Impact Evaluation of Building Resilience for Adolescent Development Programme

26th September 2025

Inception Report

Prepared for:

United Nations Children's Fund (UNICEF) Ethiopia

Prepared by:

Economic Policy Research Institute (EPRI)

unicef
for every child

Economic Policy Research Institute

Disclaimer:

This report is intended solely for the recipient's use and contains confidential and proprietary information. Unauthorised distribution or reproduction is strictly prohibited.

The views and opinions expressed in this report are those of the authors and do not necessarily reflect the official policy or position of the Economic Policy Research Institute or its affiliates.

Authors: Karim Stephan, Luca Lazzarini, José Giarola

Evaluation Manager: Dr. Mussarrat Youssuf (UNICEF)

Title of Evaluation	Impact Evaluation of Building Resilience for Adolescent Development Programme
Region	Eastern and Southern Africa Regional Office
Year of Report Completion	2025
Office	UNICEF Ethiopia
Country	Ethiopia
Period	2024-2027
Date of Submission	September 26, 2025

Table of Contents

Acronyms 6

1 Introduction 8

2 Evaluation Context 9

2.1 National Context 10

2.2 UNICEF Country Programme in Ethiopia 15

2.3 UNICEF Building Resilience for Adolescent Development Programme Background 18

2.3.1 Programme Rationale & Background 18

2.3.2 Programme Packages, Interventions, and Target Population 19

2.3.3 Geographic Coverage 21

2.3.4 Stakeholder/Implementation Modality 22

2.3.5 Programme Progress 22

3 Theory of Change and Integrated Results Framework 23

3.1 Theory of Change Narrative 25

3.1.1 Problem Statement 25

3.1.2 Pathways of Change 25

3.1.3 Long-Term Outcomes 26

3.1.4 Medium-Term Outcomes 27

3.1.5 Short-Term Outcomes 28

3.2 Outputs and Key Activities 29

3.2.1 Key Outputs 29

3.2.2 Key Activities 30

3.3 Assumptions and Risks 31

3.3.1 Key Assumptions 31

3.3.2 Key Risks 32

4 Evaluation Purpose 33

5 Evaluation Objectives and Questions 35

5.1 Evaluation Objectives 35

5.2 Evaluation Questions 35

6 Evaluation Scope 39

6.1 Thematic Scope 39

6.2 Geographic Scope 39

6.3 Chronological Scope 40
7 Evaluation Framework 41
7.1 Evaluation Matrix 41
8 Methodology 60
8.1 Data Collection Methods 60
8.1.1 Baseline, Midline and Endline Surveys 60
8.1.2 Qualitative Methods 62
8.1.3 Monitoring Data 65
8.2 Data Analysis 65
8.2.1 Impact analysis 65
8.2.2 Effectiveness Analysis 68
8.2.3 Implementation Fidelity Analysis 70
8.2.4 Framework analysis 70
8.2.5 Cost-effectiveness Analysis (CEA) 72
8.2.6 Integration of Gender, Equity, and Empowerment of Women (GEEW) and Stakeholder Perspectives 75
8.2.7 Triangulation of Findings 76
8.3 Quality assurance 76
8.4 Limitations and mitigation measures 77
9 Evaluation Workplan 79
9.1 Baseline Evaluation (Sept–Nov 2025) 79
9.1.1 Baseline Objective 79
9.1.2 Draft Baseline Overall Evaluation Questions 79
9.1.3 Enumerator Training and Piloting 80
9.1.4 Data Collection 80
9.1.5 Baseline Reporting Process 81
9.1.6 Midline Objective 81
9.1.7 Midline Focus Areas 82
9.1.8 Midline Data Collection Methods 83
9.1.9 Enumerator Refresher and Tool Refinement 84
9.1.10 Data Collection - Fieldwork 85
9.1.11 Midline Reporting Process 85
9.2 Endline Evaluation (June–Sept 2027) 86

9.2.1 Endline Objective 86
9.2.2 Endline Focus Areas 86
9.2.3 Enumerator Training and Preparation 87
9.2.4 Endline Reporting Process 88
10 Risk Mitigation 90
11 Ethical considerations 91
11.1 Ethical approach....91
11.2 Ethical approval....94
12 Consolidated Timeline of Evaluation Activities (2025–2027) 94

Annexes Provided as Separate Attachment

Acronyms

Adolescents with disabilities (AWDs)

Bill & Melinda Gates Foundation (BMGF)

Body Mass Index (BMI)

Child Protection (CP)

Children with Disabilities (CWDs)

Cognitive Behavioural Therapy (CBT)

Convention on the Elimination of All Forms of Discrimination Against Women (CEDAW)

Convention on the Rights of the Child (CRC)

Cost-effectiveness Analysis (CEA)

Cost-Effectiveness Ratio (CER)

Difference-in-Differences (DiD)

Disability-Adjusted Life Years (DALYs)

Economic Policy Research Institute (EPRI)

Ethical Research Involving Children (ERIC)

Female Genital Mutilation (FGM)

Focus Group Discussions (FGDs)

Gender-based violence (GBV)

Gender, Equity, and Empowerment of Women (GEEW)

Government of Ethiopia (GoE)

Human Papilloma Virus (HPV)

Incremental Cost-Effectiveness Ratio (ICER)

International Non-Governmental Organisation (INGO)

Key Informant Interviews (KIIs)

Knowledge, Attitudes and Practices (KAPs)

Low- and Middle-Income Countries (LMICs)

Menstrual Health and Hygiene (MHH)

Menstrual Hygiene Management (MHM)

Mental Health and Psychosocial Support (MHPSS)

Ministry of Education (MOE)

Ministry of Health (MOH)

Ministry of Women and Social Affairs (MOWSA)

Monitoring & Evaluation (M&E)

Most Significant Change (MSC)

Organisation for Economic Co-operation and Development – Development Assistance Committee (OECD-DAC)

Out-of-school (OOS)

Post-Traumatic Stress Disorder (PTSD)

Primary Healthcare (PHC)

Quality-Adjusted Life Years (QALYs)

Regional Bureau of Women and Social Affairs (BoWSA)

Reproductive, Maternal, Neonatal, Child, and Adolescent Health (RMNCAH)

School Health and Nutrition (SHN)

Sexual and Gender-Based Violence (SGBV)

Sexual and Reproductive Health (SRH)

Sexually Transmitted Infections (STIs)

Social and Behavioural Change (SBC)

Terms of Reference (TOR)

Theory of Change (ToC)

United Nations Children's Fund (UNICEF)

Water, Sanitation, and Hygiene (WASH)

Weekly Iron Folic Acid Supplementation (WIFAS)

Youth-Friendly Health Services (YFS)

1 Introduction

This document presents the Inception Report for the evaluation of the **Building Resilience for Adolescent Development Programme** (hereafter referred to as "the programme") in Ethiopia. It follows an Evaluability Assessment that examined the programme's readiness to undergo a credible, ethical, and meaningful evaluation. The programme, which combines Health, Nutrition, Education, Social and Behavioural Change (SBC), Child Protection, Water, Sanitation and Hygiene (WASH), and Mental Health and Psychosocial Support (MHPSS) interventions, is an initiative supported by the Gates Foundation (BMGF) and implemented in partnership with UNICEF Ethiopia, and the Government of Ethiopia (GoE).

The purpose of this Inception Report is to lay the foundation for the upcoming evaluation by confirming its scope, design, methods, tools, and timeline. It provides an overview of the programme context, defines the key evaluation questions, and summarizes findings from the evaluability assessment, including the availability of data, feasibility of proposed methodologies and limitations associated to them. The report presents a comprehensive evaluation plan that spans baseline, midline, and endline phases with detailing account of the sampling strategy, data collection methods and tools, and data analysis and quality assurance.

The evaluation team relied on the following methods to prepare this Inception Report:

- **Review of background documents**
    - To inform the design of the evaluation and ensure alignment with programme objectives and context, the evaluation team conducted a comprehensive review of relevant documentation. This included the programme's investment document, roll-out plan, results framework and tracker, implementation and programmatic guidelines, previous evaluations and assessments, as well as relevant national strategies and development plans.
- **Stakeholders consultations**
    - Stakeholder engagement during the inception phase was essential for validating the evaluation scope and design, clarifying expectations, and ensuring contextual relevance. The evaluation team held consultations with UNICEF Ethiopia sector leads at the Country Office, regional programme staff, and staff of the Gates Foundation.

The document review and stakeholder consultations supported the assessment and refinement of the Theory of Change (ToC), while also helping to identify key risks, assumptions, and results pathways. As part of this process, the Evaluation Team sought to understand the programme's design, scope, and intended outcomes, as well as the rationale behind delivering the integrated package of interventions. Stakeholder consultations, in particular, provided insight into the implementation status of programme components, operational challenges, and learning priorities. These activities contributed to the development of Evaluation Questions tailored to the needs of the intended users and uses of the evaluation. They also informed the selection of appropriate methodologies, along with an understanding of their limitations and trade-offs, which it is discussed in the next sub-section.

As part of the inception period, the evaluation team has conducted and developed an Evaluability Assessment of the programme, with the goal to evaluate the readiness of the programme to undergo a credible, ethical, and meaningful evaluation, which directly informed the Inception Report. The Evaluability Assessment, included in the Annexe of this document, confirmed that a rigorous and meaningful evaluation of the programme is feasible, **but some limitations require attention. These are expanded on in the methodology section of this report.**

Collectively, the activities conducted during the inception phase and the findings from the Evaluability Assessment provided a solid foundation for this evaluation and directly informed the content and structure of this Inception Report.

2 Evaluation Context

School-based programmes such as free meals, vaccinations and deworming can be highly impactful for school-age children and adolescents, especially those from poorer backgrounds. Such schemes also help keep children in school, ensuring that government investment in education is not wasted. Globally, around nine in ten countries currently implement School Health and Nutrition (SHN) programmes. Among low- and middle-income countries, around 70% of them implement school health programs for adolescents, covering topics such as nutrition, hygiene, physical activity, sexual and reproductive health and rights, and life skills. More than 100 countries have school-based vaccination programmes, and more than 450 million school-age children are dewormed every year in schools in low- and middle-income countries. SHN programs help reduce absenteeism caused by preventable or treatable illnesses and hunger. For instance, school-based deworming and micronutrient supplementation can lead to 2.5 years of additional schooling. Malaria prevention interventions have been shown to reduce absenteeism by 62%, and promoting handwashing in low-income countries can reduce absenteeism by 21% to 61%. Improving school WASH facilities can also reduce girls' absence during menstruation. Students who are frequently bullied are twice as likely to skip school.[1]

Despite progress, many children, especially in the poorest countries, are missing out. Although many countries have made efforts to integrate different schemes, for example delivering school meals along with vaccinations, relatively few have taken the next step and adopted truly comprehensive approaches in which a range of initiatives is delivered systematically. Programs are often fragmented and inadequate, lacking comprehensive scope and consistent implementation at scale. There are also significant data gaps regarding policy coverage, program quality, and impact, especially in low-income countries and for specific aspects like mental health and services in non-government schools.[2] Against this backdrop and supported by the BMGF, UNICEF Ethiopia has partnered with the GoE to implemented the **Building**

[1] UNESCO, UNICEF, WFP, 2023. Ready to learn and thrive: School health and nutrition around the world. Paris: UNESCO, UNICEF, WFP

[2] UNESCO, UNICEF, WFP, 2023. Ready to learn and thrive: School health and nutrition around the world. Paris: UNESCO, UNICEF, WFP

Resilience for Adolescent Development Programme, which combines Health, Nutrition, Education, SBC, Child Protection, WASH, and MHPSS interventions, targeting adolescents in Ethiopia.

2.1 National Context

Ethiopia is the second most populous country in Africa, with an estimated population of 126.5 million. Located in the Horn of Africa near the Middle East and its markets, Ethiopia shares borders with Eritrea, Somalia, Kenya, South Sudan, and Sudan. In 2024, Ethiopia continued to grapple with a polycrisis, including armed conflict, intercommunal violence, climate-related hazards, disease outbreaks, food insecurity, and socio-economic challenges.

Ethiopia is highly vulnerable to climate change, ranking fifteenth out of 163 countries where children are most at risk. This vulnerability is worsened by rapid population growth and dependence on rain-fed agriculture. The Somali and Afar regions have been especially affected, with prolonged droughts causing crop failures, loss of income for farming families, and severe losses of livestock due to water and pasture shortages. As a result, poverty has deepened, and many families are struggling to meet their basic needs.

Ethiopia has experienced a modest increase in overall inequality despite significant poverty reduction between 2011 and 2016. The Gini coefficient, a measure of income inequality, rose from 0.30 in 2011 to 0.33 in 2016, although it remains relatively low compared to other countries in the region. This increase was primarily driven by the growing disparity between urban and rural areas. Urban households saw their consumption grow rapidly at 6% annually, while rural households experienced less than 1% growth. Consequently, the share of total inequality attributable to differences between rural and urban areas doubled to 29% by 2016.[3]

Poverty in Ethiopia is increasingly concentrated in rural areas, accounting for 88% of the poor population in 2016. While regional disparities in overall consumption and poverty rates are generally limited, with Oromia and Amhara hosting the largest shares of the poor due to their large populations, disparities are more pronounced across agro-ecological zones. The drought-prone lowlands exhibit the highest poverty rates (32% in 2016), while pastoral areas, contrary to common perception, have relatively low monetary poverty rates (22% in 2016) but significantly lag in non-monetary welfare dimensions such as education and health. The extreme poor are particularly concentrated in Somali regions, whereas transitory poverty is most prevalent in Amhara.[4] Climate change is projected to exacerbate these regional inequalities, particularly in the lowlands, by impacting agriculture and livestock productivity, which could heighten social tensions and increase the risk of conflict over resources.[5]

In the Amhara region, ongoing conflict between federal forces and the Fano militia has led to displacement, destruction of infrastructure, gender-based violence (GBV), limited access to

[3] Assessment, Ethiopia Poverty. "Harnessing Continued Growth for Accelerated Poverty Reduction." *World Bank* (2020): 36-61.
[4] Assessment, Ethiopia Poverty. "Harnessing Continued Growth for Accelerated Poverty Reduction." *World Bank* (2020): 36-61.
[5] World Bank. *Ethiopia - Country Climate and Development Report (English).* Washington, D.C. : World Bank Group (2024).

essential services, and disrupted livelihoods. In the Tigray region, despite some improvements following the peace agreement, around 1 million people remain internally displaced and face significant food insecurity. The need for MHPSS is particularly high among GBV survivors, especially women and girls. In the Oromia region, armed and intercommunal conflicts have damaged health facilities, severely disrupted basic services, and displaced nearly 1 million people.

Inequality of opportunity in terms of access to basic services and opportunities for children, such as education and improved water sources, improved between 2011 and 2016, but significant disparities persist, raising concerns under the Convention on the Rights of the Child (CRC), which guarantees children's rights to survival, development, and non-discrimination. These inequalities are primarily driven by household location (rural vs. urban) and household wealth, which largely determine access to key opportunities.[6] An estimated 21.4 million people across the country, including 12 million children, required urgent humanitarian assistance in health, nutrition, WASH, education, child protection, and social protection.[7]

Adolescence is a critical life stage marked by rapid physical, emotional, and cognitive changes that shape an individual's life trajectory. It is during this period, typically between the ages of 10 and 19, that foundational behaviours, identities, and aspirations are formed. Adolescence is a key time for educational investment, skills development, health promotion, and social empowerment. Schools serve not only as learning environments but also as protective spaces that foster life skills, build confidence, and offer access to essential services and information. Evidence shows that investments in adolescent education, especially for girls, lead to enhanced economic opportunities, improved health and nutrition outcomes, delayed marriage and childbirth, and increased overall well-being.[8] As such, holistic, cross-sectoral investments in adolescents are among the most effective strategies for advancing equitable development and breaking cycles of poverty.

In Ethiopia, where adolescents make up nearly 25% of the total population, with 68.75% between the ages of 10 and 14,[9] strengthening investments in this group is not only critical for individual well-being but also for achieving national development goals. Without targeted investment in adolescents' health, nutrition, and education, Ethiopia is unlikely to meet the objectives outlined in the National Food and Nutrition Strategy (2022–2030) or the 2030 Sustainable Development Goals.[10] Recent data illustrate the scale of the challenge. According to the national food and nutrition baseline survey report, 50% of adolescent girls live in a food-insecure household. On average, Ethiopian adolescent girls aged 15-19 years consume three food groups per day, with just 1 in 10 meeting the recommended intake of five or more food groups. Moreover, only 20% adolescent girls aged 10-19 years reported having consumed

[6] Assessment, Ethiopia Poverty. "Harnessing Continued Growth for Accelerated Poverty Reduction." *World Bank* (2020): 36-61.
[7] UNICEF Ethiopia Annual Report 2024
[8] Wodon, Q., Montenegro, C. E., Nguyen, H., & Onagoruwa, A. (2018). Missed opportunities: The high cost of not educating girls.
[9] https://www.unfpa.org/data/world-population/ET
[10] https://sdgs.un.org/goals

green leafy vegetables the previous day. When looking at the SDGs goals, under SDG 2, Zero Hunger, Ethiopia saw a reduction in acute malnutrition among children under five from 10% in 2015/16 to 7% in 2019/20, and both crop production and smallholder productivity increased annually by 5.3% and 4% respectively during 2015/16-2020/21. However, chronic malnutrition remains high at 37%, and progress in food security has been slow, compounded by various shocks like droughts, COVID-19, and rising food prices.[11]

These limited diets have serious implications for adolescent girls' nutritional status and overall health. Almost half of adolescent girls lack essential micronutrients, with urban areas experiencing higher rates (64%) than rural areas (45%). Folate deficiency is the most common micronutrient deficiency in Ethiopian adolescent girls, with urban areas (52%) having roughly double the deficiency rate compared to rural areas (29%). In general, 1 in 3 adolescent girls in Ethiopia is folate-deficient, putting them at risk of giving birth to children with neural tube defects. Folate deficiency is particularly prevalent in five regions: Dire Dawa (88%), Afar (83%), Somali (75%), Harari (67%), and Addis Ababa (56%). Anaemia remains a public health problem in the Somali (21%) and Benishangul-Gumuz (20%) regions.[12] In addition, the prevalence of overweight and obesity is rising among adolescents, reaching 6%, with higher rates observed in urban settings—reflecting an increasing shift toward sedentary lifestyles and poor dietary habits.

Poor health, particularly issues related to nutrition and hunger, significantly undermines education outcomes and the development of human capital across Low- and Middle-Income Countries (LMICs). It is estimated that poor health contributes to an estimated loss of 200 million to 500 million school days annually, while 66 million school-age children in low-income countries go to school hungry, directly impairing their ability to learn and attend school regularly. Iron-deficiency anaemia, which affects roughly 300 million schoolchildren globally, is associated with diminishing cognitive development essential for learning and academic achievement. Undernutrition during early childhood, particularly stunting, which affects one in four children under five, has long-term consequences, including reduced school participation and lower educational attainment, and may reduce adult income by as much as 22%. These effects are compounded by the fact that brain development continues through the middle-school and adolescent years, during which nutrition, disease, and stimulation play a critical role in shaping cognitive capacity. In this context, investing in adolescent health and nutrition is not only essential for physical development but also for enabling educational participation, learning outcomes, and lifelong productivity.[13]

In Ethiopia, these health and nutritional challenges intersect with structural challenges to education. The combined impact of COVID-19, conflict, and climate change has further contributed to a severe education crisis in Ethiopia. Overall, an estimated 17 million children across pre-primary, primary, and junior secondary levels are out of school, with nearly half

[11] Ethiopia Voluntary National Review 2022
[12] Food and Nutrition baseline study
[13] ICFGEO. The learning generation: Investing in education for a changing world. New York City: The International Commission on Financing Global Education Opportunity; 2016.

being girls,[14] facing long-term consequences for their development and future opportunities. Among adolescents aged 10 to 18, about 9.3 million were out of school in 2020, out of an estimated 21 million in that age group, with 53% of them being girls.[15] In addition, around 3.5 million children at the primary and secondary levels are at risk of dropping out, further deepening the crisis. Rural children are disproportionately affected, with significantly lower rates of enrolment, retention, and learning outcomes compared to their urban peers.[16]

According to the Ethiopia Voluntary National Review 2022, Ethiopia achieved improved gross enrolment rates of 95% in primary and 42.1% in secondary education by 2020/21, and net primary enrolment reached 86.4%. Nevertheless, quality remains a concern, with increased primary school dropout rates and a substantial rise in out-of-school children due to COVID-19 and conflicts, which also destroyed thousands of schools. Significant disparities persist for persons with disabilities, who face low enrolment and limited access to accessible facilities, and access to basic amenities like electricity remains low in primary schools, especially in rural areas.[17] These gaps reflect challenges to ensure the inclusive, equitable education commitments outlined in the CRC. Intergenerational educational mobility remains low, indicating that a child's educational outcomes are strongly influenced by their parents' education level, particularly for children from extremely poor households in rural areas.[18]

Access to safe WASH in schools' further influences education outcomes, particularly for adolescent girls. While studies have increasingly explored the role of WASH interventions in reducing absenteeism and dropouts, the situation in Ethiopia points to persistent gaps. According to 2023/24 education statistics annual abstract, 31.5% of primary and middle schools have access to functional water supply, with low access reported from Somali (7%), Afar (10%) and Gambella (15%) regions. Almost 91% of primary and middle schools have functional toilets, however, out of them, only 37.8% are improved. The regions that reported the lowest improved toilets are Somali (1.23%), Afar (7.83%) and Sidama (15.51%). From the schools reporting to have functional toilets, 70% of them have separate boys and girls' toilet and only 26% had available handwash.[19] Where separate facilities are lacking, female students avoid using them out of fear of harassment.[20] In emergency-affected areas, schooling becomes extremely difficult, especially during droughts, forcing students to miss classes or drop out entirely. The absence of comprehensive water and sanitation services hampers attendance facilitates disease transmission such as intestinal worms and denies privacy and dignity to girls and female teachers.[21] Low levels of awareness, inadequate infrastructure, lack of enforcement against misuse, and poor maintenance further hinder the effective use of existing WASH facilities.

[14] Available here
[15] The Luminos fund out of school children analysis 2021
[16] Available here
[17] Ethiopia Voluntary National Review 2022
[18] Assessment, Ethiopia Poverty. "Harnessing Continued Growth for Accelerated Poverty Reduction." *World Bank* (2020): 36-61.
[19] Education Statistics Annual Abstract 2023/2024, Ministry of Education, Addis Ababa, Ethiopia.
[20] One Wash National Program (OWNP) A Multi-Sectoral Swap
[21] One Wash National Program (OWNP) A Multi-Sectoral Swap

In terms of improvements against SDG 6: Clean Water and Sanitation, the proportion of households with access to an improved source of drinking water increased from 64.8% in 2015/16 to 68.8% in 2019/20, and open defecation declined from 32.3% to 27.1% over the same period. However, access to safely managed drinking water and sanitation services remains far from universal, with only 19.5% of households having improved sanitation in 2019/20, and recurrent droughts, conflicts, and floods continue to hinder effective water resource management and service delivery.[22]

Notwithstanding Ethiopia's education challenges already discussed, additional factors such as gender-based barriers, further compound the situation and disproportionately affect adolescent girls and limit their access to and continuation in education. Traditional gender norms, the heavy burden of domestic work, and long distances to school all affect girls disproportionately.[23] These structural and social challenges are further intensified by harmful practices such as child marriage and teenage pregnancy, which remain widespread. Currently, 40% of girls in Ethiopia are married before the age of 18, and 13% experience pregnancy during adolescence. Ethiopia is home to 17.3 million women and girls who were married before turning 18, including 7.5 million who were married before the age of 15. Child marriage and teenage pregnancy are the likely causes of four-fifths of births of children to mothers younger than 18. All forms of physical, verbal, and psychological harassment are common in Ethiopia. Moreover, there are weak sexual and gender-based violence (SGBV) reporting and response mechanisms, and a lack of support from the surrounding community. In terms of SDG 5: Gender Equality, Ethiopia improved its Global Gender Gap Index rank from 124th in 2015 to 97th in 2021. There has been progress in reducing child marriage and Female Genital Mutilation/Cutting (FGM/C), but accelerated efforts are needed to eliminate these by 2030 and aligning with the Convention on the Elimination of All Forms of Discrimination Against Women (CEDAW).[24]

Regarding gender inequality, poor households are notably less likely to be headed by a woman, a finding that contrasts with conventional wisdom. Female-headed households generally exhibit higher consumption levels and a lower likelihood of falling into poverty. However, women still face significant disadvantages across many other dimensions; for example, female farmers have less access to education, land, and finance than their male counterparts.[25] The official female labour force participation rate has remained stagnant at 46.3% over the last decade, and female unemployment in urban areas is high, affecting one in four women.[26] Studies suggest that matching women's employment to that of men could increase Ethiopia's long-run GDP per capita by 19.5%.[27]

[22] Ethiopia Voluntary National Review 2022
[23] Available here
[24] Ethiopia Voluntary National Review 2022
[25] Assessment, Ethiopia Poverty. "Harnessing Continued Growth for Accelerated Poverty Reduction." *World Bank* (2020): 36-61.
[26] Assessment, Ethiopia Poverty. "Harnessing Continued Growth for Accelerated Poverty Reduction." *World Bank* (2020): 36-61. And Bertelsmann Stiftung, BTI 2024 Country Report — Ethiopia. Gütersloh: Bertelsmann Stiftung, 2024.
[27] Ethiopia Gender Landscape, World Bank (2025)

While Ethiopia has been making significant progress in promoting school-based health and nutrition interventions through various initiatives such as health promotion campaigns, Human Papilloma Virus (HPV) vaccination, expanded school feeding programs, iron-folic acid supplements, hygiene education, and community-based health and WASH interventions, these efforts have mostly focused on children under five and school-age children, while adolescents have been largely overlooked.

As an integral part of expanding health services to adolescents, Ethiopia introduced HPV vaccination in December 2018 for a single-age cohort of 14-year-old girls following a two-year demonstration project in two districts of Oromia and Tigray regions from 2015 through 2017. Later, the country implemented a two-dose schedule with an interval of 6 to 12 months. The main platform to deliver HPV vaccination for girls in Ethiopia is a one-week campaign mode, a school-based outreach strategy. In 2023, more than half (56%) of young women aged 15-16 reported receiving at least one dose of the HPV vaccine, with a higher rate in young women residing in urban areas (64%) than those in rural areas (53%). 95% of the vaccinations took place at schools.[28]

Despite advances in expanding the HPV vaccine and other adolescent health programs, uptake remains limited due to gender-related challenges, including the limited decision-making power of girls to be vaccinated (parents are decision-makers), limited autonomy to move freely on their own and the fear of violence (sexual and physical), which may hamper their ability to reach facilities to get vaccinated. These barriers are worsened by limited access to accurate information and support. Myths about infertility, vaccine safety, and side effects continue to spread, further discouraging vaccination. These issues highlight broader gaps in adolescents' access to essential services and support systems, including sexual and reproductive health, nutrition, psychosocial support, and life skills education, especially for girls.

Social and Behaviour Change Communication (SBCC) strategies offer a way to address these gaps when tailored to adolescents' specific needs and vulnerabilities. Schools serve as a vital platform for reaching adolescents with age-appropriate, culturally relevant information on topics like nutrition, health, WASH, Sexual and Reproductive Health (SRH), HPV vaccination, and gender equality. When integrated into school-based programs, SBCC can help change harmful social norms, develop life skills, and empower young people, especially girls, to make informed decisions about their health and future.

2.2 UNICEF Country Programme in Ethiopia[29]

UNICEF's goal is to support the Government of Ethiopia in its commitment to protecting the rights of children and ensuring children's basic needs are met so that they can reach their full potential. UNICEF does this by working closely with the Government, international and national partners, and engaging with children. Ethiopia is Africa's second most populous

[28] PMA Ethiopia: Results from the 2023 cross section

[29] This section describes the main points outlined in the UNICEF Ethiopia Country Programme. For a more detailed discussion, please refer to the UNICEF Ethiopia Country Programme Draft available here and the UNICEF Country Kit available here.

country, with an estimated 115 million people in 2020, and 49% of its population under the age of 18. The country faces persistent issues such as high rates of multidimensional child poverty, with 88% of children (approximately 36 million) lacking access to at least three essential social services,[30] and stark disparities between rural (94%) and urban (42%) areas. Humanitarian needs are also substantial, including the presence of around 800,000 refugees and millions of internally displaced persons due to conflict, cyclical droughts, floods, and locust infestations.

Through its programme, UNICEF aims to ensure that all girls and boys, particularly those with disabilities, those experiencing multiple deprivations, and those affected by humanitarian crises, gain greater and more equitable access to quality services. The country programme is grounded in a comprehensive analysis[31] and places strong emphasis on integrated, multisectoral programming with a wide range of partners. It promotes enhanced collaboration across humanitarian, development, and peacebuilding efforts, with increased focus on gender-transformative programming and inclusion.

Guided by the principles of the Convention on the Rights of the Child (CRC), the programme prioritises equity, gender equality, transformation, inclusion, and sustainability. It focuses on six main areas: (i) health, (ii) nutrition, (iii) education, (iv) child protection, (v) climate, energy and environment, WASH, and (vi) social policy. UNICEF also aims to enhance programme effectiveness by integrating cross-cutting priorities and strategies to drive transformative change.

Health: This component focuses on **strengthening integrated multisectoral Primary Healthcare (PHC)** to deliver comprehensive reproductive, maternal, neonatal, child, and adolescent health (RMNCAH) services. **Key efforts include** expanding community-based health insurance, promoting local solutions, and integrating humanitarian-development linkages for resilience. UNICEF will support PHC financing, advocate for gender-responsive, youth-inclusive, and disability-inclusive RMNCAH policies, strengthen data systems, improve health infrastructure, and promote local production of essential medicines. It also aims to enhance immunisation services in underserved areas, strengthen mobile health teams, and improve emergency preparedness and response through SBC strategies.

Nutrition: The focus on this area is on empowering communities and transforming food systems to optimise nutrition outcomes. This involves enhancing local and household production of nutrient-dense foods, scaling up nutrition and shock-responsive social protection, and implementing integrated approaches like school nutrition packages. UNICEF will work to improve diets, practices, and services for children (including those with disabilities) by strengthening breastfeeding counselling, promoting nutritious complementary foods, and advocating for public financing. Efforts also include improving diets for school-aged children and adolescents, promoting fortified foods, and empowering women to access

[30] For instance, health, housing, education and water and sanitation

[31] The development of the country programme was based on a Country Programme Evaluation, a Gender Programmatic Review, a Partners' Perception Study, a Situation Analysis and a United Nations Common Country Analysis.

improved diets to prevent maternal and newborn malnutrition. Additionally, it supports early detection and treatment of wasting and acute malnutrition in early childhood, strengthening health systems and supply chains.

Education: This area prioritizes quality learning and skills development for vulnerable and hard-to-reach children and adolescents, aiming to equip them for successful transitions within education and from learning to earning. It emphasizes early childhood education, foundational skills, and digital technologies. UNICEF will support improved education planning, management, and monitoring, advocate for increased education spending, and integrate climate risk mitigation into policies. The programme also seeks to expand access to inclusive learning for out-of-school children, including those with disabilities and in humanitarian settings, by providing flexible learning opportunities and psychosocial support. Initiatives include strengthening foundational literacy and numeracy, promoting positive parenting, enhancing teacher training, and helping adolescents develop various skills for the future.

Child Protection (CP): This component integrates child protection, gender-based violence prevention, disability discrimination prevention, risk mitigation, and response, along with skill building and economic resilience for at-risk adolescents, youth, and women. Key actions involve strengthening institutional and financial capacity for child protection programmes, improving national coordination, and advocating for increased public financing and legal reforms, such as raising the age of criminal responsibility. UNICEF will also promote positive social and gender norms to prevent and respond to violence, encourage birth registration, and improve access to equitable, quality, and child-friendly justice services.

Climate, energy and environment, and WASH: The focus is on improving equitable and inclusive WASH services and practices while building climate resilience across UNICEF-supported sectors. This includes strengthening the implementation of the One WASH National Programme and the Total Sanitation to End Open Defecation and Urination strategy. UNICEF will work to build climate resilience by enhancing national and local capacities for risk-informed, inclusive, and climate-resilient WASH services in various settings, promoting innovative technologies for groundwater mapping, and supporting local green energy solutions. The programme also aims to improve integrated WASH and menstrual hygiene services in rural, pastoralist, hard-to-reach, and crisis-affected communities, and enhance services in urban and refugee communities.

Social Policy: This area aims to enhance budgetary and policy environments to ensure evidence-based, gender-transformative, disability- and financially inclusive, and shock-responsive social protection and basic services for vulnerable children. UNICEF will enhance the collection, analysis, and use of high-quality data to inform policies and strategies, institutionalize child poverty monitoring, and generate evidence on hard-to-reach populations. It supports the design and implementation of an inclusive and adaptive social protection system that responds to humanitarian contexts, including through cash transfer programmes, and advocates for increased public investment in child-inclusive services.

UNICEF aims to enhance the overall effectiveness of all programme components by integrating key priorities and strategies to drive transformative changes. This includes i) rights- and results-based planning, monitoring, and reporting; ii) gender-transformative programming; iii) disability inclusion and empowerment; iv) adolescent programming; v) resilience programming to adapt services to climate change; vi) SBC initiatives; vii) strengthening communication, advocacy, and partnerships; and, viii) conducting and utilizing quality evaluations to improve programmes and policies.

2.3 UNICEF Building Resilience for Adolescent Development Programme Background

2.3.1 Programme Rationale & Background

As outlined in the previous sections, adolescents in Ethiopia face a complex set of challenges that hinder their health, development, and future opportunities. These include widespread anaemia and malnutrition, limited dietary diversity, micronutrient deficiencies such as folate, insufficient access to sexual and reproductive health information and services, high rates of child marriage and adolescent pregnancy, and barriers to accessing quality education and WASH services. In response to this multidimensional vulnerability, UNICEF implemented the Joint Adolescent Nutrition-WASH-Education Programme (2018–2022), a comprehensive package of interventions that includes various activities aimed at influencing and improving knowledge, perceptions, and outcomes in nutrition, menstrual health, and hygiene. The programme's endline evaluation concluded that it contributed to improvements in nutrition and dietary diversity, despite challenging contexts during implementation. Given that the programme was shown to be "life-changing for most" and is "highly valued by participants and their families," the evaluation and recommended expanding and integrating a minimum package of adolescent development services through schools, while also providing safe spaces for adolescents who are out of school. The endline was both summative and formative in nature, and its purpose included providing insights for potential scaling up of the programme, reflecting its objective to demonstrate how a scalable programme can provide immediate and long-term nutrition and health benefits to adolescent girls. These findings directly informed the design and rationale of the four-year pilot 'Building Resilience for Adolescent Development Programme' (2023-2027) to further enhance adolescent development, the programme which is the object of this Inception Report.[32]

The programme, which is financially supported by BMGF and has a total budget of 4,500,000 USD, aligns with both national UNICEF's priorities. It reflects Ethiopia's commitment to adolescent development, as outlined in the National Food and Nutrition Strategy (2022–2030), the National Adolescent and Youth Health Strategy (2021–2025), the National Youth Policy, and broader RMNCAH-N strategies, all of which prioritise adolescent health, nutrition, and development. The programme is also informed by UNICEF's global and regional strategic

[32] It is important to note that, because the previous evaluation could not directly attribute impact of observed outcomes to the programme, the current initiative also presents an opportunity to generate evidence on the effects of such an integrated intervention packages on adolescents' health, nutrition, WASH, protection and education outcomes.

frameworks, including the UNICEF Gender Policy (2021–2030), Gender Action Plan (2022–2025), and Adolescent Girls Programme Strategy (2022–2025). Through a coordinated package of 1) **Health**, 2) **Nutrition**, 3) **Education**, 4) **SBC**, 5) **CP**, 6) **WASH**, and 7) **MHPSS** interventions, the programme adopts the multisectoral, rights-based, and equity-focused approach promoted by the UNICEF Ethiopia Country Programme. It emphasizes convergence across sectors and integrates cross-cutting strategies such as SBC and gender equality to support holistic adolescent development with schools as the primary delivery platform.

2.3.2 Programme Packages, Interventions, and Target Population

The programme is composed of 6 distinct intervention packages that integrate components from 7 different sectors. Each package builds on a common core **Basic Package**, which includes interventions in **Health**, **Nutrition**, **Education**, and **SBC**. Additional components are layered onto this core to form more comprehensive packages, as outlined below:

1. Basic Package: Health + Nutrition + Education + SBC
2. Basic + CP
3. Basic + WASH
4. Basic + CP + WASH
5. Basic + CP + MHPSS
6. Full Package: Basic + CP + WASH + MHPSS

While the primary targets of the programme are adolescent girls enrolled in school and out-of-school (OOS) from the following age brackets, 9-14 and 10-19, certain interventions have different targets. Table 1 below identifies beneficiaries for each sectoral intervention. A more comprehensive Table compiling essential programme information can be found in **Error! Reference source not found.** in the **Error! Reference source not found.** Annex.

Table 1: Interventions & Target Groups

Target	Sector	Interventions
Adolescent girls 9-14 in school and OOS	Health	HPV vaccine; counselling;
Teachers - Health care providers	Health	Training/capacity building on SRH
Adolescent girls 10-19 in school and OOS	Nutrition	Weekly Iron Folic Acid Supplementation (WIFAS)
	Cross-cutting	Referral mechanisms
Adolescent girls 10-19 OOS	Protection	Girls Power Training Manual
Adolescent girls 9-14 in school	WASH	MHH emergency supplies (pads and underwear)
	WASH & Health	Counselling
Adolescent girls and boys 9-14 in school	Education	Life Skills Training
Adolescent girls and boys 9-14 in school and OOS	WASH	Personal hygiene promotion
	Nutrition	Deworming
		Anthropometric measurement
Adolescent girls - 10 to 14 years old - enrolled in school	WASH	Menstrual Health and Hygiene (MHH) Supplies - pads and underwear

Adolescent girls and boys 13-15 in school	SBC	Edutainment – Yegna
Information pending	MHPSS	Community-Based Psychological Interventions for Adolescents
		Teacher capacity development in low-intensity support, early case identification and referral
		Low-intensity MHPSS support to in-school adolescents
Primary schools - girls' toilet primary focus (if budget allows, also boys')	WASH	Rehabilitation of WASH Facilities in Schools
Primary schools	WASH	Safe spaces rehabilitation (washing and changing/resting units)
Caregivers and wider community members	SBC	Community dialogues

Further, the programme provides for: (i) training of adolescent service providers (health extension workers and teachers); (ii) adolescent training (MHPSS, protection, nutrition); and (iii) the rehabilitation of school WASH facilities and safe spaces (washing and resting units). Community dialogues are also foreseen as part of the SBC component to create a conducive environment in the target locations.

The programme's selection of intervention is informed by prior programmes and evaluations in Ethiopia and further supported by evidence-based literature. For instance, the inclusion of WIFAS in the package of nutrition interventions is motivated by evidence that it improves iron status and reduces anaemia among adolescents, while also being associated with better mental health and enhanced academic performance.[33] In the case of WASH activities, beyond the findings from the evaluation of the previous programme in Ethiopia, global meta-analyses show that WASH interventions improve health outcomes, reduce school absenteeism, and contribute to a more conducive learning environment. [34] Life Skills training delivered through gender clubs and structured after-school programmes is backed by findings that show improvements in social behaviours, academic achievement, and reduced problem behaviours.[35] Also, child protection interventions targeting out-of-school girls, including life skills training and community dialogue, are effective in reducing child marriage and teen pregnancy and promoting educational attainment and empowerment.[36]

Other components of the package are likewise supported by evidence. MHPSS interventions, including school-based counselling, community engagement, and awareness-raising activities, have been found to reduce Post-Traumatic Stress Disorder (PTSD) and anxiety symptoms, while enhancing social support and coping skills in low-resource and crisis-affected settings.[37]

[33] (Kedir et al. 2024; Gosdin et al. 2021)
[34] (Andres et al. 2018)
[35] (Durlak, Weissberg, and Pachan 2010; Ciocanel et al. 2017)
[36] (Feyissa et al. 2023)
[37] (Bosqui, Mayya, and Al-Kadasi 2022; Purgato et al. 2018; Yu et al. 2023)

Similarly, school-based mental health interventions have been shown to enhance emotional well-being, reduce psychological distress, and support academic success, when integrating structured mental health programs—such as Cognitive behavioural therapy (CBT) and mindfulness—into school systems, particularly in low-resource settings where access to mental health services is limited.[38] Collectively, these evidence-based interventions provide a strong foundation for the programme's integrated approach to adolescent development.

2.3.3 Geographic Coverage

The programme is being implemented in 100 schools across the following locations: Aysaita (Afar), Homosha (Benishangul Gumuz), Mareko SPW (Central Ethiopia), Kersa and Gomma (Oromia), Teticha (Sidama), and the Dire Dawa City Administration. A breakdown of package types being delivered in each of the regions can be found in Table 2. Programme target locations were selected to ensure context diversity (e.g. pastoral, semi-pastoral, agrarian, urban), thus allowing for the assessment of how the pilot works in diverse settings. Schools were selected based on the following criteria:

- School-Level Capacity and Readiness, including infrastructure availability
- Schools that are reasonably accessible for the delivery of supplies, training, and monitoring
- Security situation
- Geographical representation
- Partnership opportunities

Table 2: Package and Geographic Coverage

Location	Basic Package	Basic + CP	Basic + WASH	Basic + WASH, CP	Basic + CP, MHPSS	Full Package	Total Schools per Woreda
Kersa	3 schools	5 schools	2 schools	-	2 schools	3 schools	**15 schools**
Gomma	5 schools	5 schools	-	-	-	5 schools	**15 schools**
Aysaita	10 schools	-	-	-	-	5 schools	**15 schools**
Homosha	5 schools	6 schools	-	4 schools	-	-	**15 schools**
Teticha	5 schools	2 schools	-	4 schools	3 schools	1 school	**15 schools**
Mareko SPW	5 schools	4 schools	-	2 schools	1 school	3 schools	**15 schools**
Dire Dawa City Administration	8 schools	-	-	-	-	2 schools	**10 schools**

[38] (Fazel et al. 2014; Zhang, Wang, and Neitzel 2023)

Total	**41 schools**	**22 schools**	**2 schools**	**10 school**	**6 schools**	**19 schools**

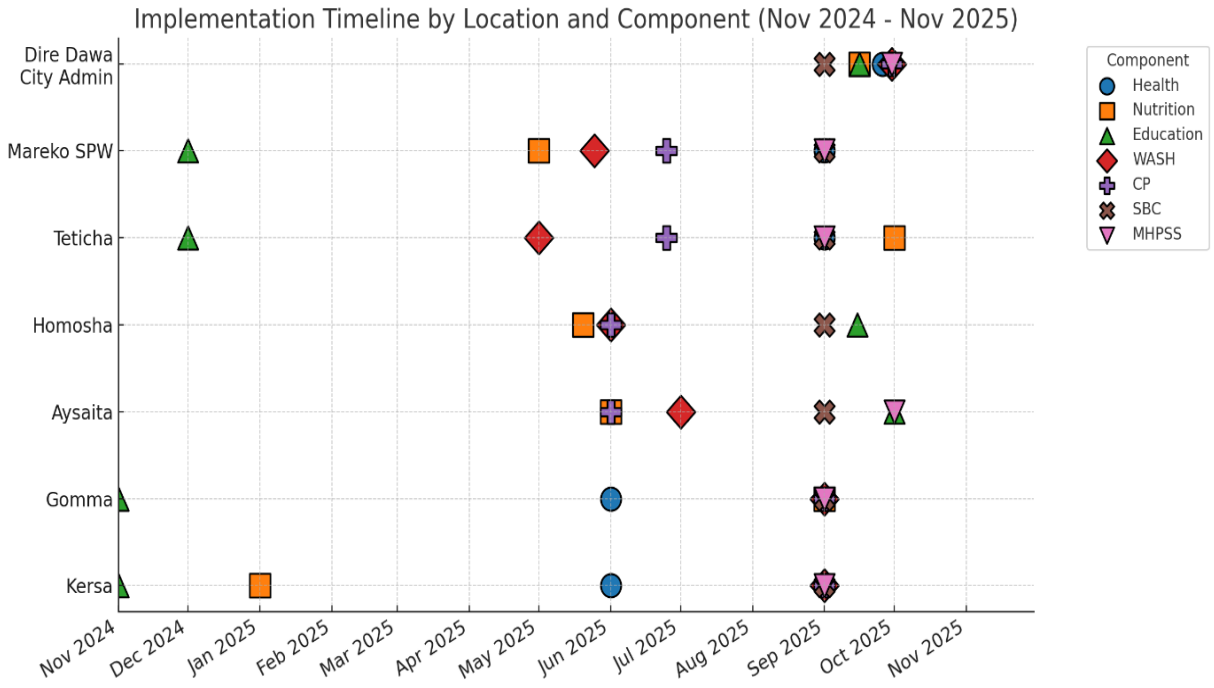
2.3.4 Stakeholder/Implementation Modality

UNICEF Ethiopia coordinates with national and international partners to design, implement, monitor, and evaluate the pilot. The primary national stakeholders are the Federal Ministry of Health (MOH), Ministry of Women and Social Affairs (MOWSA), and Ministry of Education (MOE), and their respective sub-federal structures (i.e. regional, city administration, zonal and woreda levels). Girls Effect, an International Non-Governmental Organisation (INGO), is responsible for implementing and monitoring the SBC edutainment component. Finally social workers, health workers, teachers and other local officials support the implementation at the grass-root level. Interventions are designed to leverage existing structures and personnel. For example, gender clubs are key platforms for service delivery, whereas teachers, health extension workers, and social workers are involved in service delivery and monitoring after having received specific training.

2.3.5 Programme Progress

As of June 2025, the programme team has concluded the groundwork for active implementation. The implementation of the different components has been phased, with each one rolling out at different times and in different schools, rather than all being introduced at once.

Figure 1: Programme Implementation



The graphic shows that Education, Nutrition[39] and WASH[40] were piloted early in select Woredas (e.g. Education in Gomma and Kersa in late 2024; Nutrition in Mareko SPW in May 2025; WASH in Teticha in May 2025), but most components roll out in June, coinciding with the end of the school year, and again in September with the new term. Chronologically, the last interventions (Nutrition in Teticha and Education in Aysaita) will not launch until October 2025. Moreover, as part of the WASH component, halfway through the pilot timeline, target schools will develop a business plan (e.g., training adolescents to prepare pads and underwear, setting up micro-businesses, etc.) to replenish MHH emergency kits in safe spaces for sustainability purposes. The temporal scope of the evaluation will thus run from September 2025 up to December 2027.

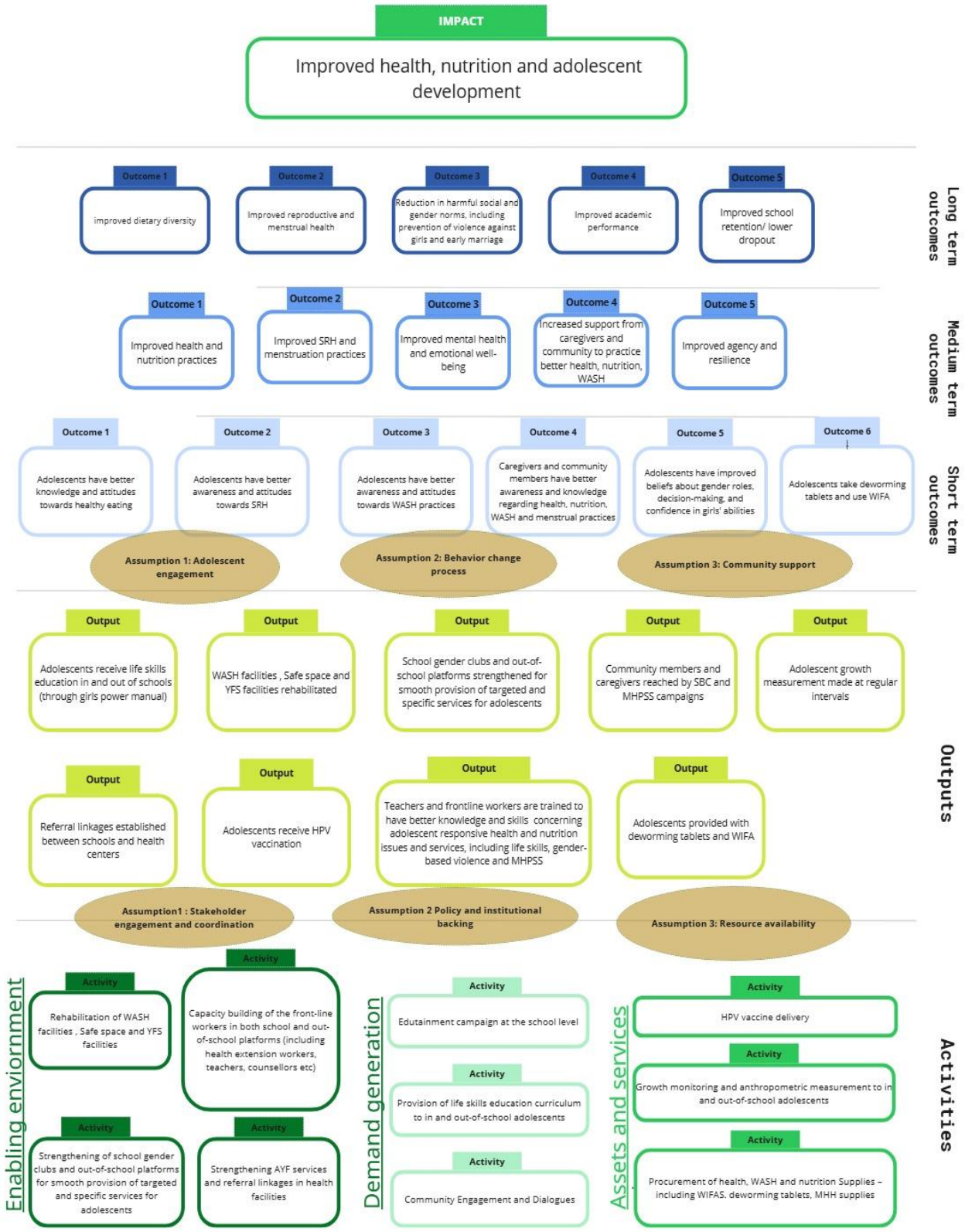
3 Theory of Change and Integrated Results Framework

The evaluation team used the pre-inception phase to review the programme's existing theory of change. Although the latter contained an overview of the key activities and outcomes, it lacked a detailed listing of activities and outputs and pathways to change that are expected to drive outcomes. Moreover, the existing theory of change was assessed to be more a general framework to guide the investment grant of the programme rather than one specifically focused on the minimum package of interventions. The team therefore worked on reformulating and expanding the ToC based on the existing one whilst incorporating inputs from the results framework and information obtained from stakeholder consultations, as well as an initial desk review. This involved adding detail on the expected outputs and adding a list of expected outcomes across the short, medium and long terms.

[39] WIFAS distribution has already started, whereas anthropometric measurement and deworming are yet to be implemented.

[40] Cost-assessments for structural interventions have concluded, and the actual rehabilitation of safe spaces and WASH facilities has begun.

Figure 2: Theory of Change



3.1 Theory of Change Narrative

3.1.1 Problem Statement

Adolescents (both in-school and out-of-school) face a range of interrelated challenges in health and development. Many adolescents suffer from poor nutrition and micronutrient deficiencies, leading to issues such as low Body Mass Index (BMI) and anaemia. Rates of early marriage and teenage pregnancy remain high in some communities, indicating gaps in reproductive health knowledge and agency among youth. Menstrual hygiene is often insufficient due to limited awareness and inadequate WASH facilities, contributing to health risks and school absenteeism for girls. Mental well-being is another concern; adolescents frequently lack psychosocial support and life skills, facing stress, GBV, and harmful gender norms that undermine their confidence and decision-making. These problems are compounded by low community awareness and engagement – caregivers and community leaders may not fully understand or support adolescent health needs. Adolescent with disabilities are also affected by these vulnerabilities, and are exposed to heightened risk, especially due to the lack of serviced adapted to their conditions. Without intervention, these issues contribute to a cycle of poor health outcomes, limited educational attainment, and constrained opportunities for adolescents, especially girls.

3.1.2 Pathways of Change

The program's Theory of Change is built on several pathways that link interventions to desired outcomes, ensuring that each component contributes to the overarching goal of improved adolescent health and development. **At the foundation**, the program provides adolescents with knowledge, life skills, and services across nutrition, reproductive health, WASH, and psychosocial support. By **increasing knowledge and awareness** (for example, through life skills education, health clubs, and edutainment campaigns), adolescents begin to adopt healthier attitudes and practices. They learn about proper nutrition (e.g. healthy eating habits and the importance of micronutrients), understand reproductive health and menstrual hygiene (including the benefits of HPV vaccination and menstrual hygiene management (MHM), and develop critical life skills such as communication, negotiation, and confidence.

In tandem, the program works on **improving services and the enabling environment**. Schools and health facilities are strengthened – WASH facilities are rehabilitated to provide safe water and sanitation, and youth-friendly health services (YFS) are enhanced with better referral linkages. Adolescents are given direct support, like WIFAS and deworming tablets, to tackle anaemia and parasitic infections immediately. Community engagement campaigns raise awareness among caregivers and community members about adolescent health, nutrition, mental health and the importance of supporting young people (for instance, discouraging early marriage and GBV). Trained teachers, health workers, and counsellors form a support network around adolescents, ensuring that guidance and services are accessible.

Through these pathways, **short-term changes** occur: adolescents, including those from marginalised groups, start using the new knowledge and resources available to them. For example, they make use of improved WASH facilities, take their nutritional supplements, and

participate in life skills sessions. Simultaneously, caregivers and community leaders become more aware of and supportive toward adolescent needs. These short-term outcomes set the stage for **medium-term changes**: with ongoing practice and reinforcement, adolescents develop sustained healthy behaviours and positive attitudes. They become more aware of health issues (like nutrition, hygiene, and vaccination) and more equitable in their gender attitudes. Girls and boys gain confidence and learn to make informed decisions regarding their health and education.

As these behavioural and social changes take root, the program expects to see **long-term outcomes**. Adolescents, including the most vulnerable, consistently practice good health and nutrition habits, exercise their sexual and reproductive health rights responsibly, and demonstrate improved resilience and life skills (e.g. effective decision-making and communication). We anticipate more adolescents, particularly girls, will remain in school and delay marriage, supported by a community that values their well-being and rights. This supportive environment and improved agency leads to measurable improvements such as better nutritional status (healthy BMI and micronutrient levels), improved mental health, and lower prevalence of issues like teenage pregnancy, GBV, and HPV. All these pathways of change converge toward the **impact** – a generation of adolescents who are healthier, empowered, and able to make positive life choices that benefit themselves and society.

3.1.3 Long-Term Outcomes

By the end of the program (long-term), the following outcomes are expected to be realized among adolescents, especially the most vulnerable (such as those with disabilities) and their communities:

- **Improved dietary diversity –** Adolescents consistently consume a wider range of nutritious foods, supported by family and community practices**.** This outcome is supported by nutrition education, growth monitoring, and the uptake of food-based and supplement-based interventions.
- **Improved reproductive and menstrual health –** Adolescents, especially girls, have greater access to menstrual hygiene products and safe sanitation facilities, demonstrate safe menstrual hygiene practices, and are empowered to manage their reproductive health effectively leading to better HPV vaccination rates and reduced teenage pregnancy among others.
- **Reduction in harmful social and gender norms, including prevention of violence against girls and early marriage –** Norms shift towards equity, reducing tolerance for violence and child marriage and school dropout.
- **Improved school retention and lower dropout rates:** More adolescents remain in school through secondary level. Girls, in particular, are less likely to drop out due to marriage, pregnancy, or lack of facilities, as interventions like life skills, community support, and WASH facilities encourage continued education.
- **Improved academic performance –** Better health, nutrition, and emotional well-being translate into stronger school engagement and learning outcomes.

These long-term outcomes together contribute to healthier, safer, and more equitable life trajectories for adolescents, aligning with the program's ultimate impact of improved adolescent health and development.

3.1.4 Medium-Term Outcomes

In the medium term, the program aims to achieve several outcomes that indicate shifts in knowledge, attitudes, and early behaviour change among adolescents – including the most marginalized- and stakeholders:

- **Improved health and nutrition practices:** Adolescents consistently adopt healthier eating habits and lifestyle practices, resulting in better overall nutrition and health. This includes maintaining a healthy BMI and consuming a diet rich in necessary nutrients.
- **Improved SRH and menstruation practices:** Adolescents, especially girls, practice proper menstrual hygiene and make informed choices about their sexual and reproductive health. They have access to and utilize menstrual hygiene products and SRH services, leading to better reproductive health outcomes.
- **Improved mental health and emotional well-being:** Adolescents experience better mental health outcomes, evidenced by higher self-esteem, lower stress or anxiety levels, and greater access to psychosocial support. They are more resilient in the face of challenges, partly due to the safe spaces and counselling services made available.
- **Increased support from caregivers and the community for healthy practices:** Families and community members actively support and reinforce adolescents' healthy behaviours. Social norms shift to favour girls' education, proper nutrition, sanitation, and the rejection of harmful practices like early marriage or GBV. This creates an enabling environment that sustains the positive changes initiated by the program.
- **Enhanced knowledge and attitudes on healthy eating:** Adolescents acquire knowledge about nutrition and healthy diets and begin to favour nutritious food choices. They understand the value of a balanced diet (including micronutrients from diverse foods or supplements) and are more willing to consume things like iron and folic acid supplements or deworming tablets for their health.
- **Improved agency and resilience among adolescents:** Adolescents have a stronger sense of agency – they feel empowered to shape their own lives. They can advocate for themselves and their peers, bounce back from setbacks, and are capable of handling social pressures. This resilience and empowerment means they are more likely to continue healthy behaviours and mentor others, ensuring the longevity of the program's impact.

These medium-term outcomes are critical steps in the change process. They reflect that adolescents are not only informed but also beginning to internalize positive behaviours and norms. Such changes set the groundwork for the behavioural practices and system-level shifts captured in the long-term outcomes. For example, once adolescents have knowledge and positive attitudes (medium-term), they are far more likely to practice healthy behaviours and influence peers and family (long-term). Similarly, awareness and acceptance of concepts like

HPV vaccination or girls' education in the medium-term lead to concrete actions (like actually getting vaccinated or keeping girls in school) as the program progresses.

3.1.5 Short-Term Outcomes

In the short term (during and immediately after the implementation of activities), the program expects to observe initial outcomes that signal the uptake of new knowledge and services:

- **Improved attitudes and practices towards WASH:** Adolescents develop positive attitudes and habits regarding water, sanitation, and hygiene. They understand the importance of practices like handwashing with soap, using latrines, and keeping their environment clean, which is reflected in their daily routines at school and home.
- **Improved knowledge and attitudes towards healthy eating:** Adolescents are better aware of the positive impact of healthy diets and the types of food to eat. They understand the importance of taking WIFAS and deworming tablets.
- **Uptake of deworming and micronutrient supplements:** Adolescents begin taking deworming tablets and WIFAS as provided by the program. This short-term outcome shows compliance and trust – adolescents understand the importance of these supplements for their health and start incorporating them into their routine (often with support from teachers or health workers during school health days). Improved supplement uptake in the short term is likely to lead to better nutritional status in the longer term.
- **Increased awareness of HPV and uptake of vaccination:** Adolescents (and their parents) become aware of HPV and its link to cervical cancer, and they gain knowledge about the benefits of HPV vaccination. This results in greater acceptance and demand for the HPV vaccine among eligible adolescents.
- **Shift in gender norms and beliefs:** Boys and girls show improved beliefs about gender roles – for instance, greater acceptance of girls' right to make decisions about their lives and confidence in girls' abilities. Adolescents start challenging harmful gender stereotypes; girls become more confident in voicing their opinions, and boys become more supportive of gender equality.
- **Improved caregiver and community knowledge:** Caregivers, parents, and community members exhibit better awareness and knowledge regarding adolescent health, nutrition, WASH, and menstrual hygiene practices. In the short term, this might be seen through increased attendance and engagement in community dialogues, asking questions about these topics, or willingness to allow adolescents (especially girls) to participate in program activities. This outcome is crucial because it indicates the community is beginning to buy into the program's messages. With enhanced knowledge, community members are more likely to support adolescents in practicing what they learn (for example, ensuring their daughters have menstrual pads or encouraging their sons to respect girls).

These short-term outcomes serve as early signs of progress, validating that the program's activities are reaching the target groups and that behaviour change is starting. They are relatively immediate and observable changes that the program can measure in real-time (e.g.,

tracking supplement distribution and usage, facility usage logs, pre/post session feedback from community workshops). Achieving these short-term outcomes builds momentum – they are the first crucial steps that lead into deeper attitude change (medium-term outcomes) and sustained behaviour change (long-term outcomes).

3.2 Outputs and Key Activities

To realize the above outcomes, the program delivers a set of outputs through carefully designed activities. The outputs are the direct results of program activities – the goods, services, and structures put in place – while the activities are the actions or interventions undertaken by the program. Together, these ensure that adolescents and their communities have the necessary tools, knowledge, and environment to change behaviours. Below are the key outputs of the program and the major activities that produce them:

3.2.1 Key Outputs

Life skills education for adolescents: Adolescents, including those with disabilities, receive comprehensive life skills education both in schools and through out-of-school platforms (using a tailored "Girls' Power" curriculum and similar tools). This output equips young people with knowledge on health, nutrition, gender equality, and personal development, as well as skills like critical thinking, communication, and problem-solving.

Rehabilitated WASH and safe spaces: Water, sanitation, and hygiene facilities in target schools (and any dedicated youth spaces) are rehabilitated or upgraded. This includes ensuring schools have gender-segregated latrines with water supply, handwashing stations, and disposal facilities for menstrual hygiene products. Additionally, safe spaces and YFS) centres are set up or improved to provide adolescents with secure, supportive environments for learning and receiving services (e.g., a private room for counselling or health check-ups).

Strengthened adolescent support platforms: School-based gender clubs and OOS youth platforms are strengthened and actively functioning. These clubs/platforms provide peer support networks and regular forums where adolescents can discuss issues, receive mentorship, and access information or services. A strengthened and inclusive club might have trained facilitators, a structured curriculum of activities, and linkages to health or counselling services.

Community and caregiver engagement reached: A significant number of community members and caregivers are reached by SBC communication and MHPSS campaigns. This output indicates the roll-out of community dialogues, media campaigns, or group sessions that raise awareness about adolescent health and development issues (such as nutrition, puberty, mental health, gender-based violence, and the value of keeping girls in school). It also covers any community-level psychosocial support initiatives that help create a more understanding and supportive environment for adolescents, especially those with disabilities.

Regular growth monitoring of adolescents: Systems are in place for regular growth measurements (height, weight, BMI) and health check-ups of adolescents at designated intervals. This output may involve school health days or clinics measuring adolescents' growth

and screening for nutrition or health issues. The data gathered not only helps in tailoring interventions (like identifying students who need extra nutrition support) but also engages adolescents in understanding their own health status.

Provision of deworming and WIFA supplementation: Adolescents are provided with deworming tablets and WIFAS on a routine basis. This output ensures that essential health supplies reach the target group to address anaemia and worm infestations, which are common health problems affecting adolescent growth and concentration.

Trained teachers and frontline workers: Teachers, health extension workers, counsellors, and other frontline personnel are trained to be more knowledgeable and skilled in adolescent-responsive health and nutrition services. This includes training on delivering life skills education, inclusive education, addressing gender-based violence, providing basic mental health and psychosocial support, and managing adolescent-friendly health services. As an output, it means a cadre of adults around adolescents are equipped to guide and support youth effectively.

HPV vaccination of adolescents: Adolescent girls receive the HPV vaccine as part of the program's support to national immunization efforts. Achieving this output means coordinating with health authorities to deliver vaccination drives or referrals so that adolescents are protected against HPV, reducing future risk of cervical cancer.

Established school-health referral linkages: Formal referral linkages are established between schools and health centres (including youth-friendly clinics). This output might include referral protocols, communication channels, or partnership agreements that ensure adolescents identified with health or psychosocial needs in the school setting can be referred to appropriate health facilities and receive care. It strengthens the continuum of care so that issues spotted by teachers or through school health screenings are followed up by medical professionals.

3.2.2 Key Activities

The achievement of outputs are based on a series of activities set out by the programme, these are described in detail in previous sections and can be categorised into the following categories:

Demand Generation	Provision of supplies and supply-side investments	Enabling Environment

• Life skills education sessions • School-based edutainment campaigns • Life skills education sessions	• Rehabilitation of facilities • Capacity building of frontline workers and officials • Growth monitoring and health screening • Provision of health and hygiene supplies • Provision of health and hygiene supplies	• Community engagement and dialogues • Strengthening youth clubs and platforms • Strengthening adolescent-friendly health services

All these activities are interconnected and mutually reinforcing. For instance, the success of life skills education is amplified by community dialogues that get parents on board, and the impact of WASH facility upgrades is maximized when paired with hygiene education and available supplies. By delivering these outputs and activities in a coordinated manner, the program ensures that knowledge translates into practice, and that adolescents have both the **motivation and the means** to improve their health and development.

3.3 Assumptions and Risks

The Theory of Change is based on certain assumptions about the context and includes awareness of potential risks that could affect implementation. It is important to make these explicit, as they represent conditions needed for success and factors that may require mitigation strategies.

3.3.1 Key Assumptions

Community and Stakeholder Support: It is assumed that parents, community leaders, schools, and local health authorities remain supportive of the program's objectives. The theory assumes that communities are willing to participate in dialogues and that cultural or religious norms will allow discussions on topics like reproductive health and gender equality. Supportive stakeholders are crucial for sustaining activities (for example, allowing vaccinations in schools or encouraging girls' participation in clubs).

Policy and Institutional Backing: The program's success assumes a stable policy environment where government policies favour adolescent health and development (such as existing mandates for life skills education or WASH in schools). It also assumes that schools and health facilities are cooperative and that trained staff remain in their positions. For instance, it's assumed that once teachers and health workers are trained, they will continue to apply their skills and not be transferred away immediately.

Resource Availability: A key assumption is that essential resources and supplies will be available when needed. This includes a steady supply of WIFA supplements, deworming tablets, vaccines, and sanitary materials, as well as the funding and materials to maintain facilities. The plan assumes that there won't be major supply chain disruptions and that the necessary financial resources will remain in place throughout the program duration.

Adolescent Engagement: It is assumed that adolescents will be willing and able to engage with the program – attending sessions, taking supplements, and utilizing facilities. The ToC assumes that the program content is culturally appropriate and interesting enough to keep them involved. In other words, the theory counts on adolescents' motivation to improve their own lives when given the opportunity and information.

Behaviour Change Process: The Theory of Change assumes that increases in knowledge and improvements in the environment will indeed lead to behaviour change, as suggested by behaviour change theory. It presumes that if adolescents know more and have better options (like facilities and supplies), they will choose healthier behaviours, and that caregivers with greater awareness will support rather than obstruct these behaviours. While generally valid, this assumption requires that the messaging and interventions are effective in addressing deeper drivers of behaviour (such as self-efficacy and social norms).

3.3.2 Key Risks

Cultural Resistance and Social Norms: One risk is that deep-rooted cultural norms and practices (like preference for child marriage or stigma around discussing sexuality) could limit the program's effectiveness. Even with awareness activities, some community members might resist change, potentially discouraging adolescents from participating or applying new knowledge. For example, if community leaders do not buy in, they might oppose life skills classes or misinterpret the program's intent. The program mitigates this by involving community influencers early, but the risk remains that change in norms can be slow.

Economic and Security Factors: Poverty and economic pressure pose a risk to outcomes such as school retention and nutrition. If households face economic crises, adolescents (especially girls) might still drop out of school to work or marry despite the program. Similarly, any deterioration in security (conflict, unrest) or public health emergencies (like a pandemic) could disrupt school schedules, supply chains, and community gatherings, hindering activity implementation. Such external shocks are beyond the program's control but could significantly impact results.

Drop-off in Participation: There is a risk that some beneficiaries do not continue participating fully – for instance, adolescents might lose interest in clubs, or trained teachers might revert to old teaching methods without continuous support. Ensuring consistent engagement is a challenge; if not maintained, the desired outcomes (like sustained behaviour change) may not materialize.

Capacity and Maintenance Issues: Another risk is that infrastructure and capacity gains might not be maintained. For instance, WASH facilities require upkeep – if schools lack funds or commitment for maintenance, facilities could fall into disrepair, reducing their use. Trained

personnel might be overburdened or could leave without replacements equally trained, causing gaps in service quality. The program must rely on local systems to carry forward the improvements; if those systems are weak, there's a risk that some outputs (like clubs or services) wane after direct support ends.

Coordination and Referral Challenges: While the program establishes referral linkages and multi-sector collaboration, there is a risk that coordination breaks down. Health centres might be understaffed or lack youth-friendly services despite training, leading referred adolescents to drop off the care continuum. If communication between schools and clinics isn't consistently maintained, some adolescents needing help might fall through the cracks.

4 Evaluation Purpose

This evaluation occurs at a particularly suitable moment, as it launches alongside the programme's rollout, capturing baseline, midline and endline data to fulfil the evaluation purposes described below, and will generate the actionable insights needed to guide and inform any future scale-up.

The purpose of this evaluation is twofold: Learning and Accountability.

Learning: This evaluation will assess both the design and implementation of the pilot through the lens of the Organisation for Economic Co-operation and Development – Development Assistance Committee (OECD-DAC) framework. The project will examine how the convergence of interventions across Health, Nutrition, Education, SBC, protection, WASH, and MHPSS produces benefits for adolescents, by delving into the various service packages delivered for adolescent development to identify which work and which require improvement. The evaluation will assess the extent to which the programme has achieved its preset objectives, intended and unintended outcomes, as well as the underlying TOC to identify gaps and refine it accordingly. This assignment will identify best practices and mitigation strategies to inform programme adjustments that better fulfil the rights of adolescents. A particular focus will be dedicated to evaluating the scalability as well as operational components for long-term sustainability, building on the positive impacts on adolescents shown by an earlier evaluation of the Ethiopia Joint Adolescent Nutrition-WASH-Education programme. Specifically, the evaluation aims to determine the appropriate inputs in terms of policy, human resources, financing, stakeholder engagement, and intervention design to cost-effectively sustain the results gained through the programme and potentially reach adolescents nationwide through similar school-based strategies. The evaluation team will also attempt to explore the programme's contribution to school retention and absenteeism, investigating whether the school health package contributes to keeping students—especially girls—engaged in education.

Accountability: By engaging right-holders and duty-bearers, the evaluation will gather perspectives on the programme's added value to their lives and well-being, ensuring their inputs guide future programming and resource allocation. Moreover, findings and lessons

learned will be actively disseminated through knowledge management systems and stakeholder forums, demonstrating to donors and partners the appropriateness, effectiveness, and impact of their investment.

Evaluation users were determined to be the following, based on interviews and review of documents shared at the pre-inception phase:

- UNICEF Ethiopia
- Federal Ministry of Health
- Federal Ministry of Women and Social Affairs
- Federal Ministry of Education
- Regional Health Bureaus
- Regional Education Bureaus
- Regional Bureau of Women and Social Affairs
- Donor - Bill & Melinda Gates Foundation
- Programme beneficiaries – i.e. adolescents in target woredas
- Girl Effect and other implementing partners

Evaluation Use: Stakeholders' potential use of the evaluation can be summarised as below:

Table 3: Stakeholders & Potential Use of the Evaluation

Stakeholder	Potential use
Federal and sub-federal government	• Use evidence to eventually guide national policy and resource allocation for the scale-up of the pilot • Incorporate evaluation insights into sectoral strategies and regional roll-out plans • Adapt operational and sustainability inputs (human resources, financing, stakeholder engagement) at the federal and regional levels
Donors and implementing partners	• Use evidence to gauge the sustainability and possible scale-up of the programme • Understand how different packages influence school retention • Fulfil donor accountability and cost-efficiency requirements through robust findings and recommendations • Inform future funding decisions and philanthropic strategy
UNICEF Ethiopia	• Adjust programme design and delivery • Advocate with the Government of Ethiopia (GoE) and partners for sustained investments in adolescent services and system strengthening • Feed lessons into internal knowledge-management systems and strategic decision-making for ongoing implementation and scale-up
Beneficiaries	• Benefit directly from improved, context-adapted service packages • Participate in feedback loops (via AYF) to ensure their needs drive operational refinements

5 Evaluation Objectives and Questions

5.1 Evaluation Objectives

The evaluation will focus on the following key objectives, as outlined in the Terms of Reference (ToR) attached:

- Assess the extent to which the programme has achieved its preset objectives, intended and unintended outcomes, with a deep dive on the impact on girls, children with disabilities, out-of-school adolescents, persons with disabilities and other marginalized individuals (Effectiveness)
- Document best practices, benchmarks and mitigation strategies to produce recommendations that can practically inform the continuous adjustment of the programme to respond contextually and relevantly to the evolving needs of the target population
- Examine the coherence and multiplier effect of multisectoral collaboration for adolescents, assessing how the combination of approaches employed contributed to enhanced health and nutritional outcomes for adolescents
- Determine the appropriate inputs in terms of policy, human resources, financing, engagement of stakeholders and intervention design to cost-effectively sustain the results gained through the programme and potentially reach adolescents nation-wide through similar school-based strategies
- Assess and analyse the Theory of Change by incorporating evaluation insights and stakeholder perspectives, showing how effectively or not the causal linkages work, identifying necessary gaps and preconditions for improved design or further scale-up
- Evaluate how gender, disability inclusion, and the needs of marginalized adolescents (girls and boys) are addressed within programme activities, assessing the effectiveness of these efforts in promoting equitable access, participation, and outcomes

5.2 Evaluation Questions

UNICEF Ethiopia has proposed in the Terms of Reference (TOR) a set of preliminary questions as a starting point to support the evaluation design process. Based on the evaluability assessment, the evaluation team has revised the preliminary set of questions by adding, modifying, merging, and reclassifying some of them, as indicated in the table below.

New question
What is the added value of the packages containing additional interventions compared to the basic one?
Modified question
To what extent have the programme ~~objectives~~ results been achieved in each site and across different sectors? Were they achieved on time?

Overlapping questions[41]
What have been the major factors influencing the achievement or non-achievement of the programme objectives in providing integrated services for boys, girls and vulnerable groups (including Children with Disabilities (CWDs)? Overlaps with *Was the programme implemented as planned? How? Why/ Why not? What have been the main barriers, enablers and adaptations to the interventions?*
Have there been any significant delays in programme implementation and achievement of results, and if so, why? Overlaps with *To what extent have the programme results been achieved in each site and across different sectors? Were they achieved on time?* And *Was the programme implemented as planned? How? Why/ Why not? What have been the main barriers, enablers and adaptations to the interventions?*
How has the program influenced changes in national policies or practices beyond the immediate program goals? Overlaps with *To what extent is the joint programme supporting long-term buy-in and ownership by duty bearers and rights holders?*
How has implementation been adapted over the course of the programme and why? Are there disparities across or within regions in the level of implementation? Overlaps with *Was the programme implemented as planned? How? Why/ Why not? What have been the main barriers, enablers and adaptations to the interventions?*
What were the strengths and gaps in achieving coherence and adding value while avoiding duplication of effort? Overlaps with *How well do the programme interventions fit together, create synergies and coherence?* And *What is the extent of partnership, coordination, and complementarity with the interventions of the Ethiopian government and other relevant actors operating within the same context*

[41] Questions in italics were maintained

around similar interventions? Are there aspects of the operation that conflict with the interventions of or one-UN programming or other actors?
How have lessons learned from this programme been integrated into other similar programs either at district level or nationally? Overlaps with *What is the extent of partnership, coordination, and complementarity with the interventions of the Ethiopian government and other relevant actors operating within the same context around similar interventions? Are there aspects of the operation that conflict with the interventions of or one-UN programming or other actors?*
To what extent has the programming approach ensured access to interventions for people living with disabilities? Overlaps with *To what extent is programming conducive to supporting children with disabilities and from the most marginalised populations (including those furthest left behind)?*
Questions reclassified as sub-questions
To what extent do beneficiaries feel the programme addresses their needs that would otherwise have been unsupported? Classified as a sub-question of *How relevant are the integrated package interventions and services to the needs of adolescent boys and girls, including those with disabilities?*
To what extent have the strategies adopted by the programme contributed to the sustainability of results, especially equity and gender-related results? And To what extent is the joint programme supporting long-term buy-in and ownership by duty bearers and rights holders? Classified as sub-questions of *What key factors will facilitate the scale-up and sustainability of the package?*

As a result of the review, the final evaluation questions, categorised by evaluation criteria, are reported in the table below, whereas the evaluation matrix (see dedicated section below) outlines for each evaluation question, related sub-questions, indicators, data collection methods and sources, and proposed analytical methods.

Relevance
What are the deprivations and barriers in relation to adolescent services?

How relevant are the integrated package interventions and services to the needs of adolescent boys and girls, including those with disabilities?
Have services been fully adapted to meet the needs of different groups and in different target areas?
Effectiveness
What systems are in place to monitor the implementation and uptake of the program? Do monitoring systems generate reliable, accurate, and consistent data?
To what extent have the programme results been achieved in each site and across different sectors? Were they achieved on time?
Was the programme implemented as planned? How? Why/ Why not? What have been the main barriers, enablers and adaptations to the interventions?
Were there any unexpected consequences of the programme?
Efficiency
To what extent is the programming approach efficient in the achievement of desired results in terms of resource utilization (human, technical, financial) and timely delivery?
What are the costs of implementing the different interventions in the package? Are these activities low in cost and affordable?
Is the programme being delivered in a cost-effective manner? Are there disparities in cost-effectiveness across areas?
OR how cost-efficient is the programme and what potential is there for efficiency savings at all stages (capacity and constraints)?
Impact
To what extent has the program contributed towards impact level change? How and why? At what level, supporting which programme indicators?
What is the difference in impact indicator achievement in the different regions and woredas?
What is the added value of the packages containing additional interventions compared to the basic one?
Coherence
How well do the programme interventions fit together, create synergies and coherence?
What is the extent of partnership, coordination, and complementarity with the interventions of the Ethiopian government and other relevant actors operating within the

same context around similar interventions? Are there aspects of the operation that conflict with the interventions of or one-UN programming or other actors?
How efficient and effective was the collaboration among various ministries involved in the programme, and whether the programme was led by and involved all relevant departments to achieve the best results?
Sustainability
What key factors will facilitate the scale-up and sustainability of the package?
What enablers and barriers are there for the GoE and other national and sub-national stakeholders to sustainably deliver results for adolescents during and beyond the project timeframe?
Gender, disability, equity and human rights
To what extent has UNICEF and the government identified and addressed the specific needs of girls and boys in the design and implementation of the programme?
To what extent is programming conducive to supporting children with disabilities and from the most marginalised populations (including those furthest left behind)?
To what extent did programming incorporate a Human Rights-Based Approach in all its phases? Were duty-bearers and rightsholders involved in advocacy, design and implementation of the programme?

6 Evaluation Scope

EPRI is conducting an impact evaluation an evaluability assessment, a baseline assessment, a midline evaluation, and an endline evaluation with the following thematic, geographic, and chronological scope:

6.1 Thematic Scope

The evaluation will examine the relevance, effectiveness, efficiency, coherence, impact and sustainability of the BMGF funded adolescent health and nutrition programme. The evaluation will assess the adolescent health, nutrition, education, and gender transformative outcomes anticipated from the intervention under the various participating sectors with due attention given to adolescents with disabilities. The sustainability of the programme should also be investigated from an angle of cost to assess what resources will be required for the GoE to potentially continue and scale up the initiative nationally.

6.2 Geographic Scope

The programme is implemented in five regions of Ethiopia and one city administration, in six woredas and 100 schools. The targeted regions and woredas are Afar (Aysaita woreda),

Benishangul Gumuz (Homosha woreda), Central Ethiopia region (MarekoSPW woreda), Oromia (Kersa and Gomma woredas), and Sidama (Teticha woreda), and Dire Dawa city administration. The evaluation will encompass all implementation woredas.

6.3 Chronological Scope

The building resilience for adolescent development programme had planned to begin on the ground implementation from September 2025 after undergoing iterative design from November 2023. It is planned to span 3 years. The scope of the evaluation will be from its launch in November 2024 to December 2027.

7 Evaluation Framework

7.1 Evaluation Matrix

Criterion	Questions	sub-questions	Collection phase	Indicators	Data collection method	Data source	Analytical approach
Relevance	What are the deprivations and barriers in relation to adolescent services?	What demand-side barriers prevent adolescents from accessing dedicated services? How does the programme design consider these?	Baseline	[FGDs, Case studies AWDs] Barriers [Desk review] Strategies to mitigate demand-side barriers (e.g. parent engagement for sensitisation, awareness campaigns) [Survey] Barriers selected in the survey (disaggregated by age, gender, and disability)	Desk review Focus Group Discussions (FGDs) Student survey	Programme design documents Grey literature Adolescents Adolescents with disabilities (AWDs) Parents Community leaders frontline workers (social and health)	Content analysis Thematic Framework analysis Descriptive statistical analysis
		What supply-side barriers prevent adolescents from accessing the dedicated services? How does the programme design consider these?	Baseline	[Desk review] Documental evidence of strategies to mitigate supply-side barriers (e.g. frontline workers training(s), facilities rehabilitation, AYF services strengthening, etc.) [KIIs, FGDs] Barriers	Desk review Key Informant Interviews (KIIs) FGDs	Programme design documents Grey literature Local government officials, UNICEF field staff, head teachers, teachers, frontline workers (social and health)	Content analysis Thematic Framework analysis

Criterion	Questions	sub-questions	Collection phase	Indicators	Data collection method	Data source	Analytical approach
	How relevant are the integrated package interventions and services to the needs of adolescent boys and girls including those with disabilities?	What are the unique needs of: A) in-school girls (including those with disabilities); B) out-of-school girls (including those with disabilities); C) in-school boys (including those with disabilities);	Baseline	[Desk review] Documental evidence of needs [KIIs, FGDs and case studies AWDs] Needs [Survey] Needs selected in the survey (disaggregated by age, gender, and disability)	Desk review KIIs FGDs Student survey	Programme design documents Grey literature Adolescents AWDs) Parents Safe space coordinator Gender club head & members	Content analysis Thematic Framework analysis Descriptive statistical analysis
		How/on which basis were interventions, intervention packages, and locations selected?	Baseline	[Desk review] Selection criteria for interventions, intervention packages, and locations Documental evidence of contextual analyses, stakeholder analyses, and vulnerability assessments [KIIs] Selection criteria / rationale	Desk review KIIs	Programme design documents Grey literature Programme designers	Content analysis Thematic Framework analysis
		To what extent do beneficiaries feel the programme addresses their needs that would otherwise have been unsupported?	Midline and Endline	[FGDs and case studies AWDs] Previously unsupported needs addressed by the programme [Survey] % of students reporting that the programme addresses needs previously not supported (disaggregated by age, gender, and disability)	Student survey FGDs	Adolescents AWDs Student surveys	Thematic Framework analysis Descriptive statistical analysis

Criterion	Questions	sub-questions	Collection phase	Indicators	Data collection method	Data source	Analytical approach
	Have services been fully adapted to meet the needs of different groups and in different target areas?		Baseline and midline	[Desk review, KIIs] Documental and anecdotal evidence of differentiated/adapted contents for A) in-school girls (including those with disabilities); B) out-of-school girls (including those with disabilities); C) in-school boys (including those with disabilities); and D) locations;	Desk review KIIs	Programme design documents Programme designers	Content analysis Thematic Framework analysis
Effectiveness	What systems are in place to monitor the implementation and uptake of the program? Do monitoring systems generate reliable, accurate, and	How, by whom, and how frequently is programme implementation and performance monitored?	Baseline	[Desk review, KIIs] Documental and anecdotal evidence on type, timing and frequency of data collected; means and responsibilities for data collection; data storage and flow; monitoring strategy and/or data collection & analysis plan; pretesting of the monitoring strategy and/or plan;	Desk review KIIs	Project progress and donor reports Data platforms / administrative data Results framework and tracker Monitoring strategy / data collection & analysis plan UNICEF M&E Personnel, teachers, frontline workers (social and health), local government officials	Content analysis Thematic Framework analysis Contribution analysis

Criterion	Questions	sub-questions	Collection phase	Indicators	Data collection method	Data source	Analytical approach
	consistent data?	To what extent is data collection consistent across time and places? To what extent do the indicators measure the process or performance intended to be measured? To what extent is the collected data accessible and shared among programme stakeholders?	Midline and Endline	[Desk review] Documental evidence showing consistency of data collected over time and target locations; quality/types of process and performance indicators; evidence of knowledge/data sharing tools and/or events [KIIs] Anecdotal evidence on data collection challenges and horizontal/vertical information flows	Desk review KIIs	Project progress reports and donor reports Data platforms / administrative data Results framework and tracker Local government officials, UNICEF M&E and field personnel,	Content analysis Thematic Framework analysis
	To what extent have the programme results been achieved on each site and across different sectors? Were they achieved on time?	Did the intervention lead to the intended outputs and short-term outcomes?	Midline and Endline	[Desk review] [SBC]#In-school adolescents exposed to Yegna/edutainment (disaggregated by age, gender, and disability); #Community dialogues [Education]#Adolescents receiving life skills training; % of adolescents graduating from life skills training (Both disaggregated by age, gender, and disability) [Protection]#OOS Adolescent girls receiving capacity building; % of OOS adolescent girls graduating from capacity building (Both disaggregated by age, gender, and disability)	Desk review	Project progress and donor reports Data platforms / administrative data Results framework and tracker	Content analysis Descriptive statistical analysis

Criterion	Questions	sub-questions	Collection phase	Indicators	Data collection method	Data source	Analytical approach
				[Health]#Adolescents receiving counselling; #Adolescents receiving one dose of HPV vaccine; #Adolescents receiving full HPV vaccine coverage (two doses) #In-school adolescents benefitting from personal hygiene education (All disaggregated by age, gender, and disability) [Mental health / psychosocial wellbeing] #Students who accessed MHPSS support; (Disaggregated by age, gender, and disability) #Key MHPSS messages disseminated at the community level [Nutrition]#Adolescents undergoing anthropometric measurement; # Adolescent receiving deworming tablets once a year; #Adolescents receiving deworming tablets twice a year; (All disaggregated by age, gender, and disability) #Adolescent girls receiving IFA (Disaggregated by cycle/frequency, in-school vs OOS, age, and disability)			

Criterion	Questions	sub-questions	Collection phase	Indicators	Data collection method	Data source	Analytical approach
		To what extent did the SBC component change knowledge, attitudes, and practices of adolescents and community members?	Baseline, Midline, and Endline	[Survey] Knowledge, attitudes and practices concerning MHM, Hygiene, mental health, GBV, Child Marriage, Teenage pregnancy, and nutrition (All disaggregated by age, gender, and disability) [FGDs and case studies AWDs] Anecdotal evidence and contribution stories concerning changes in knowledge, attitudes, and practices on gender roles, nutrition, child marriage, teenage pregnancy, mental health, GBV, SRH, STIs, MHM	Student survey FGDs	Adolescent AWDs Parents Community leaders	Descriptive statistical analysis Thematic Framework analysis Contribution analysis
		To what extent did the life skills education and capacity building developed or strengthened skills in in-school and OOS adolescents?	Baseline and Endline	[Desk review] Life skills Training and OOS capacity building scoring (if available) [Survey] Scoring on scenario-based assessments on decision-making, communication, problem solving, emotional regulation, conflict resolution, empathy, and safety awareness (All disaggregated by age, gender, and disability) [FGDs and case studies AWDs] Anecdotal evidence and contribution stories concerning changes in decision-making, communication, problem solving, emotional regulation, conflict resolution, empathy, and safety awareness	Desk review Student survey FGDs	Life Skills Training and OOS capacity building monitoring tests (if accessible/available: situational judgement tests - life skills education; group members assessments - Girls Power manual) Adolescents AWDs	Content analysis Descriptive statistical analysis Thematic Framework analysis Contribution analysis

Criterion	Questions	sub-questions	Collection phase	Indicators	Data collection method	Data source	Analytical approach
		To what extent have governmental systems and platforms improved, enhancing the minimum services' availability, accessibility and quality?	Baseline, Midline, and Endline	[Desk review] Documental evidence on # of gender clubs; Community platforms delivering MHPSS; Safe spaces rehabilitated; WASH facilities rehabilitated; MHH emergency kits procured and distributed; Anthropometric Tools procured and distributed; Teachers and health workers receiving training and capacity building; referral cases; Community MHPSS facilitators receiving training; Key MHPSS messages disseminated at the community level; [Survey] Adolescent students seeking health or AYF services; type of support sought; % of students seeking support reporting facing a barrier; barrier type; % of students reporting that the support addressed their need; % of students reporting being treated friendly and respectfully by service provider; (All disaggregated by age, gender, and disability) % of surveyed teachers receiving training; Knowledge, attitudes and practices concerning gender-responsive and inclusive pedagogy; health and nutrition integration; MHPSS; child protection and referral; WASH promotion; life skills promotion; (All disaggregated by gender, age, and training type) [KIIs, FGDs and case studies AWDs] Anecdotal evidence and contribution stories on service availability, needs, quality, barriers; training;	Desk reviewStudent and teacher surveysKIIs FGDs Direct observation	Project progress and donor reports Data platforms / administrative data Results framework and tracker Adolescents AWDs Head teachers and teachers Frontline workers (social and health) Safe space coordinator Field data collectors' direct observation	Content analysis Descriptive statistical analysis Thematic Framework analysis Contribution analysis

Criterion	Questions	sub-questions	Collection phase	Indicators	Data collection method	Data source	Analytical approach
				[Direct observation] checklist results on WASH, MHM, Disability inclusion in schools			
		To what extent the demand for minimum services package has increased among in-school and OOS adolescents?	Endline	[Desk review] U-report data, service uptake data; [KIIs, FGDs and case studies AWDs] Anecdotal evidence and contribution stories on service provision and change in service demand;	Desk review KIIs FGDs	Project progress and donor reports Data platforms / administrative data Results framework and tracker U-report data (if accessible/available) Adolescents AWDs Teachers Frontline workers (social and health) Safe space coordinator	Content analysis Descriptive statistical analysis Thematic Framework analysis Contribution analysis
		What differences can be observed in results achievement stemming from the implementation of the	Midline and Endline	[Desk review and survey] Change in outcome indicators for treatment group (beneficiaries of packages with additional components) beyond the change observed in the control group (beneficiaries of basic package)	Desk review Student survey	Project progress and donor reports Data platforms / administrative data Results framework and tracker Adolescents	Content analysis Descriptive statistical analysis Difference-in-Differences (DiD) analysis

Criterion	Questions	sub-questions	Collection phase	Indicators	Data collection method	Data source	Analytical approach
		different programme packages across the woredas?					
	Was the programme implemented as planned? How? Why/ Why not? What have been the main barriers, enablers and adaptations to the interventions ?		Midline and Endline	[Desk review] Documental evidence on implementation adherence, dosage, quality, and adaptation; programme adaptations across sites; [KIIs and FGDs] Anecdotal evidence on implementation challenges and enablers; other factors affecting quality implementation;	Desk review KIIs FGDs	Workplans (UNICEF, Government, and implementing partners) Results framework Project progress reports Adolescents Local government officials UNICEF field personnel Head teachers and teachers Frontline workers (social and health) Safe space coordinator	Content analysis Implementation fidelity Thematic Framework analysis
	Were there any unexpected consequences of the programme?		Midline and Endline	[KIIs and FGDs] Anecdotal evidence on positive and negative unexpected consequences	KIIs FGDs	Adolescents Community leaders Parents UNICEF field personnel Head teachers and teachers Frontline workers (social and health) Safe space coordinator	Thematic Framework analysis

Criterion	Questions	sub-questions	Collection phase	Indicators	Data collection method	Data source	Analytical approach
Efficiency/Cost-Effectiveness	To what extent is the programming approach efficient in the achievement of desired results in terms of resource utilization (human, technical, financial) and timely delivery?		Midline and Endline	[Desk review] Documental evidence on ratio of platforms for service delivery/target adolescent population; ratio of trained service deliverers (teachers, health workers, social workers)/target adolescent populations; Ratio of activities completed on schedule; Ratio total cost/adolescent reached ;comparative cost per adolescent in targeted woredas [KIIs and FGDs] Anecdotal evidence on (human, technical, financial) resource utilisation and delays	Desk Review KIIs FGDs	Programme documents including cost and monitoring data (Programme activity costs, Personnel costs, Operations costs, Transportation costs, Monitoring costs, Overheads) UNICEF personnel involved in budget, planning and administration of programme activities Government stakeholders from relevant ministries	Content analysis Thematic framework analysis
	What are the costs of implementing the different interventions in the package? Are these activities low in cost and affordable?		Midline and endline	[Desk review] Documental evidence on Programme activity costs, Personnel costs, Operations costs, Transportation costs, Monitoring costs, Overheads Cost of similar interventions in comparable settings [KIIs and FGDs] Stakeholders' perceptions on affordability	Desk Review	UNICEF monitoring framework UNICEF expenditure and investment tracking Local government and school budgetary data UNICEF personnel involved in budget, planning and administration of programme activities Government officials (federal and sub-federal)	Content analysis Thematic Framework analysis Cost measurement
	Is the programme being delivered in a cost-effective manner? Are there		Midline and endline	[Desk review] Cost per outcome Differences in cost per outcome across woredas Cost effectiveness Ratio	Desk Review	UNICEF monitoring framework UNICEF expenditure and investment tracking Local government and school budgetary data	Cost-effectiveness analysis

Criterion	Questions	sub-questions	Collection phase	Indicators	Data collection method	Data source	Analytical approach
	disparities in cost-effectiveness across areas?						
	OR how cost-efficient is the programme and what potential is there for efficiency savings at all stages (capacity and constraints)?		Midline and endline	Cost-benefit ratio	KIIs Desk Review	UNICEF monitoring framework UNICEF expenditure and investment tracking Local government and school budgetary data	Cost-efficiency/benefit analysis
Impact	To what extent has the program contributed towards impact level change? How and why? At what level, supporting which programme indicators?	To what extent has the programme achieved medium and long-term outcomes that can be measured within the project temporal duration?	Endline	[Desk review] Documental evidence on change in school dropout rates; life skills Training and OOS capacity building scoring (if available) [Medium term]; improved academic performance; increased re-enrolment in schools [Long Term]; [Survey] Knowledge, attitudes and practices concerning MHM, Hygiene, mental health, GBV, Child Marriage, Teenage pregnancy, and nutrition; scoring on scenario-based assessments on decision-making, communication, problem solving, emotional regulation, conflict resolution, empathy, and safety awareness (All disaggregated by age, gender, and disability) [KIIs, FGDs, and case studies AWDs] Anecdotal evidence and contribution stories on changes in knowledge,	Desk review Student survey KIIs FGDs	Project progress and donor reports Data platforms / administrative data Results framework and tracker Life Skills Training and OOS capacity building monitoring tests (if accessible/available: situational judgement tests - life skills education; group members assessments - Girls Power manual) Adolescents AWDs Community leaders Parents	Content analysis Descriptive statistical analysis Thematic framework analysis Before and after comparison Contribution analysis DiD Estimations

Criterion	Questions	sub-questions	Collection phase	Indicators	Data collection method	Data source	Analytical approach
				attitudes, and practices on gender roles, nutrition, Child marriage, teenage pregnancy, mental health, GBV, SRH, STIs, MHM, support from caregivers and community, and reduction in harmful social and gender norms			
	What is the difference in impact indicator achievement in the different regions and woredas?		Midline and Endline	[Desk review and survey] Change in medium- and long-term outcome indicators (see below) for treatment group (beneficiaries of packages with additional components) beyond the change observed in the control group (beneficiaries of basic package) [Desk review] Documental evidence on change in school dropout rates; life skills Training and OOS capacity building scoring (if available) [Medium term]; improved academic performance; increased re-enrolment in schools [Long Term]; [Survey] Knowledge, attitudes and practices concerning MHM, Hygiene, mental health, GBV, Child Marriage, Teenage pregnancy, and nutrition; scoring on scenario-based assessments on decision-making, communication,	Desk review Student survey	Project progress and donor reports Data platforms / administrative data Results framework and tracker Life Skills Training and OOS capacity building monitoring tests (if accessible/available: situational judgement tests - life skills education; group members assessments - Girls Power manual) Adolescents	Before and after comparison Contribution analysis DiD Estimations

Criterion	Questions	sub-questions	Collection phase	Indicators	Data collection method	Data source	Analytical approach
				problem solving, emotional regulation, conflict resolution, empathy, and safety awareness (All disaggregated by age, gender, and disability)			
	What is the added value of the packages containing additional interventions compared to the basic one?		Midline and Endline	[Desk review and survey] Change in medium- and long-term outcome indicators (see below) for treatment group (beneficiaries of packages with additional components) beyond the change observed in the control group (beneficiaries of basic package) [Desk review] Documental evidence on change in school dropout rates; life skills Training and OOS capacity building scoring (if available) [Medium term]; improved academic performance; increased re-enrolment in schools [Long Term]; [Survey] Knowledge, attitudes and practices concerning MHM, Hygiene, mental health, GBV, Child Marriage, Teenage pregnancy, and nutrition; scoring on scenario-based assessments on decision-making, communication, problem solving, emotional regulation, conflict resolution, empathy, and safety awareness (All disaggregated by age, gender, and disability)	Desk review Student survey	Project progress and donor reports Data platforms / administrative data Results framework and tracker Life Skills Training and OOS capacity building monitoring tests (if accessible/available: situational judgement tests - life skills education; group members assessments - Girls Power manual) Adolescents	Before and after comparison Contribution analysis DiD Estimations

Criterion	Questions	sub-questions	Collection phase	Indicators	Data collection method	Data source	Analytical approach
Sustainability	What key factors will facilitate the scale-up and sustainability of the package?	To what extent have the strategies adopted by the programme contributed to the sustainability of results, especially equity and gender-related results?	Baseline, Midline and Endline	[Desk review and survey] participant sub-groups - in-school girls (including those with disabilities); out-of-school girls (including those with disabilities); in-school boys (including those with disabilities) - with smaller gains in short-, medium, and long-term outcome indicators [case studies AWDs] evidence and case content	Desk review Student survey	Project progress and donor reports Data platforms / administrative data Results framework and tracker Life Skills Training and OOS capacity building monitoring tests (if accessible/available: situational judgement tests - life skills education; group members assessments - Girls Power manual) Adolescents AWDs	Content analysis Descriptive statistical analysis
		To what extent is the joint programme supporting long-term buy-in and ownership by duty bearers and rights holders?	Midline and Endline	[Desk review, KIIs and FGDs] Anecdotal, documentary evidence and contribution stories on: continuation of key skills practice acquired as a result of the programme; peers, parents or community influencing (programme learnings); programme learnings/data used to inform policies or programmes; scale-up strategies; emergence of spin-off initiatives by youth or community; instances of public advocacy by local leaders; linkages between programme monitoring and UNISE; integration of HPV in government routine immunization programme; scale-up strategies; youth participation in decision-making and monitoring (Youth Advisory Groups); business plan and implementation of activities to replenish MHH emergency kits in schools; formal commitments (local government or community); persistence of political/cultural barriers to programme	Desk review KIIs FGDs	Project progress and donor reports Data platforms / administrative data Results framework and tracker Government officials (federal and sub-federal) UNICEF Personnel Adolescents	Content analysis Thematic Framework analysis Contribution analysis

Criterion	Questions	sub-questions	Collection phase	Indicators	Data collection method	Data source	Analytical approach
				continuation; unmet needs; persistence of key risk factors impeding sustained outcomes;			
	What enablers and barriers are there for the GoE and other national and sub-national stakeholders to sustainably deliver results for adolescents during and beyond the project timeframe?		Midline and Endline	[KIIs and FGDs] Anecdotal evidence on enablers and barriers for sustainable programme delivery	KIIs FGDs	Government officials (federal and sub-federal) UNICEF Personnel Head teachers and teachers Frontline workers (social and health) Safe space coordinator	Thematic framework analysis

Criterion	Questions	sub-questions	Collection phase	Indicators	Data collection method	Data source	Analytical approach
Coherence	How well do the programme interventions fit together, create synergies and coherence?		Midline and Endline	[Desk review, KIIs and FGDs] Anecdotal and documentary evidence on alignment of programme activities, outputs and outcomes to goal/impact; consistency with relevant government policies/strategies; adolescent involvement in programme decision-making, implementation and monitoring (particularly U-report, Federal, Regional, Woreda Youth Advisory Groups); knowledge/data exchange, particularly to inform decision-making;	Desk reviewKIIsFGDs	Programme design documents, progress reports U-report Youth Advisory Groups minutes and other documentation Sectoral government policies and strategies Government officials (federal and sub-federal) UNICEF Personnel	Content analysis Thematic framework analysis
	What is the extent of partnership, coordination, and complement arity with the interventions of the Ethiopian government and other relevant actors operating within the same context around similar interventions ? Are there aspects of the operation		Midline and Endline	[Desk review, KIIs and FGDs] Anecdotal and documentary evidence on partnerships formed; similar interventions implemented by other actors in the target woredas; complementarity/synergies with interventions implemented by other actors in the target woredas; programme alignment with the United Nations Sustainable Development Cooperation Framework and Humanitarian Response Plan	Desk review KIIs FGDs	Project progress reports Ministerial / UN Agencies / NGOs annual reports United Nations Sustainable Development Cooperation Framework and Humanitarian Response Plan Government officials (federal and sub-federal) Frontline workers (social and health) UNICEF Personnel Community leaders	Content analysis Thematic framework analysis

Criterion	Questions	sub-questions	Collection phase	Indicators	Data collection method	Data source	Analytical approach
	that conflict with the interventions of or one-UN programming or other actors?						
	How efficient and effective was the collaboration among various ministries involved in the programme, and whether the programme was led by and involved all relevant departments to achieve the best results?		Midline and Endline	[Desk review, KIIs and FGDs] Anecdotal and documentary evidence on ad-hoc formal coordination platforms or ad hoc working groups; use of national and sub-national level multisectoral platforms to deliver minimum packages at school level; cooperation documents (Memorandums, TORs, etc.); annual Intersectoral policy dialogue on integration and provision of Minimum Service Package (MOH , MOE , MOWSA); coordination at the regional, city administration, zonal and woreda levels (including reported overlaps or confusion in roles/responsibilities); coordination enablers and challenges	Desk review KIIs FGDs	UNICEF and government documentation Results framework Project progress reports Government officials (federal and sub-federal) Frontline workers (social and health) Teachers UNICEF Personnel	Content analysis Thematic framework analysis
Gender, disability, Equity and Human Rights	To what extent has UNICEF and the government identified and addressed the specific needs of girls		Baseline, Midline, and Endline	[Desk review] Evidence of contextual analyses, stakeholder analyses, vulnerability assessments, literature explaining interventions' rationale; [KIIs and FGDs] Anecdotal evidence on need identification in programme design and needs adaptation during programme implementation; contribution stories on need addressal;	Desk review KIIs FGDs Student survey	Programme design documentsGrey literature (programme designers and implementers) Adolescents	Content analysis Thematic Framework analysis Contribution analysis Descriptive statistical analysis

Criterion	Questions	sub-questions	Collection phase	Indicators	Data collection method	Data source	Analytical approach
	and boys in the design and implementation of the programme?			[Survey] % of students reporting that the programme addresses needs previously not supported (Disaggregated by age, gender, and disability)			
	To what extent is programming conducive to supporting children with disabilities and from the most marginalized populations (including those furthest left-behind)?		Midline and Endline	[Desk review] Documental evidence on participant sub-groups - in-school girls (including those with disabilities); out-of-school girls (including those with disabilities); in-school boys (including those with disabilities) - with smaller gains in short-, medium, and long-term outcome indicators [case studies AWDs] evidence and case content [KIIs and FGDs] contribution stories from programme beneficiaries	Desk review KIIs FGDs	Project progress and donor reports Data platforms / administrative data Results framework and tracker Life Skills Training and OOS capacity building monitoring tests (if accessible/available: situational judgement tests - life skills education; group members assessments - Girls Power manual) Parents and caregivers Adolescents AWDs	Content analysis Descriptive statistical analysis Thematic Framework analysis Contribution analysis
	To what extent did programming incorporate a Human Rights Based Approach in all its phases? Were duty-bearers and rightsholders involved in advocacy,		Baseline, Midline, and Endline	[Desk review] Documental evidence on HRBA alignment in design, implementation, monitoring and evaluation; Youth participation in decision-making and monitoring (Youth Advisory Groups); beneficiary adolescents included as U-reporters; [KIIs and FGDs] Anecdotal evidence on beneficiary involvement	Desk review KIIs FGDs	UNICEF and government documentation U-report data (if accessible/available) Results framework Project progress reports UNICEF Personnel	Content analysis Thematic Framework analysis

Criterion	Questions	sub-questions	Collection phase	Indicators	Data collection method	Data source	Analytical approach
	design and implementation of the programme?						

8 Methodology

The evaluation will use various analytical methods to answer the evaluations questions. Primarily, to measure the programme's impact on outcomes of interest, the evaluation team suggests using a Differences-in-Differences approach based on a survey of primary school students. The approach will not have a pure comparison group but rather all groups (in this case schools), receive the "basic" set of interventions – while some benefit from additional interventions. The rationale for using this approach is highlighted in evaluability assessment provided in the annexe. The evaluation will compare the cost, effectiveness and impacts of different implementation modalities against the base scenario. Quantitative data will be collected before the intervention (baseline), midline and at endline from the same target population, and changes will be analysed descriptively. Qualitative methods will support the quantitative analysis and the process evaluation- this will be critical in better understanding which factors contributed to results. These will run alongside the surveys to interpret the *how and why* of those changes.

Ethical Imperatives

The evaluation team fully aligns with UNICEF's commitment to "do no harm" and affirms that the wellbeing of children supersedes all operational concerns. Given the risks associated with administering lengthy or cognitively demanding survey tools to young children—especially in fragile contexts—direct engagement with children will only be undertaken after completing the necessary ethical reviews, obtaining guardian consent, child assent and providing enumerators with robust, context-appropriate training. This includes methods for child-sensitive data collection, recognition of distress signals, and adherence to child safeguarding protocols.

8.1 Data Collection Methods

In light of ethical imperatives, technical considerations, and programmatic data needs, the evaluation team will adopt a mixed-methods approach to the baseline, midline and endline data collection for the Minimum Package Evaluation. This approach is designed to uphold UNICEF's safeguarding commitments—particularly in relation to engaging with children—while ensuring timely, useful, and ethically sound baseline evidence generation to inform the programme's implementation and learning agenda.

8.1.1 Baseline, Midline and Endline Surveys

A quantitative survey will be administered in selected intervention schools at baseline (just as activities commence) and repeated at midline (after 12 months) and endline (after 24 months). These repeated cross-section surveys will include a student survey (primary schools) of girls and boys aged 10-14 and teacher surveys in primary schools to measure selected health and nutrition, MHH, life skills and mental well-being outcomes, and

knowledge/behaviour questions on WASH and protection (e.g. handwashing practices, safety perception in school). The sample will cover a representative subset of each package modality and will allow to test for differences in outcomes between these packages. The teacher survey will consist of a representative survey of teachers across the intervention schools. The

Sampling strategy for primary student survey

To design a multi-arm experiment comparing schools receiving the Basic package against other program packages, we first need to calculate how many schools are required to detect a change in outcomes. Assuming a 50% prevalence in the Group 1 and 62.5% prevalence in Group 2 (a 12.5 percentage point difference) would lead to Cohen's h coefficient of 0.25. This effect size for this analysis (h = 0.25) was found to exceed Cohen's convention for a small effect (h =0.2) but would allow to measure small to medium effects from the programme.

Assuming that 50% of the subjects in the reference population have the factor of interest, a 5% level of significance and after applying continuity correction, the study would require a sample size of 244 for each group to detect a difference of 12.5 percentage points.

However, given that the data collection will be taking place in clusters (in this case schools), the sampling calculations should account for the fact that within these clusters, observations are more similar to each other than observations from different clusters, which means that while clusters are independent, observations within clusters are not — thus, violating the assumption of independence in standard sample size calculations. The sample should be multiplied by a design effect to account for this loss of power. Using the design effect formula below, and assuming an intra-cluster correlation of 0.08, we estimate a design effect of 4.12:

DE= 1 + (M-1)p , where M is the size per cluster set at 40 and p is the ICC

Given that most variations of the package are implemented in a small number of schools, we only apply this design effect to the arm consisting of schools receiving the basic package (which will serve as a comparison). For the rest, since all schools are captured through the data collection, the design effect is reduced to 1 and Simple Random Sampling is used in effect. The sample would be as follows for each phase (baseline, midline, endline):

Package	Number of schools receiving	Number of schools where data is collected	Number of students surveyed per phase	Number of teachers surveyed per phase
Full	19	19	245	50
Basic	41	27	1,080	57
Basic + CP + MHPSS	6	6	245	5
Basic + CP	22	22	245	21
Basic + CP + WASH	10	10	245	95
TOTAL (with 5% oversampling)	100	84	2,170	245

The sample in each school will consist of a mix of girls and boys aged 10-15. The evaluation team suggests allocating a higher share of the sample (60-70%) to girls given that a larger part of the programme components are specifically targeted at girls. Additionally, during the recruitment process, the enumerators will work with head teachers at identifying and including adolescents with disabilities in the research. The survey questionnaire includes a series of questions to capture adolescents' disabilities and their responses will be highlighted in the reporting to reflect their outcomes.

8.1.2 Qualitative Methods

A robust qualitative component will delve deeper into why changes did or did not occur and how the integrated approach functioned. **Qualitative methods – including case studies –** will run alongside the surveys to interpret *how* and *why* those observed changes took place. This mixed-method design balances measurement of what changed with in-depth inquiry into the change processes and drivers, which is crucial for evaluating an integrated programme. By triangulating survey results with qualitative evidence, the evaluation will build a nuanced understanding of the programme's effects and the extent to which observed changes can be credibly linked to the interventions of the minimum package.

This will include **FGDs** with various participant groups: students (separately by gender/age), OOS girls, parents/caregivers, gender club members, frontline workers and community members. FGDs will explore perceptions of the programme's effects changes in attitudes any persisting challenges. The FGDs will aim to capture these various stakeholders' views on needs and challenges at baseline and will delve into programmatic changes and impacts at midline and endline. FGDs with students (both in and out of gender clubs) will focus on their acquired knowledge and skills, as well as recommendations to improve package interventions. FGDs with caregivers and community members will help explain contextual factors and enablers or bottlenecks towards improvement in students' outcomes. Critically, FGDs will serve as the main approach to collect data on OOS girls and their experience with the programme.

As part of each FGD, a Most Significant Change (MSC) storytelling exercise will be incorporated: participants will be invited to recount *"Since this programme started, what is the most significant change you have experienced (or observed) in your child's schooling or community, and why is it important to you?"* Through this exercise, beneficiaries themselves identify the changes they value most since the programme began and discuss what or who contributed to those changes. This MSC technique is especially valuable in an integrated programme context – participants' stories can reveal how multiple interventions (education, WASH, protection/MHPSS) may have collectively produced a meaningful change. The team will document these MSC stories and later analyse them to understand common themes and causal pathways, which will enrich the contribution analysis by illustrating *how and why* key changes happened from the beneficiaries' perspective. FGDs will be held separately for teachers, students (segregated by gender), and parents/caregivers, using participatory tools such as ranking exercises, timelines, and perception mapping. These discussions will explore how services were experienced, perceived quality and accessibility, and any challenges

encountered. FGDs will be facilitated by trained moderators in local languages and will last 60 to 90 minutes.

KIIs will be conducted with stakeholders such as school principals, social workers, UNICEF field staff, local government officials (Education Bureau, Water Bureau, etc.), and national UNICEF and government stakeholders. KIIs will gather insights on needs and deprivations, implementation fidelity, coordination, and context (for example, how did government involvement affect the programme, enablers and bottlenecks, what external events influenced results). The qualitative sample will be purposive to ensure all geographical areas (regions and woredas) are represented. Interviews will be semi-structured following an interview guide (see Annex 2 for a summary of qualitative tools).

KIIs will gather insights on needs and deprivations and programme design at baseline. In the later stages of data collection, KIIs will focus on implementation quality, coordination among stakeholders and referral systems, and contextual factors influencing results. KII respondents will be asked about any significant changes they have observed and their views on what influenced those changes. For the implementation fidelity exercise, semi-structured KIIs will be conducted with UNICEF programme officers, implementing partner staff, local government representatives, school leadership and frontline worker. These interviews will explore timelines, delivery processes, coordination practices, adaptations, and feedback from the field. Each interview will last approximately 30 to 45 minutes.

Direct observations will be conducted at each school site using structured checklists. These observations will assess the status and use of WASH facilities (rehabilitation, accessibility, functionality), the status of gender clubs and safe spaces in the schools. As part of the observations, the teams will also collect information on school attendance and dropout. Field teams will document infrastructure quality, use patterns, and upkeep.

Case studies will be conducted with adolescents with disabilities to gather information on their experiences and challenges. At baseline, the studies will focus on these adolescents' school-related deprivations and will then focus on their access, or lack thereof, to package interventions. The case studies are expected to generate evidence for making the design of the package and its underlying interventions more accessible to vulnerable groups.

8.1.2.1 Sampling Strategy for qualitative methods

The qualitative sample will be purposive and will seek to include perspective from stakeholders across all geographic areas covered in this pilot with a focus on disadvantage and marginalized groups and those with disabilities. The sampling strategy is purposive, meaning participants are selected based on their relevance to specific programme activities and pathways in the Theory of Change, rather than for statistical representativeness. These include geographic context (all implementation woredas), various types of schools based on their package receipt, and differences in student population characteristics. The team will ensure that the voices of adolescents with disabilities are reflected, particularly through detailed case that will include them and their caregivers. Furthermore, as part of the data collection in each locality the team will work with community members and school staff to ensure that adolescents from vulnerable groups (e.g. displaced, with disabilities etc) are included in group

discussions or interviews. The diverse perspective of these groups, especially marginalised ones, will be key to inform the design of the programme and whether implementation needs to be better adapted to answer their needs. In the analysis and reporting of the qualitative data, the team will ensure to highlight the findings from these groups throughout the evaluation questions – especially if the responses highlight disparities between groups and/or between localities.

This stratification will help the evaluation capture contextual differences in implementation. **For each phase (baseline, midline and endline**), the sampling is tentatively expected to cover the following in each woreda:

Table 4: Qualitative data sampling

Method	Target Group	Participants per Group	Sampling Strategy	Total Sample Size
Key Informant Interviews	Principals of primary schools	1	2 per woreda - purposively in schools where survey is taking place	14
	School Safe space coordinators/ counsellors	1	1 per woreda	7
	UNICEF and IP field implementers	6-8 for FGDs	1 per woreda	7
	Staff from ministries, UNICEF programme teams and M&E staff, Development Partners, Organisations of Persons with Disabilities	1	National	10-12
Focus Group Discussions	Boys in primary schools 12-14	6-8 participants	1 per woreda	7
	Girls in Primary schools 10-11	6-8 participants	1 per woreda	7
	Girls in Primary schools 12-14	6-8 participants	1 per woreda	7
	Out of school girls 12-14 in nearby youth centres / community centres	6-8 participants	2 per woreda	14

Method	Target Group	Participants per Group	Sampling Strategy	Total Sample Size
	Parents/caregivers from host communities	6-8 participants	1	7
	Local leaders, community elders, SMC/PTA	6-8 participants	1	7
	Frontline social workers/community facilitators	6-8 participants	1 per woreda	7
	Local Government - Education, WASH, Health, Social Affairs officials	6-8 participants	1 per woreda	7
Case Study - Children with Disabilities	Children with disabilities + caregiver	3	Across 3 different woredas	3

8.1.3 Monitoring Data

In the first six months to one year of implementation, the evaluators will gather monitoring data from the programme implementation team to document what was implemented when and how each month. Following that, this data will be collected in the form of reports and monitoring databases as maintained by UNICEF during midline and endline data collection. This includes reviewing implementation partners' progress reports, UNICEF's and the GoE's output monitoring (e.g. number of teachers trained, infrastructure rehabilitated, vaccines administered etc.). A timeline of interventions (by school) will be compiled to correlate inputs with outcomes. This process data is crucial to explain outcome findings.

8.2 Data Analysis

8.2.1 Impact analysis

The evaluation aims to measure changes in outcomes among adolescents exposed to different school-based program. While the program packages were **pre-assigned to schools** rather than randomly allocated, the design allows for the comparison of changes in outcomes over time between treatment groups and the Basic package group.

A **Difference-in-Differences (DiD)** approach was selected because it is well-suited for settings where randomization is not feasible and where **baseline measurements can be collected before the interventions begin**. The DiD method compares the **change in outcomes from baseline (pre-intervention) to midline/endline** in the treatment groups relative to the change in the comparison group, which allows for the estimation of program impact while controlling for **time-invariant unobserved differences between groups**.

The choice of DiD was informed by the **evaluability assessment (provided in the annex)**, which examined program design, implementation coverage, and the feasibility of alternative evaluation designs. Key findings from this assessment include:

- All schools receiving non-Basic packages will be fully covered by the evaluation, ensuring complete measurement of outcomes in these groups.
- Schools receiving the Basic package serve as the **closest available comparison group**, though they are not a "pure" control group since they may still receive other interventions or be influenced by external factors.
- Traditional randomized or quasi-experimental designs were determined to be infeasible due to **pre-assignment of program packages**, limited comparability of schools not receiving interventions, and ethical considerations of withholding services.

Given these constraints, the DiD approach provides a **robust, feasible alternative** for attributing observed changes to program interventions, while accounting for baseline differences and broader trends affecting all schools.

This evaluation involves a **multi-arm intervention** in schools, with each school pre-assigned to either a Basic package (in this case, the control) or one of several enhanced intervention packages. All schools will be observed at three time points: baseline (pre-intervention), midline (tentatively 1 year after the baseline), and endline (tentatively 2 years after the baseline). Measuring the programme's impact on outcomes of interest requires careful consideration of challenges including the clustered design of the intervention, multiple treatment arms, and ensuring that observed outcome changes can be *attributed* to the interventions rather than confounders. Specifically, the analysis must account for:

- **Clustering (School-Level Assignment):** Students within the same school are more similar to each other, violating independence assumptions. Ignoring this intraclass correlation inflates Type I error.
- **Repeated Measures:** The same schools (and potentially the same students) are measured at multiple time points. This introduces within-subject and within-cluster correlations that must be modelled appropriately.
- **Multiple Arms & Baseline Imbalance:** With several intervention packages, the chance of baseline differences between groups increases as arms increase. Even with random assignment, some schools may start higher or lower on outcomes. We need to adjust for any baseline imbalance to isolate true treatment effects.
- **Outcome Heterogeneity:** Outcomes aligned to the theory of change include short-, medium-, and long-term metrics. Effects may emerge at different speeds (e.g. short-term outcomes improving by midline, long-term by endline) and could vary in magnitude. The method must capture these time-dependent effects and handle different outcome types (continuous scores, binary indicators, etc.) as needed.

A DiD approach compares the change in outcomes from baseline to follow-up in treatment groups against the change in the control group (Basic). This naturally accounts for baseline level differences and isolates the *differential* change due to treatment. In a cluster trial, DiD leverages baseline measurements to improve power over a post-test-only comparison. The

key assumption is that in absence of treatment, the trends would be parallel across groups – reasonable here due to randomization (no systematic bias in trends). DiD is conceptually straightforward and directly supports attribution by focusing on *changes* rather than absolute levels. However, implementing DiD requires a modelling approach that correctly handles clustering and repeated measures.

The core analytical model will be a **linear mixed-effects regression** that includes fixed effects to capture the DiD comparisons and random effects to account for clustering (and repeated observations). Key features of the model:

- **Outcome Variable:** Each student groups' outcome measured at baseline, midline, and endline. (If the same students are followed over time, the data are truly longitudinal; if different students are sampled at each wave, the model is still at the school level with repeated measures per cluster). We assume continuous outcomes for illustration; if binary or count outcomes exist, a generalized mixed model (e.g. logistic) can be used similarly.
- **Time Indicators:** Include time as categorical indicators for midline and endline (with baseline as the reference). This allows flexible outcome trajectories over time. These fixed effects absorb the average change over time in the **Basic** (control) group (i.e. secular trends).
- **Treatment Group Indicators:** Include indicators for each enhanced package group (for each treatment arm), with the Basic package group as the reference category. This accounts for any persistent differences between groups (though under randomization, baseline differences are random and will be handled by baseline covariates).
- **Interaction Terms (DiD effects):** Crucially, include interactions between each treatment group and the time indicators. These interaction coefficients estimate the **additional change in the outcome for that treatment group beyond the change observed in the Basic group**. In other words, the interaction term is the difference-in-differences estimator – how much more (or less) the treatment group's outcome changed from baseline to follow-up compared to the control. Each treatment arm will have its own interaction for midline and endline, allowing the treatment effect to potentially differ at 1 year vs 2 years. This aligns with the theory of change: we can detect short-term impacts at midline and longer-term impacts at endline for each package.

A critical requirement to obtaining unbiased estimates of changes in outcomes is the ability to control for cofounders – to do so, the evaluation team will do the following:

- **Observed confounders** such as age, gender, regional and woreda indicators, school size, and baseline outcome levels will be included as covariates in the DiD regression models.
- **Unobserved time-invariant confounders** are addressed by the DiD framework itself, which subtracts out effects common to both treatment and comparison groups over time.
- **Time-varying external factors** (e.g., government policy changes, seasonal shocks, local health campaigns) will be documented during the evaluation through qualitative monitoring and incorporated as control variables where feasible.

- **Process documentation** and qualitative data collection will complement quantitative findings, allowing the evaluation to assess the contribution of program outputs to observed outcomes in the context of external influences.

The team will explore additional quasi-experimental analysis options such as propensity score matching to attribute impact from the programme on key outcomes of interest. The methodology for propensity score matching approaches, which can be used together with DiD, is displayed in the annex.

8.2.2 Effectiveness Analysis

8.2.2.1 Contribution Analysis Framework

All evidence will be synthesized using a contribution analysis approach. We will begin by articulating the programme's results chain and assumptions (from the ToC) and then systematically examine the extent to which each link in the chain occurred, using both quantitative indicators and qualitative findings. For outcomes like "improved knowledge of nutrition and health practices" or "improved mental well-being," we will look for the presence of expected intermediate changes (e.g. new WASH facilities built, psychosocial activities delivered) and gather stakeholder confirmation that these were important. We will also investigate alternative explanations for observed changes and assess their influence relative to the programme. The end result will be a reasoned judgment of the programme's contribution to each outcome area, with evidence tracing the narrative from activities to outcomes. The following **non-exhaustive** table illustrates, based on the articulated Theory of Change, how outputs will logically be linked to outcomes as part of the contribution analysis – it is important to note that given that this programme relies on a multi-sectoral integrate approach, multiple outputs together are expected to contribute to the same outcomes:

Table 5: Contribution analysis linkages

Programme Output	Short-Term Outcome	Medium-Term Outcome	Long-Term Outcome	Measurement Instrument	Timing
Life-skills modules delivered to students	Increased knowledge of decision-making, problem-solving, negotiation skills	Improved self-efficacy, critical thinking, and conflict resolution skills	Enhanced life skills reflected in improved academic engagement and reduced risky behaviors	Life Skills Survey (enumerator-administered)	Baseline, Midline, Endline
WASH facilities installed (handwashi	Increased knowledge of proper handwashin	Increased use of school WASH	Sustained hygiene practices, reduced	WASH Observation Checklist, KAP Survey (Hygiene module)	Baseline, Midline, Endline

Programme Output	Short-Term Outcome	Medium-Term Outcome	Long-Term Outcome	Measurement Instrument	Timing
ng stations, clean toilets)	g, hygiene practices	facilities, adoption of hygiene behaviors	absenteeism due to illness		
SRH education delivered (sexual and reproductive health)	Increased knowledge of puberty, contraception, and reproductive health	Improved decision-making around SRH, adoption of protective behaviors	Reduced early pregnancy, improved SRH outcomes	SRH KAP Survey, FGDs	Baseline, Midline, Endline
Teacher training on life skills, nutrition, and WASH	Increased teacher knowledge and confidence	Improved delivery of modules and facilitation skills	Sustained program quality, improved student outcomes	Teacher Knowledge & Practices Survey, Classroom Observations	Baseline, Midline, Endline
Program monitoring and feedback mechanisms	Awareness of program objectives among school staff	Adaptive improvements in program delivery	Sustained quality and coverage of interventions	Implementation/Process Monitoring Reports	Continuous

Quantitative data will be analysed using **before-after comparisons** – primarily paired statistical tests or difference in proportions over time within the sample. We will use disaggregation to see if outcomes differed by sub-groups. While causality cannot be confirmed for all outcomes, a significant improvement across most sites in girls' Knowledge, Attitudes and Practices (KAPs) and attendances would be taken as evidence supporting the programme's contribution, especially if triangulated with qualitative feedback like "the new latrines made it easier for girls to attend regularly."

Qualitative data from FGDs/KIIs will be transcribed and coded to identify common themes and outliers. We will use qualitative techniques, **including participatory ranking and the MSC** storytelling approach. These methods will gauge whether beneficiaries and local officials attribute improvements to the programme and identify which changes they perceive as most significant. In particular, integrating the MSC technique allows beneficiaries to define the most important outcomes and explain how and why those changes occurred, providing rich, beneficiary-defined evidence of the programme's contribution within the broader integrated approach. Any contradictions (e.g., if surveys show improvement but some participants say nothing changed) will also be explored and explained. By using mixed methods, the evaluation

aims to build a compelling evidence base for the programme's results narrative, even without experimental rigor.

8.2.3 Implementation Fidelity Analysis

The analytical framework for assessing implementation fidelity draws on established implementation science principles and consists of four core dimensions: adherence, dosage, quality, and adaptation. Adherence refers to the extent to which interventions are delivered as designed, while dosage assesses the frequency, duration, and coverage of activities relative to what was planned. Quality examines the skill, consistency, and contextual appropriateness with which services are delivered. Adaptation considers whether deviations from the original design occurred, and if so, whether they enhanced or undermined the intended results.

In the context of the minimum package, the implementation fidelity framework is particularly relevant given the integrated and multisectoral nature of the interventions. The dimension of adherence will be assessed by comparing what was delivered on the ground with what was planned in the design documents—for example, teacher training sessions under the education component, delivery of MHH supplies, SBC edutainment interventions, delivery of life skills curriculum, rehabilitation of usability of WASH facilities, and provision of community-based MHPSS and child protection services. The dosage dimension will focus on the intensity and reach of each intervention. This includes whether target groups—especially girls—received the intended number and frequency of services such as sessions, materials, or visits. For quality, the evaluation will examine whether services were delivered with consistency, appropriateness, and skill, drawing on direct observations, user feedback, and interviews with teachers, students, and caregivers. Finally, the adaptation component will capture any changes made during implementation, exploring whether deviations from the original design were necessary responses to contextual barriers, such as access, staffing limitations, or infrastructure delays, and whether these changes enhanced the programme's relevance or undermined fidelity to its core design. This framework application will allow the evaluation to capture a nuanced picture of how well the programme is functioning in practice and where mid-course corrections may be required. This external study will serve as a critical early review to strengthen implementation quality, course-correct weaknesses, and reinforce the integrated delivery model for improved outcomes by programme end.

8.2.4 Framework analysis

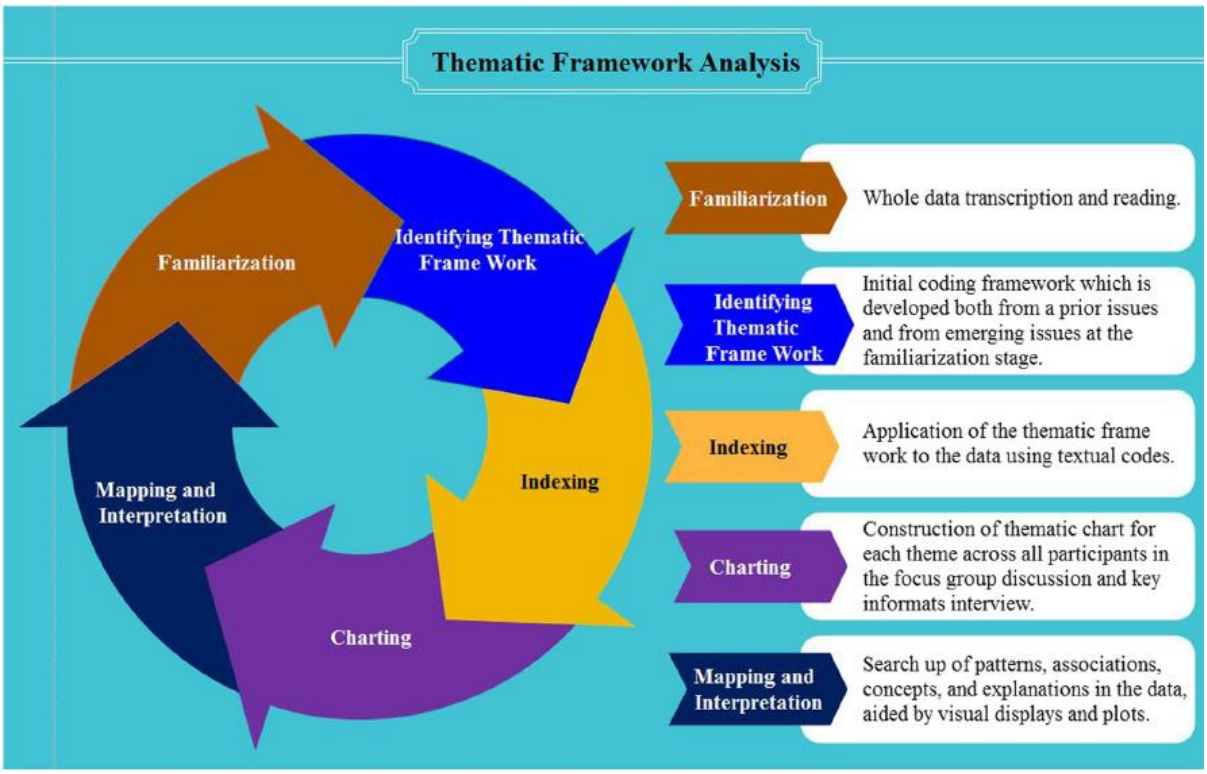
A framework analysis approach helps evaluators structure large volumes of data and interpret them in a more structured manner to answer questions mainly relating to programme **relevance, and sustainability** in evaluative research where thematic areas or a framework can be developed. In applied policies or programme evaluations, this has a unique space as it allows the study team to either collect all the data and then analyse it or do data analysis during the collection process and tweak the instruments and framework further.

Framework analysis is heavily based in, and driven by, the original accounts and observations of the people it is about; is flexible and open to amendment throughout the analytical process; is systematic and comprehensive and allows within-case and between-case analysis. Finally,

and importantly, it is accessible to others: the analytical process and interpretations derived from it can be viewed and judged by people other than the evaluators, making the process more transparent and useful for the programme team.

Framework analysis will involve five stages of work (as shown in the figure below):

Figure 3 Thematic Framework Analysis



Source: Dagnew, Z. A., Abraham, I. A., Beraki, G. G., Mittler, S., Achila, O. O., & Tesfamariam, E. H. (2020). Do nurses have barriers to quality oral care practice at a generalized hospital care in Asmara, Eritrea? A cross-sectional study. BMC oral health, 20, 1-11.

The evaluation team will use NVivo to support analysis at several stages by enabling users to store, organise, categorise, and classify qualitative data and then support structured analysis of the information during familiarisation, identifying a thematic framework, indexing, charting, and mapping and interpretation.

During **familiarisation,** EPRI will use a selection of the gathered information to gain awareness, identify themes and recurrent ideas once the data is collected. This will involve gaining an overview of the data, by reading initial transcripts or field notes from the KIIs, and FGDs.

Audio recordings of interviews and FGDs will be transcribed and translated to English (if needed) for the international team to analyse with support from the national counterparts. The raw datasets in the form of interview transcripts will be reviewed for initial impressions, cleaned, and anonymised. Identifiable information (such as names, places, districts) will be removed to preserve the privacy and anonymity of participants. Where needed, acronyms/abbreviations used during interviewing and transcribing will be expanded,

unusable data discarded, detectable errors corrected (without changing the meaning of the data, as far as possible), duplicates removed, etc.

The raw data will be consolidated into one dataset in preparation for analysis. At this stage, any questions arising from the data cleaning process may be clarified with the fieldwork team.

The evaluation team will **identify the thematic framework.** The EPRI team will build a thematic framework based on the data and areas/ideas and the cleaned dataset will be organised by topic categories/evaluation questions. A priori issues will also be built into the framework to the extent that the data indicate they are necessary.

This framework will then be used to filter and classify the rest of the data, and in this process, using intuitive and logical thinking, the team may further update the framework to reflect new and emerging themes and

During **indexing,** the full volume of collected data will be indexed in detail across the themes in the framework using NVivo (for qualitative coding/indexing). NVivo supports qualitative research by enabling users to store, organise, categorise, and classify qualitative data and then support structured analysis of the information. With the NVivo tool, the researcher can categorise single cases and observations with specific attributes and then analyse results in a structured manner.

During **charting**, initial inductive coding of data to draw out emerging themes and sorted data into themes and analysed to identify the frequency of themes, associations between themes, common experiences and perceptions. This component allows data to be sorted according to the themes identified in the thematic framework or a priori research themes.

Finally, in **mapping and interpretation**, the data will be interpreted, and the team will create typologies, find associations, provide explanations, and can develop strategies for improving interventions. The agreed-upon analytical framework will provide structure and focus to the analysis.

8.2.5 Cost-effectiveness Analysis (CEA)

The first step in conducting a CEA is to clearly define the objectives of the analysis:

- What is the specific outcome that needs to be evaluated (e.g., improved health outcomes, reduced poverty, increased educational attainment)?
- What are the different interventions or program packages being compared (e.g., basic + CP, basic + WASH)?
- What is the time horizon for the analysis (e.g., one year, five years)?

The evaluation team will work with UNICEF and key stakeholders to continuously update the methodology for measuring the programme's cost-effectiveness. This analysis is expected to be presented as part of the endline report but will require collecting outcome measurements throughout the pilot as well as cost estimates on the programme's implementation. The

methodological approach will therefore ensure that these are reflected and that the necessary assumptions and caveats are put in place before conducting the analysis.

Identify the Interventions and Outcomes

In this step, we identify the interventions or program packages that will be compared (e.g., the Basic package vs. enhanced program packages such as Full, Basic + WASH, Basic + CP, etc.).

The desired outcomes should align with the intervention objectives. For instance, the outcomes could be:

- **Health outcomes** (e.g., reduced malnutrition, improved health outcomes).
- **Economic outcomes** (e.g., poverty reduction, productivity, earnings).
- **Educational outcomes** (e.g., improvement in school attendance).

Measure Costs

The costs of the intervention are a critical part of the analysis. There are two types of costs to consider:

- **Direct Costs**: These include the **implementation costs** of the intervention, such as:
  - Program delivery costs (e.g., IFA, deworming tables, operational expenses for life skills delivery, gender club activities, etc.).
  - Administration costs (e.g., monitoring, evaluation, and reporting).
  - Capital costs (e.g., infrastructure, WASH rehabilitation etc.).
- **Indirect Costs**: These could include any **indirect costs** associated with the program, such as the opportunity cost of participants' time (e.g., time spent attending school or healthcare).

The total cost of each intervention can be calculated as:

Total Cost = (Number of Beneficiaries) × (Cost per Beneficiary)

Where:

- **Number of Beneficiaries** is the target population.
- **Cost per Beneficiary** is the total amount allocated per beneficiary (including direct and indirect costs).

Measure Outcomes

For each intervention, it is essential to define the outcome metrics that will be used to compare effectiveness. The outcome should be a measurable quantity that reflects the impact of the intervention. **The evaluation will seek to measure the change in outcomes directly as part of the data collection, but will rely on a literature review and evidence from similar interventions for outcomes that cannot be measured or directly attributed to the programme.**

Examples of **effectiveness measures** include:

- **Health Outcomes**: Number of children receiving immunizations, reduction in malnutrition.
- **Educational Outcomes**: Improvement in life skills, school attendance.
- **Economic Outcomes**: Poverty reduction, income increase, employment rates.

Effectiveness can also be measured in **quality-adjusted life years (QALYs)** or **disability-adjusted life years (DALYs)**, if deemed relevant to the programme objectives.

Calculate Cost-Effectiveness

Once the costs and outcomes have been defined, the **Cost-Effectiveness Ratio (CER)** is calculated for each program or intervention. This ratio compares the cost of achieving a specific outcome across different interventions.

Cost-Effectiveness Ratio (CER) = Outcome Measure/Total Cost of the Intervention

Compare the Cost-Effectiveness of Interventions

Once the CER is calculated for each intervention, we can compare the interventions. The intervention with the lowest cost per unit of outcome is considered the most cost-effective.

- **Incremental Cost-Effectiveness Ratio (ICER)**: The ICER compares the additional cost and additional benefit of one intervention over another:

ICER = (Effectiveness of Intervention A−Effectiveness of Intervention B)/ (Cost of Intervention A−Cost of Intervention B)

This is useful when comparing two or more interventions that might have similar outcomes but different costs.

Sensitivity Analysis

Since assumptions about costs, effectiveness, and other factors might be uncertain, a sensitivity analysis will be performed to assess the robustness of the results under different assumptions.

- **One-Way Sensitivity Analysis**: Change one variable at a time (e.g., cost, effectiveness) to see how sensitive the results are to that variable.
- **Probabilistic Sensitivity Analysis**: Use **simulations** to test the combined effect of variability in multiple factors simultaneously. This will allow for estimating the **range of possible outcomes** and **confidence intervals** for cost-effectiveness estimates.

8.2.6 Integration of Gender, Equity, and Empowerment of Women (GEEW) and Stakeholder Perspectives

The evaluation has been deliberately designed to integrate Gender, Equity, and Empowerment of Women (GEEW) considerations at multiple levels, ensuring that the analysis moves beyond simple sex-disaggregation. GEEW is reflected explicitly in the evaluation objectives and a criteria is specifically dedicated to reflecting Gender and Equity voices in the evaluation. All tools, both quantitative and qualitative, are structured to capture the gendered experiences of adolescents in schools, with a focus on nutrition, menstrual health, sexual and reproductive health (SRH), life skills development, and the use of WASH facilities. The survey instruments include questions that measure differential access to resources, knowledge, and services by gender, age, and socio-economic status. Indicators such as decision-making autonomy, menstrual hygiene management practices, awareness and uptake of SRH information, and use of school WASH facilities allow for the assessment of both outcomes and empowerment-related changes. Qualitative data collected through focus group discussions and key informant interviews will **specifically include participants from marginalized groups**, such as adolescents with disabilities, girls from low-income households, or those living in remote or underserved communities. Thematic analysis will be employed to identify barriers, facilitators, and context-specific experiences related to intervention uptake and outcomes.

The qualitative data collection, comprising focus group discussions (FGDs) and key informant interviews (KIIs), is similarly designed to capture multiple perspectives. This includes adolescent girls and boys, teachers, school administrators, caregivers, and local health or social service providers. Through these multiple lenses, the evaluation will systematically explore how the program interacts with existing power dynamics within schools, households, and communities. Divergent views and experiences will be documented and analyzed, allowing for a nuanced understanding of equity and access issues. Triangulation across these diverse sources ensures that both convergent and divergent perspectives inform the interpretation of findings. Beyond basic disaggregation, a structured gender analysis framework will be applied to link observed outcomes to underlying gender norms and structural inequities. This framework will allow the evaluation to:

- Compare outcomes across intersecting dimensions such as sex, age, socio-economic background, school type, and disability;
- Examine whether program outputs, such as life skills sessions, WASH improvements, and nutrition interventions, contribute to shifts in adolescent autonomy, decision-making, or agency;
- Identify differential constraints that affect access, participation, and benefit from program activities, particularly for girls in contexts with restrictive gender norms.

Quantitative and qualitative data will be analyzed together to provide a comprehensive understanding of GEEW-related outcomes. Survey scores, knowledge, and reported behaviors will be triangulated with qualitative insights on perceived empowerment, barriers, and facilitators.

8.2.7 Triangulation of Findings

In addition to the data analysis methods/approaches described, one of the major elements that will be utilised is triangulation— the use of data from various sources to capture different perspectives of the same phenomenon. The collection of data through multiple sources will help feed into the triangulation process, whereby acquiring similar information from various sources will allow verification of the same.

- **Method triangulation***:* The results from different methods (Surveys, Desk reviews, KIIs, FGDs) will be triangulated to validate and strengthen answers to causal questions, for example, by identifying and ruling out alternative explanations or by identifying and explaining exceptions to the main pattern observed.
- **Source triangulation**: This will involve comparing and consolidating information received from sources. For example, consolidating information from regional and subnational government agencies and implementing partners.
- **Geographic triangulation***:* This will include comparing and contrasting the implementation mechanisms and practical considerations of programmes across different regions, woredas and kebeles.

The analysis of the collected qualitative data will rely on triangulation to understand policies, regulations, relevant strategies and guidelines, service provisions, implementation practices, programme monitoring etc to distil lessons and recommendations for improvement. If any two sets of findings contradict each other, the team will investigate more deeply to understand the causes of the discrepancies.

In the event that findings from various data collection methods do not converge to the same conclusions, the evaluation team will set out a detailed list of the areas and elements where discrepancies exist to explore each more carefully using relevant methods. This could for instance entail conducting additional interviews with the stakeholders concerned to reflect on the evidence that was generated (e.g., if quantitative results differ from an informant's answers). Alternatively, the team will assess the quality of secondary data (in terms of statistical power, variance, quality assurance process etc.) to understand whether potential errors might cause the disparity in findings.

8.3 Quality assurance

Overall, quality control is treated as a core responsibility of the team leader, who is mandated to assess and ensure the quality of all deliverables. Continuous communication with UNICEF and project stakeholders throughout all phases of the project is EPRI's approach to overcoming potential challenges to the evaluation as they come along. A weekly meeting has been set up for this end, complemented by constant communication via e-mail and other platforms.

The final report and all related deliverables will be developed in line with the UNICEF-Adapted UNEG Evaluation Report Standards and will therefore be well-structured, logical, clear and complete, provide a clear and comprehensive description of the object of evaluation, its purpose, objectives and scope, a transparent description of design and methods, with evidence-based findings responding directly to the evaluation criteria and conclusions developed based on reasonable judgements, evidence and arguments. Recommendations will be relevant to the evaluation and will be developed together with relevant stakeholders. Quality assurance and related procedures will be introduced at all levels. Furthermore, the approach to the evaluation, developed by EPRI is explicitly based on the UNEG Norms and Standards for Evaluation, as follows:

a) Internationally agreed principles, goals and targets
b) Utility
c) Credibility
d) Independence
e) Impartiality
f) Ethics
g) Transparency
h) Human rights and gender equality
i) National Evaluation Capacities
j) Professionalism

At UNICEF, the Evaluation Manager provides the first level of quality assurance for key deliverables before submitting them to the Regional Office for a second review. The evaluation team reports directly to the UNICEF Ethiopia Country Office Evaluation Manager, who is accountable to the Country Representative and ensures quality assurance throughout the process.

An Evaluation Reference Group (ERG) will be established in an advisory capacity to strengthen the evaluation further and ensure quality. The ERG will comprise representatives from UNICEF, donors, the government, and other key stakeholders. Its main role will be to review and provide feedback at critical milestones, such as the inception report and the draft final report, and to advise on recommendations to ensure they are specific, relevant, and actionable.

Once deliverables have passed internal quality assurance, they will be shared with the ERG for comments and feedback, as well as with other bodies such as the Programme Management Team (PMT) and the Country Management Team (CMT). The inception report and draft evaluation reports for each phase will be presented to the ERG at validation workshops. The evaluation team will address all comments and feedback from the ERG and will incorporate these into the deliverables.

8.4 Limitations and mitigation measures

The evaluation team has identified the following limitations:

M&E Framework completeness: The evaluation team highlighted that the programme M&E framework needs to be updated: outcome and impact indicators are not finalised, and most existing indicators are process/output-focused. Monitoring roles and reporting frequency are unclear: responsibilities for tracking indicators across UNICEF, government, partners and schools are not assigned. Indicators specifically focusing on AWDs are missing. **Mitigation:** The evaluation team will work with UNICEF in continuously improving this framework to ensure that outputs and outcomes are properly tracked and that it is actionable for gathering evidence and measuring programme implementation. UNICEF, as a response, is presently reviewing the framework.

- **Implementation start dates:** The implementation of some interventions is already underway - some components started before the evaluation (education, nutrition, WASH), undermining the feasibility of attribution of outcomes to certain interventions. **Mitigation**: To the extent that the roll-out allows at baseline, the team will attempt to select schools where the least number of interventions are currently active – the team will work on deriving evidence from qualitative data to understand programme impact.

- **Absence of pure comparison group**: The design does not capture schools from localities where the programme is not active as these are determined to be contextually different from intervention woredas. Furthermore, several interventions are national campaigns limiting the ability to draw strict comparisons with targeted schools. Finally ethical concerns prevent withholding the package from nearby schools. The proposed design uses the basic package schools as a comparator limiting the ability to quantitatively attribute changes to the programme to some sectoral interventions (especially in nutrition and health). Additionally, the Basic package group may differ systematically from enhanced package schools (e.g., socio-economic status, teacher quality). **Mitigation**: Collect detailed baseline covariates and apply regression adjustment or matching methods; use Contribution Analysis to complement quantitative attribution by documenting implementation context.

- **Non-randomized allocation**: Schools are pre-assigned to intervention packages, limiting causal inference due to potential selection bias. Because schools were purposively selected on practical/operational criteria (e.g. infrastructure, accessibility, security, partner opportunities), ruling out unbiased random assignment. **Mitigation**: Use Difference-in-Differences (DiD) with baseline adjustment and school-level covariates; include sensitivity analyses to test robustness of results against unobserved confounders.

- **Programme clarity**: There is a lack of clarity on the MHPSS component, as the related guidelines/implementation details have not yet been developed. More generally, some delivery modalities remain unclear, reducing clarity on who is reached and how

to sample and measure. **Mitigation:** The team will continuously work with UNICEF to obtain the required information prior to the launch of these interventions.

Timing of follow-ups: Midline (1 year) and endline (2 years) may not fully capture long-term outcomes. **Mitigation**: Focus midline on short- and medium-term outcomes; document longer-term processes and anticipated outcome trajectories; consider follow-up beyond 2 years if resources allow.

- **Data quality and availability concerns**: school attendance, examination, and performance records may be incomplete or unreliable as secondary sources. Another layer of uncertainty is represented by the accessibility of cost data (particularly at the sub-federal level) for the government-run programme interventions, adding to personnel, overheads, and school costs. **Mitigation:** The team will work with UNICEF and stakeholders to ascertain the availability of this data, it will be proactive in ensuring that where data is not immediately available, that the team can collect it at the subsequent data collection phase. Where the data limitations affect the proposed methodologies, the team will propose alternative approaches and will document how these shortcomings affect the ability to answer the evaluation questions.

9 Evaluation Workplan

9.1 Baseline Evaluation (Sept–Nov 2025)

9.1.1 Baseline Objective

The baseline evaluation's primary objective is to establish a comprehensive pre-intervention snapshot of the programme's context and key indicators across education, nutrition, CP, WASH, and MHPSS components. The baseline will seek to better understand the relevance of the programme in the intervention areas. As such, this phase will document initial conditions and needs in target schools and provide reference values for all outcome indicators before full implementation. It will also validate the programme's theory of change assumptions against on-the-ground realities – given some interventions have already started, the baseline will assess the relevance of these interventions' design. The baseline findings will inform programme implementation by identifying critical gaps and setting a benchmark for measuring change at midline and endline.

9.1.2 Draft Baseline Overall Evaluation Questions

1. **Baseline Status of Key Outcomes:** *What is the current status (baseline values) of the programme's intended outcomes and outputs in target schools and communities? –*
2. **Needs, Barriers, and Context Relevance:** *What are the main education, protection, and WASH needs and barriers faced by adolescent boys and girls in and out of school?*
3. **Baseline Disparities and Inclusion:** *How do baseline conditions differ across various groups and locations?*

4. **Initial Implementation Readiness:** *What is the status of programme roll-out at baseline and are systems in place for implementation?* (This question documents any early activities or preparatory steps undertaken by September 2025 – such as training of some teachers or initial community meetings – and assesses the readiness of schools and partners to absorb the upcoming interventions.)

5. **Ownership:** *Has the project design adequately considered national, regional and UNICEF priorities and integrated these in its design to ensure continuity and ownership?*

9.1.3 Enumerator Training and Piloting

In August and September 2025, the evaluation team will pilot test all baseline instruments in a non-sampled school and community in 1-2 woredas bordering the ones where the package is being piloted. This piloting checks question clarity, cultural appropriateness, and logistical flow. Enumerators and facilitators will undergo a thorough training workshop before data collection. The results of pre-testing of the survey tool will the shared in a pre-test report which the evaluation team will thoroughly review – any final changes to the tools will be made at this stage. The training of enumerators will cover:

- Survey administration techniques (sampling within schools, standardising interview protocols).
- Ethical considerations when interviewing children and vulnerable populations (in line with UNICEF procedures on consent, confidentiality, and child protection referrals in case sensitive issues arise).
- Role-play exercises and practice with the tools in the local language to ensure enumerators can comfortably handle both quantitative surveys and FGD facilitation. Adjustments from the pilot (e.g. rephrasing ambiguous questions) will be finalized by early September

9.1.4 Data Collection

Baseline data collection will take place before schools close in September 2025 in 2 Phases. Data collection with Institutional stakeholders will take place in early September, in Phase 2 data collection will take place in week 3 or 4 of September after the New Year Holidays 2025, once schools reopen.

Phase	Timing	Focus	Responsible Activities
Phase 1: Institutional and Adult Respondents	Early September 2025	Collection of data for baseline indicators that do not involve direct interaction with children. This includes school infrastructure, WASH conditions, teacher and school leadership	Tool finalisation and piloting; data collection by trained enumerators; initial data cleaning and analysis.

		capacities, and administrative data from school and woreda officials.	
Phase 2: Child-Centred Data Collection	**Late September 2025**	Data collection from children on key outcome and experience indicators, including nutrition, life skills, child protection, attendance. Tools will be simplified and child-friendly.	Supervised fieldwork; real-time quality assurance; strict adherence to safeguarding and child protection

9.1.5 Baseline Reporting Process

As data analysis concludes in early October, the evaluators will synthesise findings into a baseline report. A preliminary results briefing may be held with UNICEF Ethiopia and partners to validate findings and gather initial feedback (this could be an informal presentation of key statistics and insights, ensuring no major surprises in the report). Stakeholder validation at this stage helps ensure the baseline interpretations are accurate and buy-in is built for using the data.

9.1.5.1 Outputs and Reporting

- **Baseline Evaluation Report – Draft:** A comprehensive draft report will be delivered in **October 2025**. This report will detail the baseline findings across all focus areas, with data visualisations (tables, graphs of baseline indicators) and qualitative summaries. It will not draw conclusions about programme effect (since interventions just started) but will include *observations and recommendations* relevant to programme planning (for example, highlighting which needs are most urgent or if any aspects of the intervention need acceleration based on baseline gaps).
- **Baseline Evaluation Report – Final:** After incorporating feedback from UNICEF, donors, and implementing partners (who will review the draft), the final baseline report will be finalised by **November 2025**. This final report will be a reference document for all stakeholders, serving as the officially agreed-upon baseline values for later comparison.
- **Baseline Presentation:** A presentation (PowerPoint slide deck) summarizing key baseline findings and recommendations will be prepared for a stakeholder dissemination meeting. The evaluation team will present the baseline results to UNICEF, donor representatives, implementing partners, and government officials. The presentation will focus on actionable insights (e.g. “Areas of concern identified at baseline”) so that the programme team can immediately use this information for course-setting.

9.1.6 Midline Objective

The midline evaluation aims to assess interim progress and implementation quality about halfway through the programme, providing an opportunity for course correction and learning.

The objectives at midline are two-fold: (1) Effectiveness (Early Outcomes): Determine what initial changes or trends are observable approximately 6–9 months into implementation, compared to baseline, in key outcome areas (education, WASH, nutrition, child protection, MHPSS). (2) Process Evaluation: Examine how well the programme has been implemented thus far – focusing on fidelity to the plan, the functionality of the integrated approach, and any challenges that have arisen. In essence, midline serves as a "health check" for the programme, verifying if it is on track to achieve its goals and providing actionable feedback to managers for the remaining implementation period.

9.1.7 Midline Focus Areas

- **Early Outcome Changes:** Identify any measurable improvements or changes since baseline in the programme's target outcomes. For example, have student enrolment or attendance rates started to improve in intervention schools? Are children participating in life skills sessions and making use of health and WASH facilities? Are showing signs of improved mental well-being? Are new water points functional and increasing water availability at schools? Midline will focus on detecting these early outcome "signals" even if ultimate targets are not yet met.
- **Implementation Fidelity and Quality:** Assess the extent to which activities have been delivered as planned across all components. This includes coverage (e.g. number of schools that have received teacher training or WASH facilities versus the plan), timeliness (whether activities like training, club strengthening, construction happened on schedule), and quality (stakeholder perceptions of how well these were done). This focus area addresses whether the programme is being implemented *according to design* and identifies any gaps (for instance, if certain schools or groups have not been reached as expected).
- **Integrated Approach in Practice:** Investigate how the integration of package interventions is working on the ground. Are there signs of synergy (such as teachers noting that new WASH facilities improve attendance or menstrual health or that MHPSS training makes their teaching more effective)? How are the different components coordinating at the school and community level? The midline will pay special attention to the "value-added" of the multi-sector approach at this stage – capturing any qualitative evidence of positive interaction between components, or conversely any operational inefficiencies or coordination issues that need adjusting.
- **Beneficiary Feedback and Equity (Mid-course):** Gather feedback from various beneficiary groups on their experience with the programme so far. Do in and out of school children, especially girls, feel differences in benefit at this stage? Are girls able to participate equally (e.g. are the menstrual hygiene facilities making a difference for adolescent girls' attendance)? The midline will check whether any group is being left behind in the initial roll-out and whether adjustments are needed to ensure inclusiveness in the second half of the programme.

- **Emerging Outcomes vs. External Factors:** Consider any external influences or context changes since baseline that might be affecting results. Midline will contextualize early outcome findings by noting such factors.

9.1.8 Midline Data Collection Methods

The midline will employ a similar approach to baseline, with a mix of repeat quantitative measures (to track changes in key indicators) and deep qualitative inquiry (to understand process and early outcomes):

- **Quantitative Follow-up Surveys:** A subset of the baseline survey instruments will be repeated at midline to capture trend data. Specifically:
    - *Student and Teacher Surveys:* The same schools and, where possible, the same respondents (students and teachers) from baseline will be targeted for a follow-up survey. The surveys will include core "baseline indicators" again to measure change. Since only ~9 months have passed since baseline, large outcome shifts may not yet occur, but these data will indicate direction of change (e.g. slight uptick in awareness or behaviours or none at all). Repeating the surveys also helps maintain engagement and provides additional data points for the eventual time-series analysis.
- **Qualitative Methods – FGDs and KIIs:** The qualitative component at midline is crucial for process evaluation:
    - *FGDs with Students:* By midline, students in primary schools will have experienced some programme activities (e.g. life-skills clubs, vaccination campaigns, WASH improvements). We will conduct a few FGDs with student groups and OOS adolescents. These discussions will explore how they perceive changes in their school environment: Are they noticing differences in teaching methods? Do they feel safer or more supported? What do they think of the clubs at their school or their safe spaces? Content focus is on "early progress pathways and implementation fidelity", meaning we want to hear if the programme's presence is being felt and how.
    - *FGDs with Parents:* Similar FGDs as baseline will be held, but now the topic is their *perception of changes*: Do parents observe their children benefiting (are children more eager to attend school, do they talk about new things at school)? Also, parents can comment on community-level changes (for instance, any improvements in community attitudes or inclusion).
    - *FGDs with frontline workers:* These will delve into their experience implementing the programme: Have they received the training and materials promised? Are they using new techniques and what has been the result? This also covers how workers feel about the integration – e.g. "Has having better water at school or play materials made teaching easier?" – addressing integrated approach value-add qualitatively.

- Each FGD at midline will also probe fidelity issues: for example, in a frontline workers FGD we might ask if all planned training happened and was it on time, or in a parents' FGD, whether certain activities (like community sessions) took place as they were supposed to.
- *Key Informant Interviews:*
  - **Implementing Partners/UNICEF Field Staff:** Interviews with programme managers in all woredas will gather an internal perspective on implementation progress. Topics include coordination among the three components, management of any delays or issues, and reflections on what is working well or not.
  - **Local Officials/School Management:** Interviews with officials from the education bureau or school cluster supervisors to get an external perspective on the programme at mid-point – do they see improvements in schools? How is government involvement (are they participating in monitoring, providing support)?
  - **Donor Perspective (if available):** A brief check-in KII with a donor representative or UNICEF management might be included to understand expectations and any strategic adjustments by midline.
- *Direct Observations:* Where feasible, the midline will include field observations of programme activities in action. For example, if a teacher training or a school club session coincides with the midline visit, the evaluators will observe it using a checklist (noting participant engagement, quality of facilitation, integration of themes). Similarly, evaluators may visit a sample of WASH facilities constructed to observe their usage and maintenance. These observations enrich the process evaluation by providing firsthand evidence of fidelity and quality.

9.1.9 Enumerator Refresher and Tool Refinement

Prior to midline data collection (early January 2026), the evaluation team will update the data collection tools based on lessons from baseline and the monitoring data. Some new questions may be added to capture progress and some baseline questions may be simplified if they were difficult. A **refresher training** for enumerators (likely many of the same team from baseline, ideally) will be conducted, by Assosa University. This shorter training (2–3 days) will reorient the team to any new protocols, emphasize reliability in re-measuring indicators, and reinforce ethical practices. Enumerators will be briefed on changes since baseline (for instance, if security or cultural context has shifted) and any new consent needed for follow-up. A quick pilot or pre-test of the midline tools will be done in one school to ensure the flow is smooth, but a full pilot is not needed as these are largely repeat measures.

9.1.10 Data Collection - Fieldwork

The midline data collection is slated for **May-June 2026** , with an efficient schedule since many tools are repeats:

- Survey teams will revisit the target schools to administer follow-up surveys. Because familiarity is established, this should proceed faster than baseline. Data will again be captured electronically where possible for speed.
- FGD/KII teams will conduct qualitative sessions largely in parallel, coordinating with schools and communities to gather participants (with help from school principals or community leaders to mobilize parents).
- The evaluation team will ensure that midline fieldwork does not overly disrupt classes or activities – e.g. scheduling after-school FGDs or during free periods.
- Ongoing quality control will include supervisors observing a sample of interviews/FGDs and daily data checks as done in baseline.

9.1.11 Midline Reporting Process

A critical aspect of midline is ensuring findings are immediately useful. Therefore, the evaluators will convene a midline debrief workshop in early March 2026 (with UNICEF and key partners) *before* finalizing the report. In this session, preliminary findings on what's working and not working will be discussed, and stakeholders can provide interpretations or additional context (for example, explaining a challenge that the data flagged). This participatory review helps validate the findings and directly feeds into recommendations for programme adjustments in 2026–2027.

9.1.11.1 Outputs and Reporting

- **Midline Evaluation Report – Draft:** A draft midline report will be produced by **early August 2026**. It will detail progress since baseline, covering both quantitative midline findings and qualitative insights on implementation. Importantly, this draft will include a set of *practical recommendations* for the remaining programme period (e.g. suggestions to intensify certain activities, address any lagging areas, or improve integration/coordination). The draft will be shared with UNICEF, donors, and implementing partners for review and feedback.
- **Midline Evaluation Report – Final:** The final midline report, incorporating stakeholder feedback, will be completed by **August 2026**. This report will serve both accountability and learning purposes – documenting the state of the project at mid-term and explicitly pointing out any course corrections agreed upon. It will likely be shorter than the endline report, focused on interim findings and recommendations, but sufficiently detailed to justify the recommendations made.
- **Midline Presentation:** The evaluation team will prepare a concise presentation of midline results for a broad stakeholder meeting in late August 2026. The presentation will highlight key achievements to date (to reassure donors of progress), as well as

challenges and recommended actions (to spur decision-makers towards necessary changes). UNICEF, government officials, and partner NGOs in the region will be invited to this presentation to ensure a common understanding of midline learnings. The presentation can also be used internally by UNICEF to brief their senior management or the donor on progress.

9.2 Endline Evaluation (June–Sept 2027)

9.2.1 Endline Objective

The endline evaluation's objective is to **assess the overall performance, outcomes, cost-effectiveness, scalability and lessons of the programme upon its completion**. It is a summative evaluation that will determine the extent to which the integrated programme achieved its intended results and the degree to which observed changes can be attributed to the programme interventions. The endline also aims to evaluate sustainability of outcomes and extract learnings for future programming. In summary, the endline will provide accountability to donors and stakeholders by measuring results (effectiveness and impact), while also analysing *how* those results were achieved (process and contribution) and *what* can be sustained or replicated. A contribution analysis will be applied to evidence the programme's role in observed changes, and a thorough process documentation will explain the implementation journey. A cost-effectiveness analysis will assess which combination of interventions offer the best value for money and which are the most feasibly scalable. This final evaluation phase will consolidate findings to answer all key evaluation questions derived from the theory of change and inception report.

9.2.2 Endline Focus Areas

- **Achievement of Outputs and Outcomes:** Verify which planned outputs were delivered and the level of outcome achievement across all programme components by 2027. This includes final outcomes in health, education and WASH, child protection and MHPSS (improved dietary and nutrition behaviour, mental health, menstrual health, functional water and sanitation facilities, improved hygiene behaviour etc). The focus is on quantifying and describing the changes from baseline to endline for each result area and comparing achievement across different combinations of the package

- **Contribution and Attribution:** Examine the extent to which the programme contributed to the observed changes, distinguishing programme influence from other external factors. This focus area uses contribution analysis to build a credible narrative of causality: linking outputs to outcomes and checking if outcomes occurred in the absence of other major drivers. The evaluation will probe questions like "Would these changes have likely happened without the programme, or did the programme make a crucial difference?" Evidence for contribution will come from both quantitative change data and qualitative stakeholder attributions (e.g. what do beneficiaries themselves credit for the changes in their lives?). Attribution to the programme will rely on DiD estimation methods and specifically on those schools receiving additional components of the minimum package.

- **Process Learnings and Implementation Performance:** Reflect on how the programme was implemented over the full duration – what worked well operationally and what challenges were encountered. This includes assessing the integrated delivery model: how effectively did UNICEF and government partners manage coordination throughout? Were there adaptations made along the way (perhaps as a result of the midline findings) and did they improve results? The endline will document the implementation process, highlighting critical success factors and bottlenecks of the integrated approach over the two years.
- **Impact on Equity and Inclusion:** Determine how the programme impacted different groups. Did vulnerable groups attain the same outcome improvements as other children? Did girls benefit as much as boys in terms of education and safety? How did children with disabilities fare? This focus area ensures the evaluation addresses equity by analysing disaggregated outcomes and gathering perspectives from those sub-groups about their experience. It will identify if any group was left behind and why, thereby informing future inclusive strategies.
- **Cost-effectiveness:** Determine whether the programme is delivering value for money and establish which combination of interventions achieves the highest impact at the lowest cost per beneficiary. The team will compare cost-effectiveness across schools implementing different package variations. Lessons from the cost-effectiveness analysis will also help inform the sustainability and scalability of the programme.
- **Sustainability and scalability:** Assess the likelihood that the benefits of the programme will continue after funding ends. This includes looking at institutionalization (e.g. are trained teachers now part of the regular system and continuing new practices? Have local authorities taken ownership of maintaining WASH facilities? Are community structures like gender clubs and community centres likely to continue activities?). The evaluation will also review any plans for scale-up or handover: for instance, how UNICEF or the government plan to maintain or expand the interventions using other resources.

9.2.3 Enumerator Training and Preparation

By 2027, it is likely some original enumerators may not be available, so the team will recruit/train a fresh (or refreshed) cohort. A comprehensive training similar to baseline will be conducted in May 2027. This covers all new and revised tools (including training on any participatory methods like MSC story collection, which requires good facilitation skills). The training will emphasize consistency with baseline measures to ensure data comparability. A pilot test of the endline instruments will be done in one school to ensure clarity and timing (especially if new sections are added to capture final). Ethical protocols will be reiterated, particularly since some sensitive questions (e.g. on protection or psychosocial well-being) might be included in endline tools to assess changes.

9.2.4 Endline Reporting Process

9.2.4.1 Validation and Interpretation

Before finalizing the report, a stakeholder validation workshop will be conducted in late August or early September 2027. Here, the evaluation team will present preliminary endline findings to UNICEF, government representatives, and implementing partners. The purpose is to:

- Verify factual accuracy (ensure no major output or event was missed or misrepresented).
- Discuss the interpretation of results, especially for contribution and sustainability. Stakeholders can confirm or challenge the evaluators' emerging conclusions (for instance, if the evaluation attributes improved attendance to the programme, stakeholders might add insight like "yes, and also the government's free meal program helped" which will be incorporated).
- Build consensus on the lessons and recommendations, increasing the likelihood they will be acted upon.

The validation feedback will be used to refine the analysis and craft well-founded recommendations.

9.2.4.2 Outputs and Reporting

- **Endline Evaluation Report – Draft:** A comprehensive draft endline report will be prepared by **early September 2027**. This report will be the most extensive, covering the entire evaluation scope. It will include an executive summary for high-level audiences, a full description of the evaluation approach, detailed findings for each question, and evidence-backed conclusions and recommendations. The draft will be shared with UNICEF, donor, implementing partners, and key government stakeholders for review. Given the broad audience, the draft will be written in clear, non-technical language (with technical annexes for detailed methods and data).
- **Endline Evaluation Report – Final:** The final endline report will be completed by **late September 2027** after incorporating feedback. This final report will serve as the definitive record of the programme's performance and lessons. It will likely be formally published or at least disseminated among all partners. The final version will ensure that all stakeholder comments are addressed and that the report is balanced and evidence-based (any disagreements or differing interpretations can be noted in the report if needed).
- **Endline Presentation:** The evaluation team will deliver a presentation of the endline results to stakeholders around the time of the final report release (Sept 2027). This presentation will cater to both technical and non-technical audiences, summarizing key findings, success stories, and critical recommendations. It will be used in a dissemination workshop or closing meeting with UNICEF, donors, government

ministries (Education, Water, etc.), and community representatives. The aim is to communicate what was achieved and what should happen next (post-program).

- **Policy Brief (Endline):** In addition to the report and presentation, a concise Policy Brief (2-4 pages) will be produced at endline. This brief will distil the most important findings and recommendations relevant to policy makers and high-level stakeholders. It will focus on actionable insights: for example, policy implications for integrating services for refugee education, or recommendations for government on sustaining school WASH. The language will be non-technical, highlighting achievements (to garner support for similar approaches) and critical needs for sustainability (to influence policy/resource allocation). This brief will be shared widely – including with regional and federal government officials, donor headquarters, and potentially made public to inform the broader development community about the integrated approach's results.

10 Risk Mitigation

Table 6: Risk mitigation matrix

Risk Category	Description of Risk	Likelihood	Impact	Mitigation Measures
Ethical Risks (Children)	Emotional distress or fatigue among children due to inappropriate tools or poor facilitation.	Medium	High	Use child-friendly, brief tools; train enumerators in child safeguarding; secure ethical clearance.
Sampling and Access	Inability to reach target participants due to absenteeism, school closures, or displacement.	High	High	Coordinate with schools and local officials in advance; oversample in planning; use replacements.
Security and Instability	Heightened risk of data collection disruption in conflict-affected or fragile areas.	Low	High	Conduct risk assessments prior to deployment; avoid high-risk zones; partner with local actors.
Enumerator Capacity	Enumerators may lack the necessary training to engage ethically and sensitively with children.	Medium	Medium	Deliver intensive in-person training with practical roleplays and child engagement protocols.
Language and Translation	Misinterpretation of questions or concepts due to dialect/language mismatch.	Medium	Medium	Hire local-language speakers; pre-test tools in multiple dialects; review translations carefully.
Data Quality (Quantitative)	Measurement errors, non-responses, or social desirability bias affecting validity.	Medium	Medium	Pilot surveys; build rapport; include quality checks and daily field team debriefings.
Data Quality (Qualitative)	Power dynamics and cultural norms may suppress honest views in FGDs or KIIs.	Medium	Medium	Use skilled facilitators; separate FGDs by age/gender; triangulate with other methods.
Timing and School Calendar	Misalignment of data collection phases with academic calendar or seasonal migration.	High	Medium	Align with school calendar; phase data collection to avoid exam periods and migration cycles.
Stakeholder Engagement	Limited cooperation or buy-in from local government, schools, or community members.	Medium	Medium	Engage REBs, WEOs, and school leadership early; explain value of the evaluation.
Data Storage and Privacy	Breaches in confidentiality or unauthorised data access affecting participant safety.	Low	High	Encrypt data, anonymise identifiers, restrict access to authorised team members only.
Inclusion and Representation	Exclusion of specific subgroups (e.g. minorities, children with disabilities) may bias results and cause community tensions.	Medium	High	Map all vulnerable groups during sampling design; ensure intentional inclusion in survey and qualitative tools.
Ethical Clearance Delays	Delays in securing required ethical and administrative clearances from national or local authorities.	Low	High	Prepare ethical protocols in advance; engage with authorities early; use

Risk Category	Description of Risk	Likelihood	Impact	Mitigation Measures
				UNICEF facilitation for clearances.
Stakeholder Review Delays	Slow review of the inception report and evaluation tools by stakeholders delays start of fieldwork.	Medium	Medium	Agree on timelines and non-objection protocols with stakeholders; engage ERG early.
Programme Staff Turnover	Organisational changes limit access to institutional memory or knowledgeable staff.	High	Medium	Use programme documentation to fill gaps; triangulate with newer staff and administrative records.
Cultural Sensitivity (Difference Blindness)	Lack of awareness of local customs, norms or sensitivities among external staff causes distrust or offence.	Low	Medium	Deploy local enumerators; sensitise external team members; provide cross-cultural orientation.
Informed Consent in Children with Disabilities	Some children may not be cognitively or legally capable of providing informed consent independently.	Low	High	Obtain guardian and child assent; use community workers for children with disabilities.
Confidentiality for Vulnerable Populations	Sensitive information from vulnerable children may lead to identification or unintended harm even with anonymisation.	Low	High	Use neutral pseudonyms and aggregate reporting; assess harm risk per case study before publication.
Pandemic-Related Disruptions	New waves of COVID-19 or other infectious diseases may interrupt fieldwork.	Low	Medium	Plan for remote/phone-based data collection; follow public health protocols; stagger fieldwork if needed.

11 Ethical considerations

11.1 Ethical approach

The evaluation team recognises that safeguarding and ethics are not ancillary concerns but foundational principles in conducting responsible evaluation. The team will adopt a “do no harm” approach throughout, guided by MDS's Safeguarding Policy and aligned with UNICEF’s Ethical Guidelines for research and evaluation, the Ethical Research Involving Children (ERIC) guidance, and UNEG principles. Our detailed ethical protocols and guidelines to be followed by us and out partners.

Core Ethical Framework: Our ethical model is grounded in three interdependent commitments:

1. **Anticipate** potential harms—whether physical, emotional, or social—arising from participation in the evaluation;
2. **Mitigate** these risks through preventive measures, inclusive methods, and tailored protocols; and

3. **Address** any harms swiftly, through robust, survivor-centred reporting, referral, and redress systems.

Duty of Care: We hold a dual duty of care—to our evaluation team and to all stakeholders participating in the baseline and evaluation:

- **To the team**: All team members will undergo risk assessment prior to field deployment. They will be trained on safeguarding, business ethics, and, where relevant, hostile environment protocols. A risk register will be maintained throughout, and appropriate duty-of-care procedures will be in place to prevent or manage operational risks.
- **To stakeholders**: Ethical reciprocity underpins our engagement, especially with vulnerable populations. We will ensure participants—including children and marginalised groups—are treated with dignity and agency. Reciprocity will be contextually defined in collaboration with participants to maximise mutual benefit and minimise burden or extraction.

Alignment with Ethical Standards: The evaluation will fully comply with UNICEF's ethical and safeguarding standards. Table X below outlines how the evaluation operationalises the UNEG principles:

UNEG Principle	Implementation in the Evaluation
Integrity	The team will operate independently from programme implementation actors and will maintain objectivity. Any conflicts of interest will be declared. Multiple viewpoints will be presented transparently if findings diverge.
Accountability	Stakeholders will be fully informed of the evaluation approach through consultations (e.g., with the ERG). A safeguarding referral protocol will be activated immediately upon any reported violation. Regular updates will be shared with the Evaluation Manager.
Respect	Gender-sensitive methods will be used, including same-gender enumerators where appropriate. Separate interviews will be conducted with women and men. Tools will be tailored for children's age and comprehension.
Beneficence	All voices—including those of vulnerable groups—will be reflected in findings. An updated risk register will be used during each data collection round to ensure ongoing adherence to "do no harm" principles.

Ethical Engagement with Children: Given that children are primary stakeholders in the integrated programme, they will be included in the evaluation using child-friendly, rights-based approaches. Ethical protocols for child engagement include:

- **Risk and Benefit Assessment**: Risks of child participation will be assessed before each data collection round. Inaccessible or insecure contexts may warrant excluding children to protect their safety.

- **Privacy and Confidentiality**: All interviews with children will ensure privacy without compromising child protection. Identities will be anonymised in all data and reporting.
- **Informed Consent and Assent**: Consent will be obtained from caregivers, and assent from children, with a minimum two-day window prior to data collection. Children will be informed of their rights, including the right to withdraw at any point.
- **Compensation**: Participation will not involve payment. However, basic refreshments and breaks will be provided. Interviews will be limited to a maximum of 1.5 hours to minimise disruption.

These safeguards will also apply, with appropriate adjustments, to adult respondents, particularly those who are vulnerable or from marginalised groups.

EPRI actively seeks to include persons with disabilities in research that concerns them or where their perspectives are often neglected, in line with the **UNEG Guidance on disability inclusion**. Key accommodations and ethical considerations include:

- *Accessibility:* All research materials (surveys, consent forms, interview guides) are assessed for accessibility. For those with visual impairments, we provide forms in Braille or have consent scripts read aloud. For those with hearing impairments, we engage sign language interpreters (who also sign confidentiality agreements) or use written communication if suitable. Venue selection for interviews or focus groups takes into account physical accessibility – e.g., choosing wheelchair-accessible locations, scheduling at ground-floor venues, etc. If an area has difficult terrain, we arrange transport for those with mobility challenges so they can participate without undue hardship.
- *Time and Patience:* Interviews or surveys might take longer with some participants with disabilities (depending on the nature of disability). EPRI researchers allocate extra time and are patient, ensuring the person does not feel rushed. We allow caregivers or aides to be present if the participant wants them (while still ensuring we get the individual's own responses as much as possible). We also respect if a person uses assistive devices or technology to communicate; our team will be trained in how to interact via those means.
- *Respect and Agency:* Just as with others, we treat adults with disabilities as autonomous individuals capable of providing consent and insight. We avoid paternalistic assumptions. However, if someone's disability impacts their cognitive ability to consent, we follow a similar protocol as with children: get consent from a legal guardian or caretaker *and* assent from the individual to the extent they can provide it. We apply the principle of **"Nothing About Us Without Us"**, meaning if the research topic involves issues affecting people with disabilities, we ensure they or their representative organisations are consulted in study design and included in the sample, rather than making inferences about them from others' reports.

- *Privacy:* Some disabilities come with stigma. We ensure privacy for participants with disabilities, especially in contexts where stigma is high (e.g., a person with HIV-related cognitive impairment in a community might fear others knowing). The consent process will include discussion on what information may be shared and with whom, giving the participant control. If, for instance, a caregiver is required to assist in the interview, we check with the participant how comfortable they are with that person hearing their answers, and arrange alternatives if needed (maybe a different assistant).
- *Sensitization of Researchers:* EPRI trains its researchers on disability inclusion etiquette and awareness. This covers respectful language (e.g., using person-first language like "person with disability" rather than labels, unless the individual indicates a preference), not making assumptions about abilities, and how to ask questions in a way that is sensitive. We refer to resources such as the **UNFPA disability inclusion guidance**, which emphasizes human rights-based and inclusive evaluation approaches, to ensure our team's approach is up-to-date and respectful.

Team Composition and Training: All evaluation team members will be trained on research ethics, child protection, and cultural sensitivity. They will be required to sign EPRI's safeguarding code of conduct. A gender-balanced team will facilitate culturally appropriate engagements, especially when discussing sensitive topics.

Data Protection and Ethical Approval: All personal data will be anonymised, encrypted, and securely stored. Data access will be limited to authorised personnel, and all data will be deleted from systems upon submission to UNICEF.

11.2 Ethical approval

The evaluation team obtained ethical approval from UNICEF Ethical Review Board – Health Media Lab. The approval is valid until September 2026 at which point the team will provide updated materials to extend ethical clearance until the endline phase. The Ethical Approval Letter is provided in the annex. The team will seek additional national-level ethical clearance where required.

12 Consolidated Timeline of Evaluation Activities (2025–2027)

The table below maps the key activities for each evaluation phase across the timeline. It ensures coordination of tasks such as tool development, data collection, training (in collaboration with Assosa University), analysis, and reporting milestones.

Table 7: Consolidated timeline of activities

Timeframe	Baseline Phase (2025)	Midline Phase (2026)	Endline Phase (2027)
August 2025	– Finalize evaluation framework and tools for baseline (surveys, FGD/KII guides). – Pilot test instruments in target area for cultural relevance. – Train enumerators & facilitators on baseline tools, ethics, and logistics.	*(No midline activities yet – programme implementation period).*	*(No endline activities yet – programme ongoing).*
September 2025	– Conduct **baseline data collection** in the field: surveys in 84 schools; FGDs with children/parents/community/officials; KIIs with stakeholders. – Monitor data quality daily; address any field issues (replacement schools/respondents if needed).	*(No midline activities).*	*(No endline activities).*
October 2025	– Complete baseline fieldwork by early October. – Perform data cleaning and preliminary analysis of baseline quantitative and qualitative data. – Internal team review of initial findings; prepare summary statistics and transcripts.	*(No midline activities).*	*(No endline activities).*
October-November 2025	– **Draft Baseline Report (60 pages not including annexes)** writing (Late October) – compile findings, baseline values, and initial recommendations. – Submit draft baseline report for UNICEF/partner review. – **Finalize Baseline Report** (November 2025) after feedback. – Deliver **Baseline Presentation** to UNICEF, donors, government in a dissemination meeting (late Sept).	*(No midline activities).*	*(No endline activities).*
Oct–May 2026	– *No evaluation field activities (programme implementation continues; ongoing monitoring by programme).*	*(No midline activities; monitoring of programme continues).*	*(No endline activities; monitoring continues).*

Timeframe	Baseline Phase (2025)	Midline Phase (2026)	Endline Phase (2027)
April 2026	*(Baseline phase complete; programme ongoing).*	– **Midline planning:** Review and refine midline data collection instruments (update surveys, design FGD guides focusing on progress/fidelity). – **Enumerator refresher training** (early May) to prepare for midline. – **Midline data collection** begins: field teams revisit schools to conduct follow-up surveys and new FGDs/KIIs (progress and process focus).	*(No endline activities).*
May- June 2026	*– (No baseline activities).*	– Complete **midline data collection** by mid-June. – Data cleaning and analysis of midline surveys (baseline vs midline comparisons) and qualitative process findings. – Begin drafting midline findings and identifying recommendations.	*(No endline activities).*

Timeframe	Baseline Phase (2025)	Midline Phase (2026)	Endline Phase (2027)
July 2026	– *(No baseline activities).*	– **Draft Midline Report (60 pages not including annexes)** delivered (mid-July) detailing interim findings and proposed course corrections. – Stakeholder review and feedback on draft. – **Finalize Midline Report** (August 2026) incorporating comments. – **Midline Presentation** held with UNICEF, donors, partners to discuss progress and agreed adjustments	*(No endline activities).*
Apr–Dec 2026	*– No baseline activities (baseline data being used as reference).*	*– No midline activities (programme implements adjustments from midline, continuous monitoring).*	– **Preparatory monitoring:** Evaluation team keeps track of programme monitoring reports and context changes to inform endline design. Initial endline planning may start in
January–May 2027	*– No baseline activities.*	*– No midline activities.*	– **Endline planning:** Re-engage with

Timeframe	Baseline Phase (2025)	Midline Phase (2026)	Endline Phase (2027)
			UNICEF and partners to refine endline approach. Update evaluation matrix and develop endline data collection tools (surveys, KII/FGD guides, MSC story protocols, etc.) by April 2027. – Logistical arrangements for endline (reconfirm school samples, recruit any new enumerators) in May 2027.
May- June 2027	*– No baseline activities.*	*– No midline activities.*	– **Enumerator training & pilot (endline):** Train field teams on endline instruments, including participatory methods, and refresh ethics training (early June). – Pilot test endline questionnaires and discussion guides in one community; adjust tools as needed (mid-

Timeframe	Baseline Phase (2025)	Midline Phase (2026)	Endline Phase (2027)
			June). – **Begin endline data collection** in May 2027.
June 2027	*– No baseline activities.*	*– No midline activities.*	– **Endline fieldwork ongoing:** Conduct comprehensive endline data collection through June: repeat surveys in all target schools, FGDs with all stakeholders – Ongoing data quality checks and daily team debriefs to ensure completeness and reliability.
July 2027	*– No baseline activities.*	*– No midline activities.*	– Complete endline fieldwork by early July. – **Data cleaning:** Verify and clean survey datasets, transcribe and translate qualitative data. – **Analysis phase:** Quantitative analysis of changes (baseline vs endline comparisons), qualitative coding and synthesis, and integrated

Timeframe	Baseline Phase (2025)	Midline Phase (2026)	Endline Phase (2027)
			contribution and attribution analyses. – Team drafts initial findings and begins writing endline report sections. – Conduct a **stakeholder validation workshop** (Aug) to review preliminary results and discuss interpretations (feedback incorporated into analysis).
September 2027	*– No baseline activities.*	*– No midline activities.*	– **Draft Endline Report (60 pages not including annexes)** completed and circulated for feedback (early Sept). – UNICEF, donors, and stakeholders review draft; provide comments. – **Finalize Endline Evaluation Report** by mid/late Sept, addressing all inputs. – Prepare and disseminate

Timeframe	Baseline Phase (2025)	Midline Phase (2026)	Endline Phase (2027)
			Policy Brief summarizing key findings and recommendations (by late Sept). – **Endline Presentation** and dissemination event in late Sept 2027, sharing final results and lessons with UNICEF, government, donors, and community representatives.