

Supplementary variables

Supplementary Table S1. Participant and cluster demographic characteristics (extended)
(student-level supplementary variables)

Variable	Category	Intervention (n = 120)	Control (n = 120)	Statistical test	P-value
Socioeconomic indicators					
Parental education — father, n (%)				χ^2	
Primary or below		38 (31.7)	41 (34.2)		0.68
Secondary		52 (43.3)	50 (41.7)		
Tertiary or above		30 (25.0)	29 (24.2)		
Parental education — mother, n (%)				χ^2	
Primary or below		34 (28.3)	37 (30.8)		0.72
Secondary		56 (46.7)	54 (45.0)		
Tertiary or above		30 (25.0)	29 (24.2)		
Parental occupation — agriculture, n (%)				χ^2	0.69
Household size, mean (SD)		4.2 (1.4)	4.1 (1.3)	<i>t</i> -test	0.55
Technology access					
Personal smartphone ownership, n (%)		108 (90.0)	105 (87.5)	χ^2	0.53
Home internet access, n (%)		94 (78.3)	96 (80.0)	χ^2	0.74
Average daily internet use, hours, mean (SD)		6.8 (2.3)	6.6 (2.2)	<i>t</i> -test	0.48
Primary internet device: smartphone, n (%)		101 (84.2)	98 (81.7)	χ^2	0.60
Mental health history					
Previous mental health consultation, n (%)		18 (15.0)	16 (13.3)	χ^2	0.71
Family history of mental		22 (18.3)	24 (20.0)	χ^2	0.73

Variable	Category	Intervention (n = 120)	Control (n = 120)	Statistical test	P-value
illness, n (%)					
Prior suicide-related ideation (lifetime), n (%)		14 (11.7)	13 (10.8)	χ^2	0.84
Academic profile					
School grade (mathayom), n (%)				χ^2	0.61
M.1–M.3 (age 12–14)		54 (45.0)	57 (47.5)		
M.4–M.6 (age 15–18)		66 (55.0)	63 (52.5)		
School attendance rate previous semester, %, mean (SD)		87.4 (8.6)	88.1 (8.2)	<i>t</i> -test	0.52

All $P > 0.05$ confirm adequate baseline balance. Standardized mean differences for all continuous variables were < 0.20 , indicating negligible between-group differences consistent with successful cluster randomization. SD, standard deviation.

Cluster (school-level supplementary variables)

Variable	School 1 (I)	School 2 (I)	School 3 (I)	School 4 (I)	School 5 (C)	School 6 (C)	School 7 (C)	School 8 (C)
Total enrollment	342	298	412	287	371	312	265	395
Students enrolled in trial, n	30	30	30	30	30	30	30	30
School stratum (enrollment)	≥ 300	< 300	≥ 300	< 300	≥ 300	≥ 300	< 300	≥ 300
Distance to district hospital, km	8.2	14.7	6.4	18.3	11.2	9.8	21.4	7.6
Student-to-counselor ratio	342:1	298:1	412:1	287:1	371:1	312:1	265:1	395:1
Internet connectivity quality†	Moderate	Low	Moderate	Low	Moderate	Moderate	Low	Moderate
PHQ-9 cluster mean at baseline	11.6	12.1	11.4	12.2	11.3	11.8	11.9	11.4
IAT cluster mean at baseline	57.8	59.1	57.4	58.9	57.2	58.4	58.6	57.4

I, Intervention cluster; *C*, Control cluster; † Connectivity quality classified as Moderate (>10 Mbps) or low (≤10 Mbps) per Ministry of Digital Economy and Society district-level data.

Supplementary Table S2. Psychometric properties of all study instruments across measurement time points.										
Instrument	Domain	Items	Scoring Range	Baseline α	Post-intervention α	3-month α	6-month α	ICC (test–retest)	CVI	Validation Reference
PHQ-9 (Thai)	Depression severity	9	0–27	0.89	0.88	0.87	0.87	0.83 (95% CI: 0.77–0.88)	—	Lotrakul <i>et al.</i> , 2008
IAT (Thai)	Internet addiction	20	20–100	0.92	0.91	0.90	0.91	0.86 (95% CI: 0.81–0.90)	—	Young, 1998; Thai validation Saipanish <i>et al.</i>
Mental Health Literacy Scale	Knowledge, attitude, help-seeking	24	24–120	0.86	0.88	0.87	0.86	0.81 (95% CI: 0.74–0.87)	0.92	Developed for this study
Knowledge subscale		8	8–40	0.82	0.84	0.83	0.82	—	0.94	
Attitude subscale		8	8–40	0.79	0.81	0.80	0.79	—	0.91	
Help-seeking subscale		8	8–40	0.84	0.85	0.84	0.84	—	0.90	
Family APGAR (Thai)	Family functioning	5	0–10	0.83	0.84	0.84	0.83	0.79 (95% CI: 0.71–0.85)	—	Smilkstein, 1978; Thai validation

Instrument	Domain	Items	Scoring Range	Baseline α	Post-intervention α	3-month α	6-month α	ICC (test–retest)	CVI	Validation Reference
PedsQL 4.0 (Thai)	Health-related quality of life	23	0–100	0.88	0.89	0.89	0.88	0.85 (95% CI: 0.80–0.89)	—	Varni <i>et al.</i> , 2001
Physical functioning		8	0–100	0.86	0.87	0.86	0.86	—	—	
Emotional functioning		5	0–100	0.81	0.82	0.81	0.81	—	—	
Social functioning		5	0–100	0.83	0.84	0.83	0.83	—	—	
School functioning		5	0–100	0.80	0.81	0.80	0.80	—	—	

α , Cronbach's alpha coefficient; ICC, Intraclass Correlation Coefficient (test–retest reliability over 2-week interval); CVI, Content Validity Index. All α values ≥ 0.79 indicate good-to-excellent internal consistency throughout the study period. No clinically meaningful degradation of reliability was observed across measurement time points; supporting measurement invariance.

Confirmatory factor analysis for mental health literacy scale (Baseline sample, n = 240).

Model fit index	Obtained value	Acceptable threshold	Interpretation
CFI (comparative fit index)	0.94	≥ 0.90	Acceptable
RMSEA	0.06	≤ 0.08	Acceptable
SRMR	0.07	≤ 0.08	Acceptable
χ^2/df	2.34	≤ 3.00	Acceptable

Supplementary Table S3. Itemized direct intervention cost breakdown.

Costs reported in Thai Baht (THB) per participant based on 120 intervention participants across 4 school clusters. Reference year: 2024–2025. Perspective: Societal.

Cost category	Item	Unit cost (THB)	Units per participant	Total per participant (THB)	Proportion of total cost (%)
Human resources					
Registered nurses	Salary allocation (pro-rated FTE)	42,000/month	0.15 FTE $\times$ 3 months	1,890	19.4
School counselors	Salary allocation (pro-rated FTE)	36,000/month	0.20 FTE $\times$ 3 months	2,160	22.1
Clinical psychologist	Consultation fees (6 sessions)	800/session	6 sessions/group ($\div$ 7 students)	686	7.0
Peer educators	Honorarium	500/month	3 months	1,500 ($\div$ 30) = 50	0.5
Research/supervision staff	Salary allocation	45,000/month	0.10 FTE $\times$ 3 months	450	4.6
HR subtotal training				5,236	53.6

Cost category	Item	Unit cost (THB)	Units per participant	Total per participant (THB)	Proportion of total cost (%)
Pre-intervention training workshop	Venue, materials, meals (3 days)	45,000 total	÷ 120	375	3.8
Monthly booster sessions	Facilitator fees, materials (3 sessions)	8,000/session	÷ 120	200	2.0
Training subtotal				575	5.9
Materials and resources					
Psychoeducation workbooks	Printing and design	85/unit	1	85	0.9
Family therapy manuals	Printing	120/unit	1	120	1.2
Fidelity checklists and assessment forms	Printing	45/unit	1	45	0.5
Materials subtotal				250	2.6
Technology infrastructure					
Digital screening platform (licensing/maintenance)	SaaS subscription	24,000/school/year	× 4 schools ÷ 120	800	8.2
Tablets/devices (amortized over 3 years)	Device cost 8,000 × 10 units/school	8,000	× 4 schools ÷ 3 ÷ 120	889	9.1
Internet connectivity upgrade	Mobile data top-up	299/month/school	× 3 months × 4 ÷ 120	30	0.3
Technology subtotal				1,719	17.6
participant-related costs					

Cost category	Item	Unit cost (THB)	Units per participant	Total per participant (THB)	Proportion of total cost (%)
Transportation reimbursement (families)	200 THB per session × eligible families	200	2.4 sessions on average	480 (÷ 120 allocation)	4.9
Incentives (completion certificates, refreshments)	Per participant	250	1	250	2.6
Crisis referral coordination	Per case requiring referral	1,200	0.05 cases per participant	60	0.6
Participant subtotal				790	8.1
Overhead and miscellaneous					
Institutional overhead (15.0%)	Applied to HR and training	—	—	872	8.9
Miscellaneous administrative	Printing, communication	318	1	318	3.3
Overhead subtotal				1,190	12.2
Total direct cost per participant				9,760 THB	100.0

Costs reflect 2024–2025 Thai economic conditions. Indirect participant costs (time, transportation not reimbursed) and productivity losses during the intervention period were estimated at 2,180 THB per participant from a societal perspective, incorporated into the full economic model but excluded from the direct cost calculation above. Healthcare utilization savings were estimated from routine claims data at Wangchin Hospital at 4,370 THB per participant over the 6-month follow-up period. Sensitivity analyses varying direct costs by ± 20% did not change cost-effectiveness conclusions (ICER range: 10,620–15,040 THB per depression-free year, remaining well below the 34,400 THB WHO-CHOICE threshold).*

Supplementary Table S4. Prespecified subgroup analyses at 6-month follow-up
PHQ-9 primary outcome — Differential treatment effects.

Subgroup	Intervention mean (95% CI)	Control mean (95% CI)	Between-group difference (95% CI)	Cohen's d	Interaction <i>P</i> -value
Sex					
Female (n = 139)	6.9 (6.0, 7.8)	10.0 (9.1, 10.9)	−3.1 (−4.3, −1.9)	0.70	0.56
Male (n = 101)	7.4 (6.2, 8.6)	10.5 (9.3, 11.7)	−3.1 (−4.8, −1.4)	0.65	
Age group					
M.1–M.3 (12–14 years; n = 111)	7.3 (6.2, 8.4)	10.4 (9.3, 11.5)	−3.1 (−4.6, −1.6)	0.66	0.42
M.4–M.6 (15–18 years; n = 129)	6.9 (5.9, 7.9)	10.1 (9.1, 11.1)	−3.2 (−4.6, −1.8)	0.71	
Baseline depression					
Severity					
Mild (PHQ-9 5–9; n = 94)	4.2 (3.4, 5.0)	6.8 (6.0, 7.6)	−2.6 (−3.7, −1.5)	0.54	0.009*
Moderate-to-severe (PHQ-9 ≥10; n = 146)	9.1 (8.0, 10.2)	13.0 (11.9, 14.1)	−3.9 (−5.4, −2.4)	0.81	

IAT Primary outcome - Differential treatment effects.

Subgroup	Intervention mean (95% CI)	Control mean (95% CI)	Between-group difference (95% CI)	Cohen's d	Interaction <i>P</i> -value
Sex					
Female	43.1 (39.8, 46.4)	54.2 (50.9, 57.5)	−11.1 (−15.7, −6.5)	0.66	0.48
Male	44.3 (40.2, 48.4)	55.7 (51.6, 59.8)	−11.4 (−17.0, −5.8)	0.69	
Age Group					
12–14 years	44.8 (41.0, 48.6)	55.6 (51.8, 59.4)	−10.8 (−16.2, −5.4)	0.64	0.37
15–18 years	42.6 (39.2, 46.0)	54.1 (50.7, 57.5)	−11.5 (−16.3, −6.7)	0.69	
Baseline IAT					
Severity					
Mild-moderate (IAT 40–59; n = 142)	39.4 (36.6, 42.2)	50.3 (47.5, 53.1)	−10.9 (−14.9, −6.9)	0.65	0.19
Severe (IAT ≥60; n = 98)	49.8 (45.6, 54.0)	61.8 (57.6, 66.0)	−12.0 (−17.8, −6.2)	0.71	

Interaction terms tested within mixed-effects models adjusted for clustering and baseline values.

$P < 0.05$ indicates a statistically significant differential treatment effect. Cohen's d values represent effect sizes within each subgroup. The significant interaction by baseline depression severity for PHQ-9 outcomes ($P = 0.009$) indicates that students with moderate-to-severe depression derived larger absolute and relative benefit from the SHIELD intervention, with implications for targeted prioritization in resource-constrained settings. No significant moderation was observed for IAT outcomes across any subgroup, suggesting uniform benefit of the intervention for internet addiction regardless of baseline characteristics.

Supplementary Table S5. mediation analysis — mechanisms of intervention effect.

Bootstrapped mediation analysis (10,000 resamples) examining pathways through which SHIELD reduced PHQ-9 at 6-month follow-up. Analyses conducted using the PROCESS macro (Hayes, 2022) adapted for clustered data.

Proposed mediator	Total effect (c)	Direct effect (c')	Indirect effect (ab) [95% Boot CI]	Proportion mediated (%)	Significance
Mental health literacy improvement (T0→T1)	−3.1	−2.1	−1.0 [−1.68, −0.42]	32.3%	Significant
Family APGAR improvement (T0→T1)	−3.1	−2.2	−0.9 [−1.54, −0.31]	29.0%	Significant
PedsQL improvement (T0→T1)	−3.1	−2.6	−0.5 [−0.94, −0.12]	16.1%	Significant
Full parallel mediation model	−3.1	−1.4	−1.7 [−2.68, −0.84]	54.8%	Significant

Bootstrapped 95% confidence intervals that exclude zero indicate statistically significant indirect effects. The full parallel mediation model, including all three mediators simultaneously, accounted for approximately 54.8% of the total intervention effect on depression at 6-month follow-up, suggesting that improvements in mental health literacy, family functioning, and quality of life together explain the majority of the mechanism through which SHIELD achieves its antidepressant effects. The remaining unexplained direct effect ($c' = -1.4$) may reflect unmeasured mechanisms such as therapeutic alliance, peer support dynamics, or direct behavioral activation effects not captured by current mediators.

Supplementary Table S6. Intervention fidelity monitoring — Component-level detail.

SHIELD component	Planned sessions/activities	Sessions delivered, n (%)	Mean fidelity score (%) [SD]	Observed ICC (fidelity across clusters)	Assessment method
S — Screening & surveillance	12 weekly digital check-ins	11.4/12 (95.0%)	93.2 (4.8)	0.12	Platform logs + supervisor audit
H — Health literacy enhancement	8 group psychoeducation sessions	7.4/8 (92.5%)	88.6 (6.1)	0.09	Session recording review (20% sample)
I — Internet wellness program	6 structured CBT sessions	5.6/6 (93.3%)	87.1 (7.3)	0.11	Fidelity checklist (100.0% sessions)
E — Empowering Families	6 SFBT sessions + 3 parent workshops	8.2/9 (91.1%)	84.3 (8.2)	0.14	Observer rating + supervisor evaluation
L — Linkage to care	Continuous (stepped-care pathway)	100.0% pathway adherence	91.4 (5.6)	0.08	Referral documentation audit
D — Data-driven monitoring	Weekly dashboard review + supervision	11.6/12 (96.7%)	90.8 (4.9)	0.07	Dashboard access logs
Overall SHIELD Fidelity	—	91.7%	88.4 (6.2)	0.10	Composite index

Fidelity scores represent the percentage of required protocol elements implemented correctly per session/activity, rated on standardized checklists by independent supervisors. The overall fidelity of 88.4% exceeded the pre-specified threshold of $\geq 80.0\%$. The component with the lowest fidelity score, Empowering Families (84.3%), corresponded to the component with the highest initial implementation barrier — parental engagement — as detailed in the main text. ICC values for fidelity across clusters (range: 0.07–0.14) are modest, indicating that fidelity was relatively consistent across schools rather than driven by one particularly high- or low-performing site.*

Supplementary Table S7. Missing data analysis and multiple imputation diagnostics.

Time point	Intervention completers	Intervention missing (%)	Control completers	Control missing (%)	MCAR test (little's χ^2)	<i>P</i> -value	Imputation method
Post-intervention (T1)	112	6.7%	115	4.2%	$\chi^2 = 14.3$, df = 12	0.28	Multiple imputation (m = 20)
3-month follow-up (T2)	108	10.0%	112	6.7%	$\chi^2 = 16.8$, df = 12	0.16	Multiple imputation (m = 20)
6-month follow-up (T3)	105	12.5%	110	8.3%	$\chi^2 = 18.2$, df = 12	0.11	Multiple imputation (m = 20)

Comparison of complete case vs. multiple imputation estimates (6-Month PHQ-9):

Analysis approach	Between-group difference	95% CI	Cohen's d	<i>P</i> -value
Complete case analysis	-3.2	-4.1 to -2.3	0.70	< 0.001
Multiple imputation (primary)	-3.1	-4.0 to -2.2	0.68	< 0.001
Best-case scenario (all missing = improved)	-3.6	-4.6 to -2.6	0.79	< 0.001
Worst-case scenario (all missing = no change)	-2.4	-3.3 to -1.5	0.53	< 0.001

Little's MCAR test $P > 0.05$ across all time points support the assumption of Missing at Random (MAR), justifying multiple imputation as the primary analytic strategy. The consistency of estimates across complete case, multiple imputation, and sensitivity scenarios demonstrates robustness of primary findings to missing data assumptions. Even under the most conservative worst-case scenario (all missing assumed to show no improvement), between-group differences remained statistically significant and clinically meaningful ($d = 0.53$).*