445 students participated in the evaluation. This study represents one of the world's largest longitudinal cRCTs of a LMIC school-based wellbeing intervention to date.

In both arms, the same cohorts of Grade 7 and 8 students were surveyed at three timepoints[2]:
- **Baseline (T0):** prior to program implementation
- **First follow-up (T1):** 8 months post-baseline
- **Second follow-up (T2):** 20 months post-baseline

In the treatment group, Grade 7 students received two years of the intervention (as Grade 7 and Grade 8), and were assessed at baseline (T0), after their Grade 7 intervention year (T1) and after their Grade 8 intervention year (T2). Grade 8 students received one year (as Grade 8), and were assessed at baseline (T0), after their Grade 8 intervention year (T1), and at one-year follow-up post-intervention (T2).

Table 1 below lists the study's primary and secondary outcomes, associated measures, and how these constructs are referenced in the study findings.

Table 1. Study Primary and Secondary Outcomes and Their Associated Measures

Outcome (Type)	Measure(s)	Construct
Improved physical health (Primary)	KIDSCREEN Physical Wellbeing Scale (Ravens-Sieberer et al., 2014)	Physical health
Improved mental health (Primary)	KIDSCREEN Psychological Wellbeing Scale (Ravens-Sieberer et al., 2014)	Psychological wellbeing
	KIDSCREEN Social and Peer Support Scale (Ravens-Sieberer et al., 2014)	Social wellbeing
Improved assets and skills that contribute to positive mental health (Secondary)	Schwarzer's Generalized Self-Efficacy Scale (Schwarzer & Jerusalem, 1995)	Self efficacy
	Connor-Davidson Resilience Scale - 10 (Campbell-Sills & Stein, 2007)	Emotional resilience
	Strength and Difficulties Questionnaire Internalizing Factors Subscale (Goodman, 1997)	Presence of internalizing behaviors like anxiety, depression, and social withdrawal
Improved gender equality attitudes	International Center for Research on Women (ICRW) Gender Roles Subscale (Achyut et al., 2017)	Attitudes and beliefs regarding gender roles
	ICRW Gender Violence Subscale (Achyut et al., 2017)	Attitudes and beliefs regarding gender-based violence
Improved sense of empowerment	Global Early Adolescent Study Empowerment Scale (Zimmerman et	Sense of empowerment regarding freedom of movement and voice

---

[2] A third follow-up was originally planned for 32 months post-baseline but ultimately cancelled due to school closures resulting from the COVID-19 pandemic.

Outcome (Type)	Measure(s)	Construct
	al., 2019)	
Healthier substance use behaviors	Whether respondent smoked tobacco in previous 30 days	Tobacco smoking habits

A qualitative study component conducted at the end of the evaluation gathered focus group data from boys and girls in a subsample of treatment and control schools to contextualize the quantitative results. This qualitative study is described elsewhere (Leventhal et. al, 2022).

Data Analysis

Quantitative data gathered from student surveys were analyzed using a difference-in-differences approach, comparing changes between baseline and each follow-up for treatment versus control groups.

Analyses employed multiple imputation to address missing data and used linear mixed regression models with robust standard errors for primary and secondary outcomes, with the exception of tobacco use, which was modeled via mixed logistic regression due to its binary nature. Models controlled for background characteristics, including age, gender, marital status at baseline, and continuous school enrollment at baseline.

Key Findings[3]

After one year of the program, Youth First led to statistically significant, meaningful improvements across multiple areas of students' wellbeing, assets, attitudes, and behavior (see Figure 1).

Compared to students in control schools, Youth First participants showed stronger **social wellbeing**, greater **self-efficacy**, improved **mental health** (internalizing symptoms), more equitable **gender attitudes**, greater **empowerment**, and reduced **substance-use** behaviors (tobacco).

For example, students in Youth First experienced a 5 percent increase in social wellbeing, while social wellbeing declined by 2 percent among students in control schools. Youth First students' self-efficacy increased modestly, while it remained largely unchanged for control students. Internalizing mental health symptoms declined more among Youth First students than among control students. Youth First students also demonstrated greater gains in empowerment, while no change was observed among students in comparison schools.

In addition, Youth First participants showed more positive gender attitudes over time compared to students in the comparison group. Risk behaviors also improved, as the share of students who reported smoking tobacco in the past month declined more sharply among Youth First students than among students who did not participate in the program.

Of note, differences between Youth First and non-Youth First participants were not statistically significant after one year of the program for measures of physical wellbeing, psychological wellbeing, emotional resilience, or attitudes towards gender-based violence.

---

[3] Tables with the measures used in this study and the detailed results appear in the Annex.

Figure 1. One-Year Change in Pooled Mean Outcomes by Study Arm (Grade 7 and 8 Combined; Statistically Significant Outcomes Only)
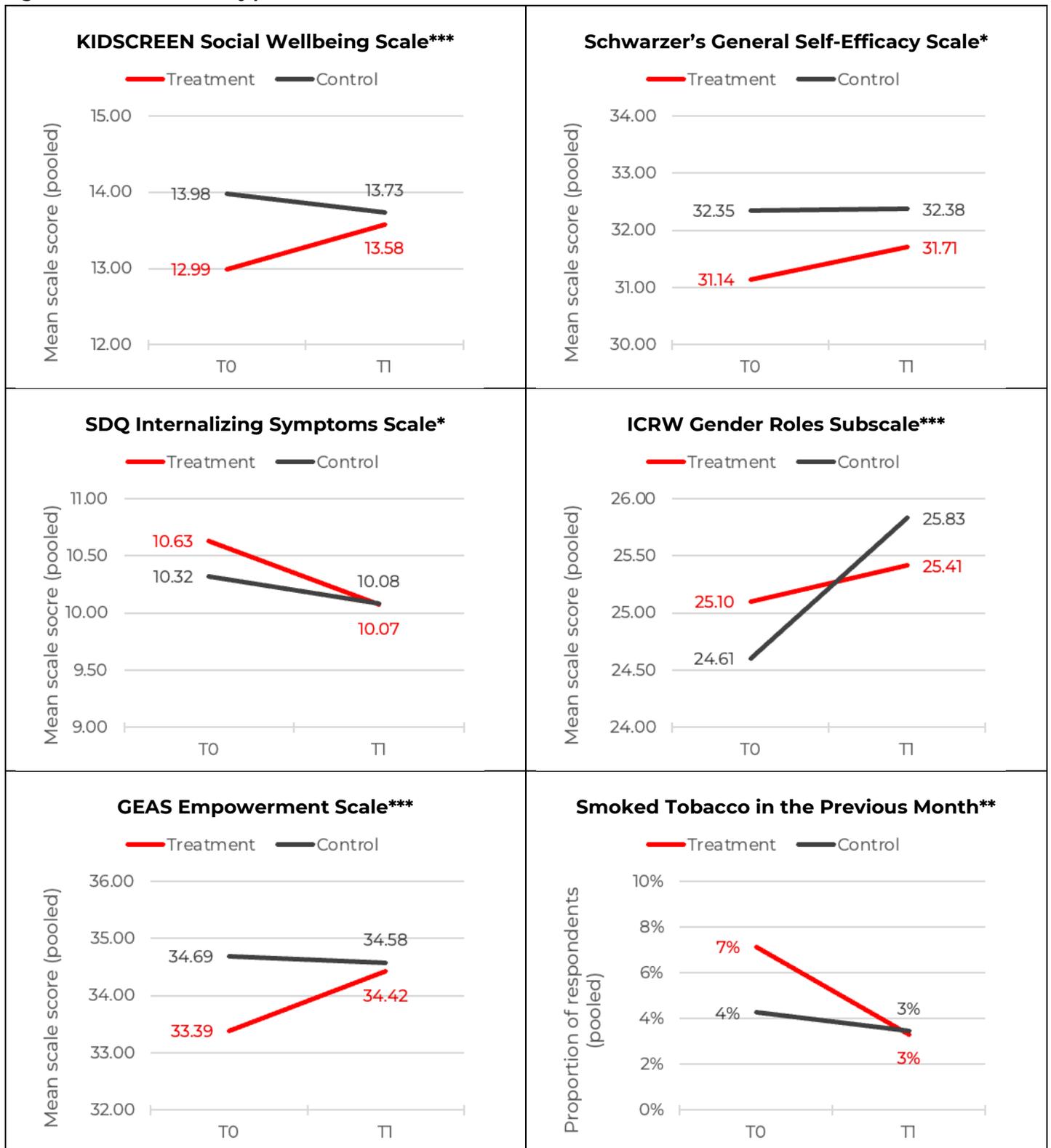
Note: One asterisk (*) indicates a statistically significant positive difference-in-difference between groups at *p* < .05. Two asterisks (**) indicates a statistically significant positive difference-in-difference between groups at *p* < .01. Three asterisks (***) indicates a statistically significant positive difference-in-difference between groups at *p* < .001.

After two years of the program, Youth First continued to produce meaningful and sustained benefits for students (see Figure 2). Students from the Grade 7 cohort who participated in Youth First for two years showed greater improvements in wellbeing, personal assets, and attitudes than students in comparison schools. In particular, Youth First students demonstrated stronger gains in **social wellbeing**, **self-efficacy**, and **empowerment**.

For example, Youth First students experienced a 57 percent increase in social wellbeing, compared to a 46 percent increase among students in control schools. Youth First students' self-efficacy remained stable, while it declined among students in control schools. In addition, Youth First students reported significantly greater increases in their empowerment than control students.

Figure 2. Two-Year Change in Pooled Mean Outcome Scores by Study Arm (Grade 7 Cohort; Statistically Significant Outcomes Only)

Note: One asterisk (*) indicates a statistically significant positive difference-in-difference between groups at *p* < .05. Three asterisks (***) indicates a statistically significant positive difference-in-difference between groups at *p* < .001.

Differences between students who participated in Youth First for two years and students from control schools were not statistically significant on measures of physical wellbeing, psychological wellbeing, emotional resilience, internalizing factors, gender attitudes, or smoking habits.

One year after completing the yearlong Youth First program, positive effects were sustained and, in several areas, strengthened (see Figures 3 and 4). Students from the Grade 8 cohort who were surveyed one year after participation continued to show greater improvements in wellbeing, personal assets, attitudes, and behaviors than students in comparison schools. In particular, one year after program completion, Youth First students demonstrated stronger **psychological and social wellbeing**, greater **self-efficacy** and **emotional resilience**, improved **mental health**, a stronger sense of **empowerment**, and healthier **substance-use** behaviors.

For example, Youth First students experienced a 39 percent increase in psychological wellbeing, compared to a 34 percent increase among students in control schools. Youth First students' emotional resilience remained stable, while it declined among control students. Internalizing mental health symptoms declined more sharply among Youth First students than among their control-school peers. Youth First students also reported greater gains in their sense of empowerment, while empowerment declined among students in comparison schools. Risk behaviors improved as well: the share of students who reported smoking tobacco in the past month declined among Youth First students, while it increased among students who did not participate in the program.

Figure 3. One-Year Post-Program Change in Pooled Mean Outcome Scores by Study Arm for Psychological and Social Wellbeing (Grade 8 Cohort)
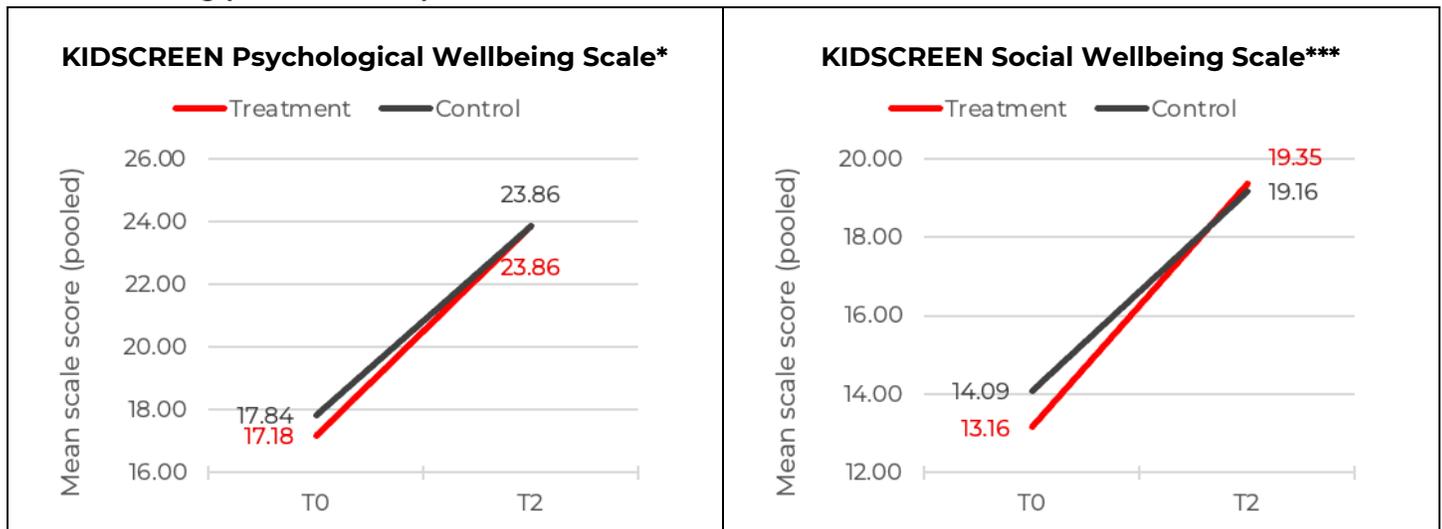

Note: One asterisk (*) indicates a statistically significant positive difference-in-difference between groups at *p* < .05. Three asterisks (***) indicates a statistically significant positive difference-in-difference between groups at *p* < .001.

Figure 4. One-Year Post-Program Changes in Pooled Mean Outcome Scores by Study Arm for Self-Efficacy, Resilience, Internalizing Factors, Empowerment, and Substance Use (Grade 8 Cohort)
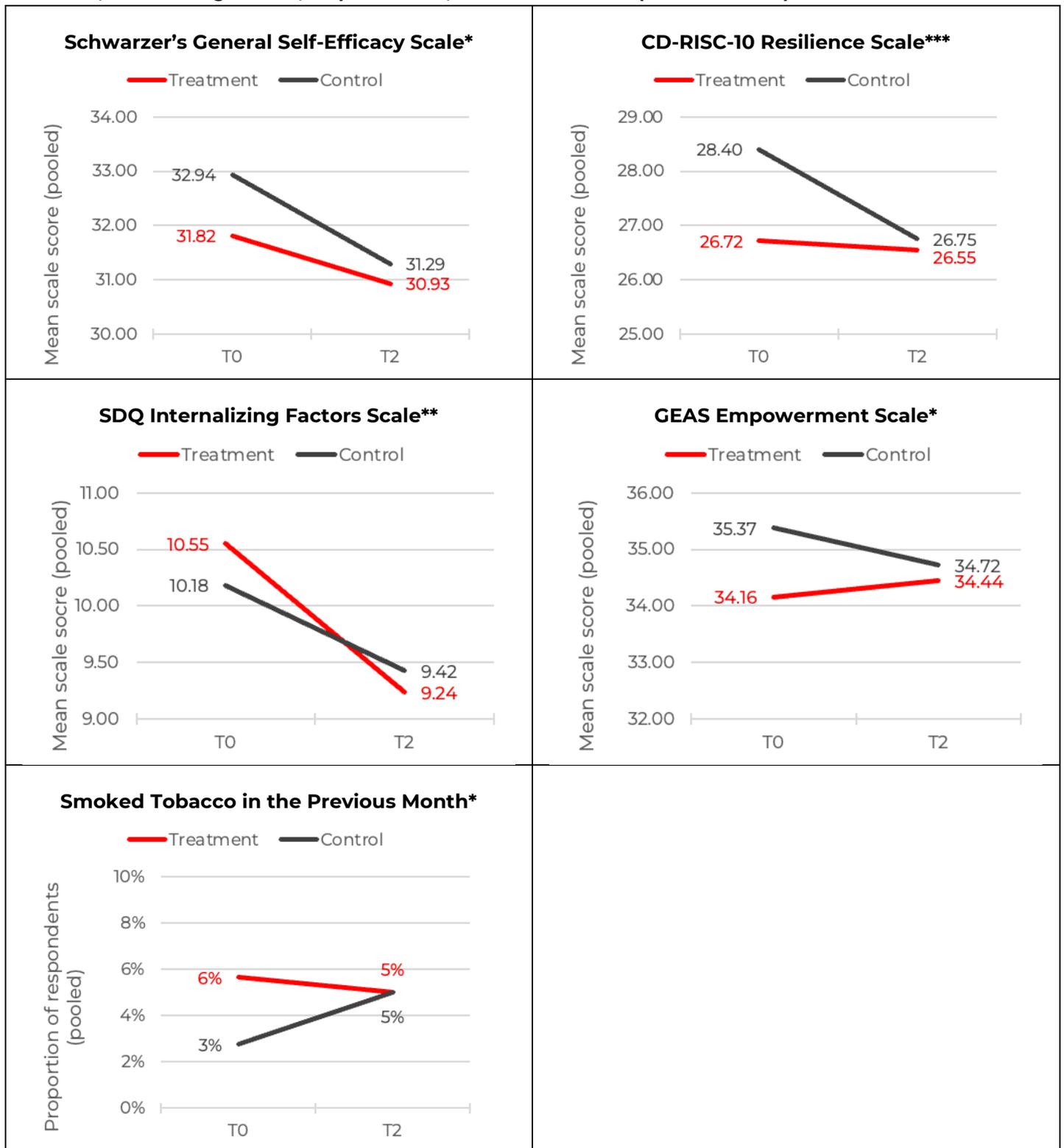
Schwarzer's General Self-Efficacy Scale*

Treatment — Control

Mean scale score (pooled)

34.00
33.00 — 32.94
32.00 — 31.82
31.00 — 31.29 / 30.93
30.00

T0 — T2

CD-RISC-10 Resilience Scale***

Treatment — Control

Mean scale score (pooled)

29.00
28.00 — 28.40
27.00 — 26.72 — 26.75 / 26.55
26.00
25.00

T0 — T2

SDQ Internalizing Factors Scale**

Treatment — Control

Mean scale socre (pooled)

11.00
10.50 — 10.55
10.18
10.00
9.50 — 9.42 / 9.24
9.00

T0 — T2

GEAS Empowerment Scale*

Treatment — Control

Mean scale score (pooled)

36.00
35.00 — 35.37
34.72 / 34.44
34.00 — 34.16
33.00
32.00

T0 — T2

Smoked Tobacco in the Previous Month*

Treatment — Control

Proportion of respondents (pooled)

10%
8%
6% — 6% — 5%
4% — 3% / 5%
2%
0%

T0 — T2

Note: One asterisk (*) indicates a statistically significant positive difference-in-difference between groups at *p* < .05. Two asterisks (**) indicates a statistically significant positive difference-in-difference between groups at *p* < .01. Three asterisks (***) indicates a statistically significant positive difference-in-difference between groups at *p* < .001.

Differences in two-year changes between students who participated in Youth First for one year and students from control schools were not statistically significant on measures of physical wellbeing or gender equality attitudes.

Discussion

This evaluation demonstrated that Youth First produced measurable, positive effects on students' mental health, self-efficacy, emotional resilience, internalizing symptoms, sense of empowerment, and substance use. Positive effects were realized after one year of the program, two years of the program, and were sustained and in some cases strengthened one year after program completion. The number and magnitude of effects at each timepoint (first and second follow-up) with each cohort (Grade 7 or Grade 8) suggest possible differences in how program exposure, age, or school context influence outcomes, warranting further investigation.

Conclusion and Next Steps

Findings from this evaluation provide robust evidence that Youth First positively impacts adolescents' mental health, empowerment, and key psychosocial assets. This evidence can support engagement with state governments to inform program scale-up in Bihar and Jharkhand and expansion to Assam and Meghalaya. Moreover, the sustained positive effects, especially for Grade 8 students one-year post-program, suggest that Youth First's benefits are likely to endure, even as the program is integrated into government systems.

Next steps include:
- **Disseminating findings** to state and national stakeholders to inform policy and program decisions.
- **Conducting further research** to explore differences in program effects by grade level and exposure duration.

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*. New Delhi, India: 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

Campbell-Sills, L., & Stein, M. B. (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. https://doi.org/10.1002/jts.20271

Goodman, R. (1997). *The Strengths and Difficulties Questionnaire: A research note*. Journal of Child Psychology and Psychiatry, 38(5), 581–586. https://doi.org/10.1111/j.1469-7610.1997.tb01545.x

Leventhal, K. S., Cooper, P. L., DeMaria, L. M., Priyam, P., Shanker, H., Andrew, G., & Leventhal, S. (2022). *Promoting wellbeing and empowerment via Youth First: Exploring psychosocial outcomes of a school-based resilience intervention in Bihar, India*. Frontiers in Psychiatry, 13, Article 931449. https://doi.org/10.3389/fpsyt.2022.931449

Pulerwitz, J., & Barker, G. (2008). *Measuring attitudes toward gender norms among young men in Brazil: Development and psychometric evaluation of the GEM Scale*. Men and Masculinities, 10(3), 322–338. https://doi.org/10.1177/1097184X06298778

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 Research, 23(3), 791-803. https://doi:10.1007/s11136-013-0428-3

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

Zimmerman, L. A., Li, M., 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

.

Annex

Table 2 Mean Pooled Scale Scores by Study Arm at First and Second Follow-Up – Grade 7 and 8 (combined)

Measure	Arm	Baseline (T0)		1st Follow-Up (T1)		Difference	Adjusted Difference-in-Difference	
		Mean	SE	Mean	SE	T0-T1	T0-T1	Sig.
KIDSCREEN Physical Health Scale	Treatment	18.06	0.08	18.35	0.08	0.29	-0.24	
(Min-max: 5-25; higher score = better result)	Control	18.48	0.10	19.01	0.09	0.54		
KIDSCREEN Psychological Wellbeing Scale	Treatment	16.69	0.09	17.13	0.09	0.44	0.09	
(Min-max: 0-24; higher score = better result)	Control	17.47	0.10	17.82	0.10	0.35		
KIDSCREEN Social Wellbeing Scale	Treatment	12.99	0.10	13.58	0.09	0.59	0.84	***
(Min-max: 0-24; higher score = better result)	Control	13.98	0.11	13.73	0.10	-0.25		
Schwarzer's General Self-Efficacy Scale	Treatment	31.14	0.11	31.71	0.11	0.56	0.54	*
(Min-max: 10-40; higher score = better result)	Control	32.35	0.13	32.38	0.12	0.02		
Connor-Davidson Resilience Scale-10	Treatment	26.09	0.14	26.25	0.14	0.16	0.18	
(Min-max: 0-40; higher score = better result)	Control	27.56	0.16	27.55	0.16	-0.01		
SDQ Internalizing Factors Scale	Treatment	10.63	0.07	10.07	0.07	-0.56	-0.32	*
(Min-max: 0-20; lower score = better result)	Control	10.32	0.08	10.08	0.08	-0.24		
ICRW Gender Roles Subscale	Treatment	25.10	0.12	25.41	0.12	0.31	-0.91	***
(Min-max: 10-40; lower score = better result)	Control	24.61	0.13	25.83	0.13	1.22		
ICRW Gender-Based Violence Subscale	Treatment	13.14	0.06	13.55	0.07	0.41	0.09	
(Min-max: 5-20; lower score = better result)	Control	13.39	0.08	13.70	0.07	0.32		
GEAS Empowerment Scale	Treatment	33.39	0.15	34.42	0.15	1.03	1.15	***
(Min-max: 0-54; higher score = better result)	Control	34.69	0.18	34.58	0.16	-0.11		

Legend of statistically significant differences: * $p < .05$, ** $p < .01$, *** $p < .001$

Table 3 Mean Pooled Scale Scores by Study Arm and Timepoint – Grade 7

Measure	Arm	Baseline (T0) Mean	SE	1st Follow-Up (T1) Mean	SE	2nd Follow-Up (T2) Mean	SE	Difference T0-T1	T0-T2	Adjusted Difference-in-Difference T0-T1		T0-T2	
KIDSCREEN Physical Health Scale	Treatment	17.81	0.12	18.32	0.11	18.61	0.10	0.51	0.80	-0.25		-0.49	
(Min-max: 5-25; higher score = better result)	Control	18.07	0.14	18.83	0.13	19.36	0.12	0.76	1.29				
KIDSCREEN Psychological Wellbeing Scale	Treatment	16.24	0.13	17.09	0.12	23.78	0.11	0.85	7.54	0.13		0.38	
(Min-max: 0-24; higher score = better result)	Control	17.11	0.14	17.83	0.14	24.27	0.12	0.72	7.16				
KIDSCREEN Social Wellbeing Scale	Treatment	12.83	0.14	13.68	0.13	20.00	0.12	0.85	7.18	0.75	*	0.83	***
(Min-max: 0-24; higher score = better result)	Control	13.87	0.15	13.98	0.15	20.22	0.15	0.10	6.34				
Schwarzer's General Self-Efficacy Scale (Min-max: 10-40; higher score = better result)	Treatment	30.52	0.16	31.32	0.16	30.31	0.15	0.80	-0.22	0.58		0.78	*
	Control	31.79	0.18	32.01	0.17	30.79	0.16	0.22	-1.00				
Connor-Davidson Resilience Scale-10	Treatment	25.51	0.20	25.77	0.21	26.02	0.18	0.27	0.51	-0.52		0.68	
(Min-max: 0-40; higher score = better result)	Control	26.75	0.23	27.54	0.23	26.58	0.20	0.79	-0.17				
SDQ Internalizing Factors Scale	Treatment	10.71	0.09	10.15	0.10	9.66	0.10	-0.55	-1.05	-0.58	*	-0.27	
(Min-max: 0-20; lower score = better result)	Control	10.45	0.11	10.43	0.11	9.68	0.11	-0.03	-0.78				
ICRW Gender Roles Subscale	Treatment	24.95	0.16	25.35	0.16	25.75	0.16	0.40	0.81	-1.22	***	-0.39	
(Min-max: 10-40; lower score = better result)	Control	24.47	0.19	26.09	0.18	25.67	0.18	1.62	1.20				
ICRW Gender-Based Violence Subscale	Treatment	13.13	0.08	13.42	0.10	13.62	0.09	0.28	0.49	-0.07		-0.19	
(Min-max: 5-20; lower score = better result)	Control	13.26	0.10	13.61	0.10	13.94	0.10	0.35	0.68				
GEAS Empowerment Scale	Treatment	32.68	0.21	33.99	0.21	33.96	0.19	1.31	1.28	1.30	**	0.86	*
(Min-max: 0-54; higher score = better result)	Control	34.03	0.26	34.05	0.23	34.46	0.20	0.02	0.42				

Legend of statistically significant differences: * $p < .05$, ** $p < .01$, *** $p < .001$

Table 4. Mean Pooled Scale Scores by Study Arm and Timepoint – Grade 8

	Arm	Baseline (T0) Mean	SE	1st Follow-Up (T1) Mean	SE	2nd Follow-Up (T2) Mean	SE	Difference T0-T1	T0-T2	Adjusted Difference-in-Difference T0-T1		T0-T2	
KIDSCREEN Physical Health Scale	Treatment	18.33	0.12	18.39	0.13	18.75	0.11	0.06	0.42	-0.24		0.22	
(Min-max: 5-25; higher score = better result)	Control	18.90	0.13	19.20	0.13	19.10	0.12	0.30	0.20				
KIDSCREEN Psychological Wellbeing Scale (Min-max: 0-24; higher score = better result)	Treatment	17.18	0.13	17.18	0.13	23.86	0.12	0.00	6.67	0.02		0.66	*
	Control	17.84	0.14	17.82	0.13	23.86	0.14	-0.02	6.01				
KIDSCREEN Social Wellbeing Scale (Min-max: 0-24; higher score = better result)	Treatment	13.16	0.14	13.47	0.13	19.35	0.14	0.30	6.18	0.92	**	1.11	***
	Control	14.09	0.15	13.47	0.15	19.16	0.15	-0.62	5.08				
Schwarzer's General Self-Efficacy Scale (Min-max: 10-40; higher score = better result)	Treatment	31.82	0.15	32.13	0.15	30.93	0.15	0.31	-0.88	0.50		0.76	*
	Control	32.94	0.18	32.75	0.16	31.29	0.17	-0.18	-1.65				
Connor-Davidson Resilience Scale-10 (Min-max: 0-40; higher score = better result)	Treatment	26.72	0.19	26.77	0.20	26.55	0.19	0.05	-0.17	0.90	*	1.49	***
	Control	28.40	0.22	27.55	0.21	26.75	0.20	-0.85	-1.65				
SDQ Internalizing Factors Scale	Treatment	10.55	0.10	9.98	0.10	9.24	0.10	-0.57	-1.31	-0.11		-0.55	**
(Min-max: 0-20; lower score = better result)	Control	10.18	0.11	9.73	0.11	9.42	0.12	-0.45	-0.76				
ICRW Gender Roles Subscale	Treatment	25.26	0.17	25.48	0.17	25.62	0.17	0.22	0.35	-0.60		-0.66	
(Min-max: 10-40; lower score = better result)	Control	24.75	0.19	25.56	0.19	25.76	0.17	0.81	1.01				
ICRW Gender-Based Violence Subscale	Treatment	13.15	0.09	13.70	0.09	13.64	0.10	0.55	0.49	0.27		0.19	
(Min-max: 5-20; lower score = better result)	Control	13.52	0.11	13.80	0.11	13.81	0.10	0.28	0.30				
GEAS Empowerment Scale (Min-max: 0-54; higher score = better result)	Treatment	34.16	0.22	34.89	0.22	34.44	0.19	0.73	0.28	0.98	*	0.99	*
	Control	35.37	0.24	35.12	0.23	34.72	0.21	-0.25	-0.65				

Legend of statistically significant differences: * $p < .05$, ** $p < .01$, *** $p < .001$

Table 5. Pooled Frequency of Students Smoking of Tobacco in Previous 30 Days by Study Arm and Timepoint

Grade	Arm	Baseline (T0)	1st Follow Up (T1)	2nd Follow Up (T2)	Difference T0-T1	T0-T2	Adjusted Difference-in-Difference T0-T1	T0-T2
Grade 7 and 8	Treatment	7%	3%	-	-4%	-	-3% **	-
(combined)	Control	4%	3%	-	-1%	-		
Grade 7	Treatment	8%	3%	5%	-5%	-3%	-3% *	-1%
	Control	6%	4%	4%	-2%	-2%		
Grade 8	Treatment	6%	3%	5%	-2%	0%	-2%	-3% *
	Control	3%	3%	5%	0%	2%		

Legend of statistically significant differences: * $p < .05$, ** $p < .01$, *** $p < .001$