Impact evaluation of the Home-Grown School Feeding programme in Malawi

School-based Programme Impact Evaluation Window

Inception Note

March 2024

Key Personnel for the Evaluation

WORLD FOOD PROGRAMME – OFFICE OF EVALUATION

Anne-Claire Luzot, Director of Evaluation

Jonas Heirman, Senior Evaluation Officer

Simone Lombardini, Evaluation Officer

Minh Phuong La, Monitoring and Evaluation Officer

WORLD BANK – DEVELOPMENT IMPACT EVALUATION (DIME)

Florence Kondylis, Research Manager

Benedetta Lerva, Economist

Astrid Zwager, Research Officer

Roshni Khincha, Research Analyst

Cox Bogaards, Research Assistant

WORLD FOOD PROGRAMME – MALAWI COUNTRY OFFICE

Simon Denhere, Deputy Country Director

Nicole Carn, Head of Programmes

Danie Svanlund, Head of RAM

Fortune Maduma, Head of School Feeding

Takondwa Moyo, Field Coordinator

March 2024

Contents

1. Background and reasons for evaluation

1. This inception note was prepared by the WFP's Office of Evaluation (OEV) and World Bank's Development Impact Evaluation (DIME) department based upon an initial feasibility assessment undertaken during an in-country mission, which included consultations with key stakeholders in March 2023. Building on the evaluation's Feasibility Assessment, Workplan and MoU, the purpose of the inception note is to summarise and externally present key information about the evaluation, inform stakeholders, guide the evaluation team, and lay out expectations during the various evaluation phases.

1.1. INTRODUCTION

2. This inception note covers the Norwegian Agency for Development Cooperation (Norad)-funded impact evaluation of the Home-Grown School Feeding (HGSF) expansion in the Kasungu district in Malawi. This evaluation will be conducted in partnership with the WFP Office of Evaluation (OEV), the WFP Malawi country office (CO), the WFP School-based Programme (SBP) division, and the World Bank's DIME department.

3. The evaluation takes place from March 2023 to March 2026, covering programme activities from September 2023 to June 2025.

1.2. CONTEXT

4. WFP has supported the government of Malawi with the provision of school meals since 1999. The WFP Malawi CO currently covers 476 schools in eight districts, 88 of which are in the Kasungu district. Norad has committed to funding a Home-Grown School Feeding (HGSF) Programme school feeding expansion of approximately 40 new schools in the Kasungu district. The expansion of the programme provides an excellent opportunity to identify a credible counterfactual to assess the impact of HGSF on children's outcomes, the local economy (specifically on farmers and cooperatives), and provide programme-relevant information to optimise interventions.

1.3. IMPACT EVALUATION IN WFP

5. The WFP Evaluation Policy 2022 defines impact evaluation as "*measuring changes in development outcomes of interest for a target population that can be attributed to a specific programme or policy through a credible counterfactual.*" WFP defines the counterfactual as estimating what would have happened in the absence of the intervention – or establishing that outcomes for the beneficiaries would not be present without the intervention. WFP impact evaluations are prospective, meaning they are planned and designed prior to programme delivery or a new phase of intervention[1]. Impact evaluations align with the timeline of a programme or pilot and usually cover one or more years.

6. The WFP Evaluation Policy (2022) identifies impact evaluations as a third category of evaluation alongside centralised and decentralised evaluations. The Policy states that impact evaluations are managed by OEV and delivered with external technical partners (for example, the World Bank's DIME department) in close coordination with the WFP programme teams at headquarters, regional bureaux, COs, and cooperating partners.

7. In line with the WFP Impact Evaluation Strategy (2019-2026), impact evaluations are primarily delivered through thematic impact evaluation windows in partnership with programme teams and co-funded by participating country offices. Windows are portfolios of impact evaluations managed and co-funded by OEV, that aim to stimulate and shape demand for impact evaluations in priority areas and enable OEV to prepare cross-regional evidence portfolios (evidence syntheses) that meet WFP's global evidence

[1] In exceptional cases where data is available, an IE may be retrospective.

needs. Starting in 2019, WFP has opened three impact evaluation windows, the first on cash-based transfers and gender, the second on climate and resilience, and a third on school-based programmes.

1.4. SCHOOL-BASED PROGRAMME IMPACT EVALUATION WINDOW

8. School-based programmes are one of the most extensive social safety nets worldwide, with an estimated 418 million children currently benefiting from school meals (State of School Feeding Worldwide 2022). Such interventions are intended to promote children's health, nutrition, education, and learning; make communities more resilient; promote gender equality; and support national economies and social stability. There is an urgent need for more evidence to inform the trade-offs in school-based programmes' designs and implementation and to support governments as they scale up their programmes. The SBP IE window was launched in 2021 by OEV, WFP School-Based Programmes (SBP) and the World Bank's DIME department.

9. The SBP IE window offers an opportunity for WFP to answer key questions about school-based programme interventions. While specific evaluation questions for each impact evaluation largely depend on country office priorities, it is expected that impact evaluations conducted as part of the window will answer at least one question within the following three areas of interest:

Health and education systems

- To what extent do different programmes' interventions, and complementary activities, contribute to children's (e.g., nutritional, health and/or learning) outcomes? How do these effects vary by age and gender? What is their relative cost-effectiveness?
- To what extent do the benefits of school feeding programmes vary throughout the year depending on seasonal fluctuations, shocks, and stressors?

Food systems and local economies

- To what extent do different procurement models impact the local economy?
- To what extent can different procurement models be combined with crop and livelihood interventions to support farmers and communities in increasing their resilience and climate adaptation?

Optimization and cost-effectiveness

- To what extent can programmes' characteristics be optimised (including in conflict-affected, fragile, food-insecure, and humanitarian settings)? Which ones are the most cost-effective?

10. In addition to this impact evaluation in Malawi, the window includes three ongoing impact evaluations in The Gambia, Jordan, and Burundi and two pilot impact evaluations in Burundi and Guatemala. OEV is assessing demands from COs for additional impact evaluations to join the window.

1.5. RATIONALE

11. School meal programmes are multisectoral interventions and an essential component of health and education systems that contribute to achieving children's development. Approximately 41 percent of children enrolled in primary school now have access to a free or subsidised daily school meal worldwide. While there is already strong evidence that school feeding impacts children's attendance, more evidence is needed on whether this translates into higher health, nutrition, and human capital outcomes, as well as into improvements in gender, social protection, and social cohesion outcomes.

12. The global annual investment of US$48 billion in school meal programmes creates a huge and predictable market for food, offering an extraordinary opportunity to transform food systems. Evidence from this evaluation will investigate the extent to which different procurement systems can impact the local economy (such as market prices, cooperatives' sales, and farmers' agricultural practices, revenues, and income).

2. Evaluation context and programme description

2.1. CONTEXT

13. Malawi is a landlocked country in Southern Africa with an estimated 22 million people (51 percent female). Average annual income is around US$567 per person (IMF 2023). About 50.7 percent of the population are poor (20 percent extremely poor) and relied on less than $1.90 a day. Food insecurity is rampant, and 37 percent of the children are stunted. About 3.8 million people suffer from hunger. Malawi is ranked 169 out of 191 in the Human Development Index (which is a summary measure of achievements in key dimensions of human development such as life expectancy, education and per capita income), and most of the population lives in rural areas. Primary education has been free since 1994 and in the past decade the primary net enrolment rate has increased to almost 100 percent, and primary school net attendance ratio is 94.3 percent for girls and 93.4 percent for boys. The pupil teacher ratio is high at 1:73, impeding the delivery of quality education. Among children between 5 and 10 years of age, 70 percent go to school without having breakfast (see WFP Decentralized Evaluation of the School Meals Programme in Malawi, 2019).

2.2. PROGRAMME DESCRIPTION

14. WFP has supported the government with the provision of school meals since 1999. The WFP Malawi CO currently covers 523 schools in eight districts. In total 676,000 children in 523 schools have received school meals in 2023. In the Kasungu district, 132 schools are receiving school meals in 2023, for a total of 94,490 children.

15. By providing a hot meal to school-going children, the programme is expected to increase school attendance and enrolment, as well as children's health and nutritional outcomes. Volunteers nominated by the community prepare the meals. School committees choose menus from recommended sets of menus indicated by the government, based on geographical and seasonal availability of products.

16. Schools involved in the home-grown school feeding programme procure food from local smallholder farmers' associations and cooperatives within a radius of 20 km from the school. The procurement process is conducted every three months at the school level. The Kasungu District Council assign to each school an amount to procure food based on student enrolment (Figure 1). Each school's committee decides the menus based on commodities locally available during the season and the available allocated resources (Figure 2). Every three months, the school launch a public tender indicating the quantities needed for each product. The procurement is conducted by each food item independently. These are maize, soybeans, groundnuts, tubers, root crops, and vegetables. Other school-feeding ingredients, such as sugar and salt, are purchased from local shops.

17. Farmer Organizations (FOs), including Farmer Association (FAs) and cooperatives, are invited to participate in the tender process, which usually takes two weeks. Bids from FOs are placed in a tender box (Figure 3), indicating the price for each requested item. At the end of the tender process, the box containing all the bids is opened in front of the school feeding committee and the participating organisation(s). The best FOs with the bidding offer per item are awarded with the contract and asked to supply to the school for the next three months.

18. Figure 4 shows the example of two FOs named that placed a bid for the tender process in December 2022. In all four items, the cooperative on the left was the cheapest bidder. It was, therefore awarded with the contract.

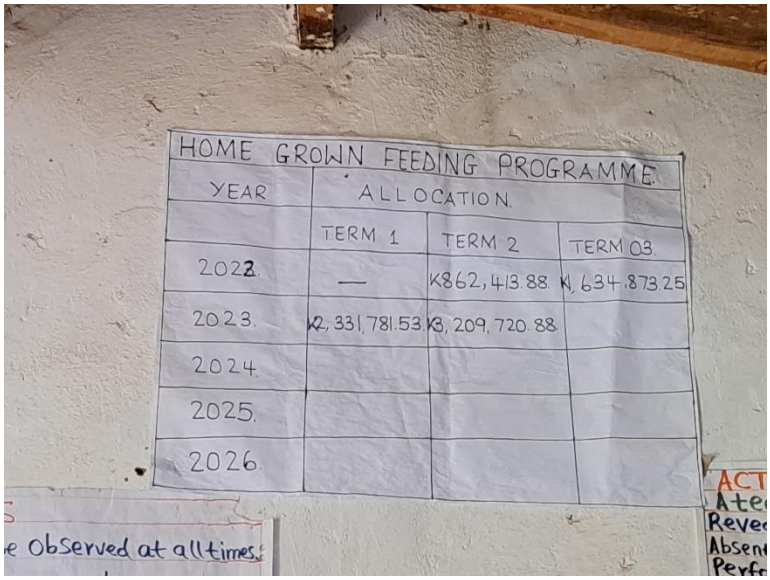

Figure 1: School allocations

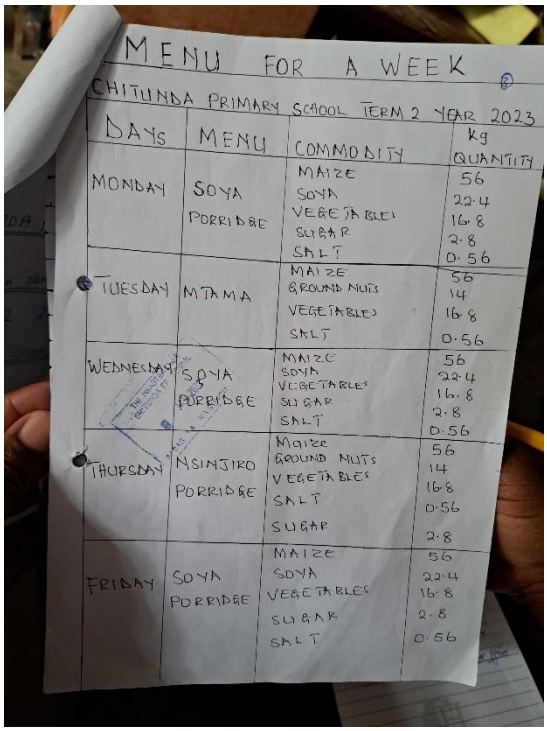

Figure 2: Example school menus

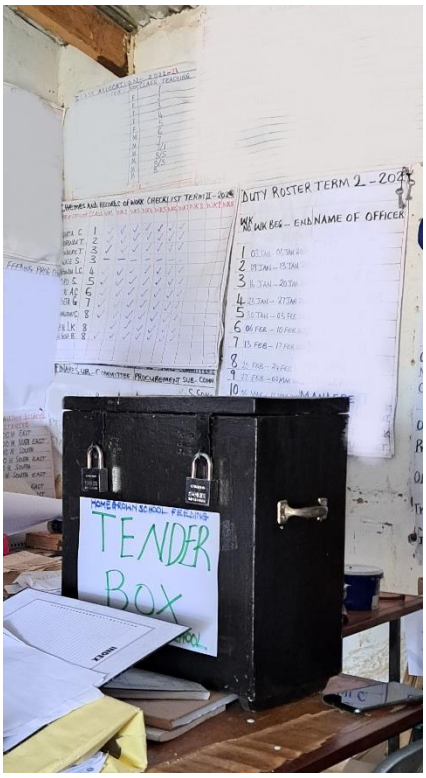

Figure 3: Tender box

Figure 4: Bid documents from two cooperatives

3. Evaluation questions and design

3.1. EVALUATION QUESTIONS

19. This impact evaluation will assess the following main questions:

    a. To what extent is the HGSF programme able to deliver school meals as expected? What is the meals quantity, quality and diversity?

    b. What is the impact of HGSF on children's outcomes (e.g., education, nutrition, and health outcomes)? How do these effects vary by gender?

    c. To what extent do the benefits of school feeding programmes vary throughout the year depending on seasonal fluctuations, shocks, and stressors?

    d. What are the challenges and constraints preventing more farmers and FOs from benefitting from the income opportunities provided by the HGSF programme?

    e. What is the impact of HGSF on the local economy (e.g. farmers and FOs in Kasungu)?

    f. To what extent can different delivery models maximise the frequency of meal distribution?

    g. What does it cost to achieve this impact? Were the interventions cost-effective?

3.2. EVALUATION DESIGNS AND PROGRAMME IMPLEMENTATION

20. Impact evaluation design and programme implementation need to be fully aligned. This section will present how programme activities are expected to be implemented to enable the evaluation designs described in the next section.

21. Following in-depth discussions and considerations between the Malawi CO, government officials at the Ministry of Education and the Kasungu District Council, and the impact evaluation team, three embedded designs are proposed to answer the six questions above. First, a school-level evaluation design will answer questions a, b and c and is based on a **school-level randomised controlled design**. Second, an FO-level evaluation design will answer questions d and e, and use a **randomised encouragement design** potentially complemented by focus group discussions (FGD) and a Discrete Choice Experiment (DCE). Finally, the optimization design will answer question f and it is based on an **A/B testing design**.

3.2.1 School-level evaluation design[2] (using a school-level randomised controlled design)

22. The first component of the evaluation leverages the planned expansion to 44 new schools to assess the impact of HGSF on children's outcomes through a randomised controlled trial design. HGSF activities expanded to 44 new schools in Kasungu district in 2023 and cover a total of 31,000 new students. School feeding activities in additional schools started in September 2023. The expansion to these new schools provided the opportunity to identify a credible counterfactual and answer whether home-grown school feeding will have an impact on children's outcomes.

23. To identify the potential evaluation sample, the evaluation team first conducted a desk review to identify schools in the district that were not already receiving school meals. Kasungu district has a total of 383 public schools. In 2023, WFP was already implementing HGSF activities in 88 schools, of which 54 are Norad-funded and 31 are financed through the UN Joint Programme on Girls Education. In addition to WFP, other organisations are providing school feeding or recently concluded providing school

[2] This component of the study has been pre-registered in the American Economic Association Randomized Control Trials (AEA RCT) registry at: https://www.socialscienceregistry.org/trials/13049

feeding. Overall, 207 schools were identified as not receiving (or not recently receiving) school feeding in the Kasungu district in 2023.

24. Out of the 207 schools with no current or recent school feeding experience in recent years, 108 schools were identified to satisfy the eligibility criteria for the home-grown school feeding intervention defined by the WFP Malawi CO. The criteria included schools having functioning committees (parents-teachers association, school management, financial, and procurement committees), having a water source and woodlot on site, and being within a 20 km of distance from an FO. Out of these 108 schools, 88 were selected for inclusion in the impact evaluation. Out of the 88 included schools, 44 are randomly assigned to receive school meals (intervention schools), and the remaining 44 serve as a comparison group.

25. The remaining eligible schools, outside the impact evaluation, are considered priority schools in the event additional funding is made available to increase school feeding activities in the Kasungu district. Three schools also started in 2023 outside the school impact evaluation design.

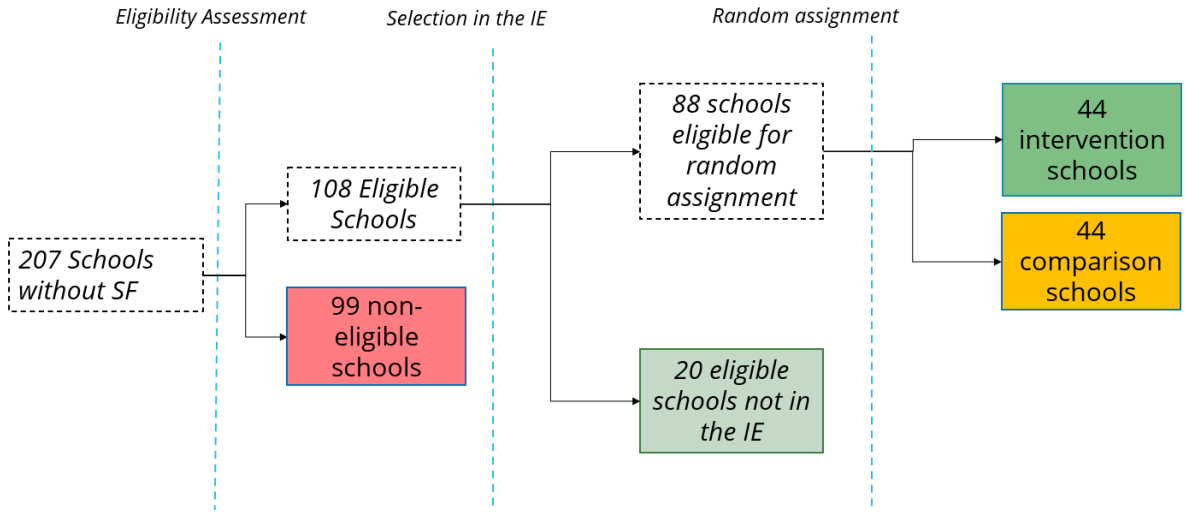

Figure 5: Selection process intervention and comparison schools

Question A: To what extent is the HGSF programme able to deliver school meals as expected?

26. The first question for this impact evaluation is to assess whether the HGSF programme can deliver school meals as expected. Assessing school meal distribution is defined by the quantity, quality, and diversity of meals distributed.

27. The evaluation will answer this question by estimating the following equation:

$$y_i = \beta_0 + \beta_1 SF_i + \varepsilon_i,$$

28. where y_i represents the outcome of school *i* (e.g., feeding days, quantity of meals distributed, school-level enrolment and attendance); SF_i is a dummy variable indicating whether the school is an intervention school (dummy equals 1) or a comparison (dummy equals 0); and ε_i is the error term. Standard errors will not be clustered since the school is the unit of randomization. In this equation, as in the equations below, β_1 is the parameter of interest, capturing the impact of the SF programme.

Question B: What is the impact of HGSF on children's outcomes (e.g., education, nutrition, and health outcomes)?

29. The main econometric analysis to answer this question will consist of linear regression to estimate the intent-to-treat effect (ITT). The impact on school performance is estimated by the following equation:

$$y_{is} = \alpha + \beta_1 SF_s + \sum X_{is} + \varepsilon_{sf}$$

30. where y_{is} indicates the outcome of for student *i* in school *s* at follow-up; α is a constant; SF_s is a variable equal to 1 if the individual is in a school providing school feeding, and 0 otherwise; X_{is} is a vector of controls; $\boldsymbol{\varepsilon_{sf}}$ is the error term. $\boldsymbol{\beta_1}$is the coefficient of interest, and it measures the impact of being enrolled in a school providing school feeding.

31. While there is already strong evidence that school feeding impacts children's attendance, this question will give an indication of whether there might be spillover effects across schools.

32. The differential effects for boys versus girls are analysed by interacting the SF dummy with the child's gender as below:

$$y_{is} = \alpha + \beta_1 SF_s + \beta_2 Female_s + \beta_3 SF * Female_s + \sum X_{is} + \varepsilon_{sf}$$

33. where y_{is} indicates the outcome of student *i* in school *s* at follow-up; α is a constant; SF_s is a variable equal to 1 if the individual is in a school providing school feeding, and 0 otherwise; $Male_s$ is a variable equal to 1 if the individual is male, and 0 otherwise; X_{is} is a vector of controls; $\boldsymbol{\varepsilon_{sf}}$ is the error term. $\boldsymbol{\beta_3}$ is the coefficient of interest, and it measures the differential impact of the school feeding programme on female students relative to males.

Question C: To what extent do the benefits of school feeding programmes vary throughout the year depending on seasonal fluctuations, shocks, and stressors?

34. To assess whether impacts are larger during the lean season, the evaluation estimates a fully interacted model with a vector of round dummies indicating four child survey rounds.

$$\boldsymbol{y_{isr} = \alpha + \beta_1 SF_s + \sum_{k=2}^{4} \beta_k R_k + \sum_{k=2}^{4} \beta_{4+k} R_k \times SF_s + \sum X_{is} + + \varepsilon_{sf}}$$

35. $\boldsymbol{\beta_k}$ capture the seasonality within the comparison group, while $\boldsymbol{\beta_{4+k}}$ are the coefficients of interest as they measure the differential impact of the school feeding programme based on the season.

3.2.2 FO-level evaluation design[3] (using a randomised encouragement design)

36. The second component of the evaluation will use a combination of a randomised encouragement design, focus group discussions (FGD), and a Discrete Choice Experiment (DCE).

37. The evaluation aims to help build an evidence portfolio around measuring the impacts on local economies as well as service delivery in schools and understanding possible bottlenecks in the setting up of linkages between schools and local markets. This will contribute to the literature on the ability of school feeding programmes to improve the economic well-being of local farmers through their impacts on market access and demand reliability for local producers.

Question D: What are the challenges and constraints preventing more farmers and FOs from benefitting from the income opportunities provided by the HGSF programme?

38. According to survey data collected in September 2023, 98 percent of the FOs in the Kasungu district are aware of the possibility of delivering to schools, and they would be interested in doing so. However, only 35 percent of them have submitted a bid in the past.

39. This question will use a combination of FGDs, survey data, and a DCE to explore such a discrepancy and identify strategies to increase the proportion of FOs submitting a bid to schools.

40. A DCE is a stated-choice survey method that consists of presenting hypothetical scenarios to respondents to uncover their implicit preferences regarding various trade-offs. These hypothetical scenarios (choice sets) will be designed to explore farmers' preferences around access to markets, production and sales. The DCE will be administered as part of the farmer survey to explore small-

[3] This component of the study has been pre-registered in the American Economic Association Randomized Control Trials (AEA RCT) registry at: https://www.socialscienceregistry.org/trials/12932

holder farmers' market preferences. The insights gained from FGDs and survey data will guide the design of choice sets for the DCE.

41. The evidence gained as part of this question will guide the design and implementation of the randomised encouragement design in Question E.

Question E: What is the impact of HGSF on the local economy (e.g. farmers and FOs in Kasungu)?

42. To assess the impacts of the HGSF programme on FOs and their members, the evaluation will rely on a randomised encouragement design intended to increase the likelihood for FOs to access market opportunities created by the HGSF programme.

43. While schools are transparent in advertising the procurement call on message boards, the short procurement timeframe and limited trust in the procurement system mean that only a few FOs submit bids to the schools' tender process. In the schools visited during the inception phase, only two cooperatives placed an offer during the previous tender process. In addition, schools lack resources to run effective advertisement campaigns, which are often limited to signs affixed on the school buildings' walls.

44. As part of this evaluation, a set of training activities, aiming to increase knowledge of existing market opportunities and build farmers' technical capacities, will be delivered to a group of randomly selected eligible FOs. Activities will include information campaign messages, capacity building to strengthen bid planning, bid writing, and bid submission and support for meeting the criteria applied by schools to select supplier FOs. If these activities create a large enough difference in participation and likelihood of being awarded a school contract between encouraged and not-encouraged FOs, this will enable the evaluation to assess the impact of HGSF on the production, sales and income of FAs and their farmer members.

45. 274 FOs have been identified as eligible to take part in the school tendering process in the Kasungu district. The evaluation will randomly encourage half of them (137 FOs) to participate in the school tendering process by providing a set of additional activities expected to increase the likelihood for the FO to succeed in the bid. The remaining 137 FOs will still be eligible to submit bids to schools. The experimental variation in the additional activities will be used to assess the impact on cooperatives and farmers.

46. As a first step, the evaluation will check if the additional training programme was successful in supporting FOs in submitting successful bids, by testing:

$$S_{ist} = \beta_0 + \beta_1 TR_i + \delta_i + \varepsilon_i,$$

47. Where S_{ist} represents whether the FO *i* was successful in a bid for school *s* during the procurement process *t*. TR_i is a dummy variable indicating whether the FO was offered the encouragement activities; δ_i are strata fixed effects accounting for distance to the closest SF school, bank account ownership, and having more than 1,000 members, which are strong predictors of FO participation in the school tendering process, and ε_i is the error term. In this equation, as in the equations below, β_1 is the parameter of interest, capturing the impact of the additional activities on the likelihood of winning a bid.

48. If the encouragement design is successful in increasing the likelihood of winning activities, the evaluation will adopt the following two estimation strategies.

49. Firstly, the evaluation will assess the impact on FOs and their members by estimating the ITT effect, estimating the following equation:

$$y_i = \beta_0 + \beta_1 TR_i + \sum X_{is} + \delta_i + \varepsilon_i,$$

50. Where y_i represents the outcome of FO or farmer *i* (e.g., revenues, sales volume, income, production, etc); TR_i is a dummy variable indicating whether the FO was invited to participate in the additional activities, or the farmer is a member of an FO invited to participate. Since TR was randomly assigned, β_1 has a causal interpretation of providing the encouragement.

51. Secondly, the evaluation will use the experimental variation of the additional activities as instrumental variable to assess the impact on FOs and their members using a local average treatment effect (LATE) effect, estimating the following equation:

$$y_i = \beta_0 + \beta_1 \hat{S}_i + \sum X_{is} + \delta_i + \varepsilon_i,$$

52. Where the parameters are defined as above, and $\hat{S}_i$ is derived as follows:

$$\hat{S}_i = \gamma_0 + \gamma_1 TR_i + \delta_i + u_i$$

53. Since the additional activities are randomly assigned, β_1 estimates the impact of being awarded a bid by using (TR_i) as an instrumental variable.

54. On the other hand, if the encouragement design does not appear to be successful in increasing the FOs' likelihood of winning bids in the first year of implementation, then the evaluation will adopt the following evaluation strategies.

55. Firstly, the preferred estimation strategy will be based on a regression discontinuity design (RDD), which will compare FOs winning a bid with those FOs who barely lost the bid. To ensure greater statistical power, the evaluation will expand the encouragement activities to all 274 FOs.

56. Secondly, depending on data availability, the evaluation will explore making use of satellite data and estimate the impact on small-holder farmers by estimating the following equation:

$$y_i = \beta_0 + \beta_1 S_i + \delta_i + \varepsilon_i,$$

57. where y_i represents outcome of farmer *i* (e.g., income, production, etc); S_i represents whether the FO *i* was successful in at least one bidding process in the last 12 months. δ_i are time-invariant controls and ε_i is the error term. In the previous equation, β_1 is the parameter of interest, capturing the impact of belonging to an FO who won a bid. As S is likely to be endogenous, the evaluation will make use of satellite data to test for parallel trends (if data is available).

3.2.3 Optimization design (Using A/B testing)

58. The third component of this evaluation will aim to combine all 132 schools where the HGSF programme is implemented in Kasungu and compare different delivery models to maximise service delivery.

59. The current model relies on parents volunteering to cook for the children on a given school day, using a rotation system in which parents from different communities are in charge of cooking on set dates. As meals are expected to be served at 6.30am, volunteers often have to wake up as early as 1am to ensure meals are ready before school begins. There is a concern that this might lead to a high number of meals not being delivered due to volunteers not being able to show up on time to prepare the meals. Moreover, as most of the volunteers tend to be women, this also poses a gender issue on potential negative effects among women volunteering to cook meals and disproportionally being exposed to risks associated with working at such early hours.

60. A previous WFP decentralized evaluation[4] of the School Meals Programme in Malawi funded by USDA found that stakeholders agreed that there was a need for more studies investigating the consequences of early feeding time.

61. The evaluation will engage with key stakeholders to explore the agreement to randomly select 66 schools out of 132 and invite them to provide the meals at mid-morning or lunch time, instead of 6.30 am. Strategies will be explored to ensure that school meals are not distributed during class, which might negatively affect students' learning. This component will seek agreement with all the stakeholders involved, particularly the Malawi National Government and the Malawi Kasungu District Council. Alternative delivery models will be explored if it is not feasible to randomly vary the timing of the meal.

[4] Decentralized Evaluation: Evaluation of the School Meals Programme in Malawi with financial support from United States Department of Agriculture (USDA) 2016 to 2018.

Question F: To what extent do different delivery models maximise the frequency of meal distribution?

62. The evaluation will aim to answer this question by estimating the following equation:

$$y_i = \beta_0 + \beta_1 Time_i + \delta_i + \varepsilon_i,$$

63. where y_i represents the outcome of school *i* (e.g., feeding days); $Time_i$ is a dummy variable indicating whether the school was instructed to serve school meals at mid-morning instead of before class In this equation, β_1 is the parameter of interest, capturing the impact of changing the time of meal, compared with the status quo.

3.2.4 Cost-effectiveness analysis

Question G: What is the cost to achieve the impact of this set of interventions? What is cost-effective to achieve this outcome?

64. The fourth and final component of this evaluation will aim to perform a cost-effectiveness analysis (CEA). A CEA provides a complementary perspective, examining the relationship between costs and outcomes. It helps decision-makers prioritise interventions that deliver the greatest impact relative to their costs, combining both the magnitude of change attributable to the intervention (estimated by the impact evaluation) and the efficiency of resource utilization.

65. The evaluation will calculate the Incremental Cost-Effectiveness Ratio (ICER) for selected children's outcomes and the Benefit Cost Ratio for farmers' outcomes and compare it with other similar interventions in Malawi and other countries.

66. Dhaliwal et. Al (2012) will provide guidance on how to conduct a comparative CEA.

4. Data collection and measurement

4.1. DATA SOURCES, SAMPLING STRATEGY AND SAMPLE SIZE

67. The evaluation will collect data from multiple sources and points in time.

School component

68. Four **child surveys** will be conducted with 1,760 children randomly selected in 44 intervention schools (880 children) and 44 comparison schools (880 children). In each school 20 children will be randomly selected to take part in four waves of child surveys, two per school year. Table 1 provides the full list of indicators to cover. Two child surveys will be collected around February 2024 and February 2025, at the end of the lean season. Two child surveys will be collected around June 2024 and June 2025 at the end of the school year, outside of the lean season. This will give the opportunity to test whether the impact of school feeding is different depending on seasonality. It will also enable to capture the short- and medium-term effects on children's outcomes. No baseline will take place for children's outcomes. Instead, school-level measures and time-invariant child-level variables (such as parental education, pre-programme family size, and pre-programme attendance, if available) will be used to confirm the balance between intervention and comparison groups.

69. One round of **anthropometric** measures (height and weight) will be collected at the end of the second school year during the last round of child surveys (June 2025).

70. Three **school surveys** will be conducted in 44 intervention schools and 44 comparison schools to collect information about attendance and dropouts. A monitoring system will be set up to standardize and digitize attendance data collected in schools[5,6].

FO component

71. Three **school surveys** will be conducted in the 176 schools receiving the HGSF programme; this includes the 88 schools surveyed as part of the school component mentioned above, and 88 different schools that had started school feeding before September 2023. The school survey will collect information on procurement and school meal delivery. The school survey will take place in September 2023, June 2024, and June 2025.

72. Three **FO surveys** will be conducted among all the 274 eligible cooperatives involved in the encouragement design component. The FO survey will take place in September 2023, August 2024 and August 2025.

73. Two **farmer surveys** will be conducted with 1,644 farmers per data collection. Six farmers will be randomly selected from each of the 274 eligible FOs. Table 1 below provides the full list of indicators to cover. Farmers' surveys are expected to take place in August 2024 and August 2025, following the second round of school survey data collection.

74. The evaluation team will also explore the availability of **satellite data** and combine it with FO and farmers' GPS locations to potentially capture information about crop production.

Optimising interventions component

[5] The first school survey is expected to be conducted in September 2023 to collect attendance and school information data for the scholastic year 2022/23. Follow-up school surveys will take place on a regular basis, expected no later than June 2024 and June 2025 to collect data for the scholastic years 2023/24 and 2024/25.

[6] Given the importance of measuring intra-household reallocation as a consequence of the school feeding programme, a parent survey measuring individual-level food consumption (administered to parents of the children attending the 88 schools in the school surveys above) will be carried out if funding resources become available in the future.

75. One additional **school survey** will be conducted among the 88 schools that are already receiving school feeding in 2023. This will be conducted one year after the introduction of the changes in the timing of the menu.

Cost-effectiveness

76. The evaluation will collect costing and beneficiary number data throughout the programme to conduct a CEA.

4.2. OUTCOMES OF INTEREST

The outcome indicators are standardized across the SBP window to allow for comparison across countries. Additional measures can be proposed upon interest and will be reviewed by all stakeholders.

Table 1: Outcomes of interest

Evaluation question Evaluation component	Unit of observation	Indicators	Type data collection	Sample
i) School design	Children	**Nutrition:** Dietary diversity (Food Consumption Score and Dietary Diversity score food groups), Food insecurity (Food Insecurity Experience Scale); **Malnutrition:** Anthropometrics (stunting, thinness, underweight, overweight, and obesity); **Mental health**: Life satisfaction, stress, depression, agency; **Health status:** Number of days ill, Washington Group Short Set on Functioning (Disability)[7]; **Learning:** Reading skills (EGRA), Numeracy skills (EGMA); **Cognitive ability:** Attention span (SCWT or Day/Night Stroop), Working memory (digit span), Fluid intelligence (Raven's Progressive Matrices (shortened) or Raven's Colour Progressive matrices); **Social cohesion:** Trust, belonging, and inclusion.	Child survey	1,760 children per data collection exercise. 20 children per school in 44 intervention schools (880 children) and 44 comparison schools (880 children).
	Schools	Attendance, dropouts, grate progression, repetition, new enrolment	School survey	88 schools (44 intervention and 44 comparison)

[7] For statistical descriptive purposes only, no comparisons will be made between treatment and comparison.

ii) FOs design	Farmers	**Agricultural production:** Production, sales and revenues; **Non-agricultural production:** Income and wages, sales and revenues; **Shocks and coping strategies:** Savings and investments, Shocks and coping strategies.	Farmer survey	1,650 farmers per data collection exercise.
	FOs	FO membership **Service delivery:** Access to bank account, Training; **Bidding data:** Number of bids submitted, Crop value and quantity for all bids (successful and unsuccessful); **Income & investments:** Sales, Markets, Quantities, Share of crops sold to schools, Office and storage space, Investments (tools, machinery, land).	FO survey	All 274 eligible FOs
	Schools	**Bidding data:** Number of tenders, number of bids received, void bids, successful bids, Crop value and quantity for all bids (successful and unsuccessful), delivery; **Service delivery:** Commodity tracking, Meal distribution days (as compared to planned meal days).	School survey	132 supported schools (44 intervention schools in the school design + 88 schools previously involved in school feeding)
iii) Optimising interventions	Schools	**Enrolment and attendance** (at grade level); **Service delivery:** Commodity tracking, Meal distribution days (as compared to planned meal days).	School survey	132 supported schools

4.3. POWER CALCULATIONS

School component

77. We base our power calculations for child-level outcomes on five measures: household food consumption scores (FCS, mean 45.5 and standard deviation 17.39), household food expenditures (mean 8.46 and standard deviation 9.55), an indicator for whether the household is worried about lack of food (mean 0.61 and standard deviation 0.49), weight-for-age (mean -1.43 and standard deviation 1.64), and height-for-age (mean -0.64 and standard deviation 1.22). Data for the household FCS, food expenditure, and the indicator of being worried about lack of food are obtained from the 2019-2020 wave of the Malawi Living Standard Measurement Survey (LSMS), while the weight-for-age and height-for-age data come from the 2016 wave of the Malawi Demographic and Health Survey (DHS).

78. Assuming 44 schools for the intervention group and 44 for the comparison group, 20 children per school, one follow-up observation, and power of the test of 80 percent, yields a minimum detectable effect (MDE) of 0.19 standard deviations for FCS, food expenditures, weight-for-age, and height-for-age, and of 0.21 standard deviations for the worry-about-lack-of-food indicator.

79. With two follow-up measurements and assuming a correlation of 0.2 between follow-up measurements, the child design is powered to detect an effect of 0.14 standard deviations for FCS and food expenditures, 0.15 for weight-for-age and height-for-age, and 0.16 for the worry indicator. Increasing the correlation between follow-up rounds to 0.5 yields an MDE of 0.16 for all indicators, excluding weight-for-age, whose MDE is 0.17 standard deviations.

80. In conclusion, the experiment is likely to be powered to detect impacts on food consumption score, food expenditures, worry indicator, and anthropometrics with at least two follow-up rounds of data collection at the child level.

Table 2: Power calculation – child-level outcomes

Outcome	Intervention	Comparison	Number of children per cluster	Comparison Mean	Comparison Standard Deviation (SD)	MDE	MDE (% of mean)	MDE (in SD)
Food Consumption Score (0-112)	44	44	20	45.52	17.40	3.25	0.07	0.19
Household food expenditures (USD)	44	44	20	8.46	9.55	1.78	0.17	0.19
Worried about lack of food	44	44	20	0.61	0.49	0.10	0.14	0.21
Weight-for-age z-score	44	44	20	-1.43	1.64	0.31	-0.28	0.19
Height-for-age z-score	44	44	20	-0.64	1.22	0.23	-0.56	0.19

Note: Child-level outcome power calculations assume a baseline to one follow-up with auto correlation of 0.5 and an intra-cluster correlation of 0.05.

FO component

81. We base our power calculations for FO-level outcomes on three measures for FOs: total number of submitted bids (mean 3.17 and standard deviation 2.96), the share of successful bids (mean 0.68 and standard deviation 0.45), and crop sales in thousands of USD (mean 264.44 and standard deviation 467.66). The FO-level outcomes come from the FO survey conducted in October 2023.

82. Assuming 137 FOs, a baseline and one follow-up observation, and power of the test of 80%, yields a MDE of 0.34 standard deviations for number of bids and crop sales and 0.36 standard deviations for successful bids.

83. With three follow-up measurements and assuming a correlation of 0.2 between follow-up measurements, our experiment is powered to detect an effect of 0.23 standard deviations for the total number of bids and crop sales, and 0.25 standard deviations for the share of successful bids. Increasing the correlation between follow-up rounds to 0.5 yields an MDE of 0.28 standard deviations for the total number of bids and crop sales, and 0.29 standard deviations for the share of successful bids.

84. In conclusion, the experiment is unlikely to be able to detect impacts on number of bids, share of successful bids, and sales, mainly due to the high variability of these indicators.

85. For farmers, we also base our power calculations on three measures: crop sales (mean 35.25 and standard deviation 71.32), organic fertilizer used (mean 0.21 and standard deviation 0.41), and quantity of seeds used (mean 18.07 and standard deviation 21.84). Farmer outcomes come from the 2019 wave of the Malawi LSMS.

86. We are assuming 137 FOs, six famers per FO, one follow-up observation, and power of the test of 80%, yields a MDE of 0.15 standard deviations for crop sales and quantity of seeds and 0.17 standard deviations for organic fertilizer used. With three follow-up measurements and assuming a correlation of 0.2 between follow-up measurements, our experiment is powered to detect an effect of 0.11 standard deviations for crop sales and quantity of seeds and 0.12 standard deviations for organic fertilizer used. Increasing the correlation between follow-up rounds to 0.5 yields an MDE of 0.13 standard deviations for crop sales and quantity of seeds and 0.15 standard deviations for organic fertilizer used.

87. In conclusion, the experiment is likely to be able to detect impacts on crop sales, quantity of seeds used, and use of organic fertilizer on farmers involved in FO sales to schools. This assessment of statistical power relies on the assumption that it is possible to sample at least six farmers participating in a sale (for treatment FOs) or that would have participated in a sale if a sale had happened (for control FOs).

Table 3: Power calculation – FO and farmer-level outcomes

Outcome	Intervention	Comparison	Number of children per cluster	Comparison Mean	Comparison SD	MDE	MDE (% of mean)	MDE (in SD)
FOs								
Total number of bids submitted	137	137	-	3.17	2.96	0.82	0.26	0.28
Share of bids successful	137	137	-	0.68	0.45	0.13	0.19	0.29
Crop sales (USD thousands)	137	137	-	264.44	467.66	130.00	0.49	0.28
Farmers								
Crop sales (USD)	137	137	6	35.25	71.32	9.00	0.26	0.13
Organic fertilizer used	137	137	6	0.21	0.41	0.06	0.28	0.15
Quantity of seeds (KGs)	137	137	6	18.07	21.84	2.76	0.15	0.13

Note: FO-level outcome power calculations assume a baseline and 3 follow-ups -up with auto correlation of 0.5. Farmer-level outcome power calculations assume 3 follow-ups with an auto correlation of 0.5.

Optimizing interventions component

88. We base our power calculations for school-level outcomes for the optimization component on three measures: children attendance rate per month (mean 0.91 and standard deviation 0.11), number of school feeding days per month (mean 12.73 and standard deviation 16.6) and share of school feeding days as compared to planned per month (mean 0.70 and standard deviation 0.33). These indicators are obtained from monitoring data collected from schools and digitized by the WFP country office in Malawi.

89. Assuming 66 schools for the intervention group and 66 for the comparison group, one follow-up observation with a correlation of 0.2, and power of the test of 80%, yields a minimum detectable effect (MDE) of 0.49 standard deviations for the number of school feeding days and 0.54 for the child attendance rate. With six follow-up measurements and assuming a correlation of 0.2 between follow-up measurements, our experiment is powered to detect an effect of 0.28 standard deviations for the number of school feeding days, 0.36 standard deviations for child attendance and 0.30 standard deviations for the share of school feeding days. Increasing the correlation between follow-up rounds to 0.5 yields an MDE of 0.37 and 0.40 standard deviations for the number and share of school feeding days, respectively, and 0.45 standard deviations for child attendance.

90. In conclusion, the optimization component seems to be powered to detect only considerable changes in child attendance rate and school feeding days.

Table 4: Power calculation – School-level outcomes for the optimization component

Outcome	Intervention	Comparison	Comparison Mean	Comparison SD	MDE	MDE (% of mean)
Child attendance rate	66	66	0.91	0.11	0.05	0.45
Number of school feeding days	66	66	12.73	6.49	2.42	0.37
Share of school feeding days (as compared to planned)	66	66	0.70	0.33	0.13	0.40

Note: School-level outcome power calculations assume six follow-up surveys with an auto correlation of 0.5.

5. Ethical considerations

91. Evaluations must conform to the 2020 United Nations Evaluation Group (UNEG) ethical guidelines. Accordingly, OEV and DIME are responsible for safeguarding and ensuring ethics at all stages of the evaluation cycle. This includes but is not limited to, ensuring informed consent, protecting privacy, confidentiality, and anonymity of participants, ensuring cultural sensitivity, respecting the autonomy of participants, ensuring fair recruitment of participants (including women and socially excluded groups), and ensuring that the evaluation results in no harm to participants or their communities.
92. The evaluation will obtain ethical clearance from a recognised Institutional Review Board (IRB) before collecting survey information from schools, children, parents, farmers and farmers' organizations.
93. The following additional considerations have been made when designing the three evaluation components:
94. Children enrolled in the 88 schools enrolled in the impact evaluation are all given an equal chance to receive school meals. Only 44 schools will be selected due to budget constraints. If more funding is available to expand school feeding to more schools, priority will be given to eligible schools not involved in the impact evaluation. When this group is fully covered, the evaluation will not prevent interventions in comparison schools if additional resources are available. Planned data collection exercises will instead be anticipated before school-feeding activities begin in comparison schools.
95. All eligible cooperatives/FOs will have the opportunity to apply to the tender. The evaluation will conduct an information campaign for a randomly selected group of cooperatives/FOs. This is expected to increase the number of organisations participating in the tender process, providing greater choices to the schools. Cooperatives/FOs not exposed to the information campaign will still be able to apply to tender processes.
96. The current model relies on volunteers to cook meals throughout the night to be ready early in the morning. This brings up issues around equity, gender (the vast majority are women) and security. The proposed design aims to find possible alternative solutions to these issues.

6. Governance

6.1. EVALUATION TEAM

97. The **Impact Evaluation Team** is responsible for designing, managing, and delivering the evaluation throughout all its steps and maintaining relationships with the country office and governing bodies. The IET includes an Evaluation Manager (EM), Technical Lead (TL), research analyst (RA), and Field Coordinator (FC). In general, the EM is based at OEV, while the TL and RA are from an external partner (e.g., DIME). The IET may also include external academic partners with attempts made to identify and include local/regional academic researchers as part of the IET, in a case-by-case situation.

98. The **Evaluation Manager** is a WFP Impact Evaluation Officer. S/he is responsible for the overall implementation throughout the evaluation process and for ensuring that the evaluation responds to WFP evidence priority needs. The EM provides the first level of quality assurance. In line with the WFP Evaluation Policy 2022, the OEV EM can also play a more significant role in an evaluation, such as team leader, who is responsible for the overall technical quality of the evaluation. All IET members shall not have vested interest in the evaluand (i.e., subject under evaluation). In cases of WFP staff, they should come from an independent evaluation unit with clear and distinct career paths and career progression incentives that are different from the programme's performance.

99. The **Field Coordinator** (FC), or Field Manager in the case of large multi-year impact evaluations, is based either at the country or regional level and is responsible for liaising with programme team and implementing partner throughout the IE. Generally, the Field Coordinator is based in the country where the intervention is implemented. Arrangements on how the Field Coordinator is recruited will vary on a country-to-country base, as they can be hired by WFP's CO or the evaluation partner(s). The Field Coordinator will need to have access to the field and WFP data and information, including access to WFP systems and WFP duty of care.

Table 5: IE Team

Name	Role	Organisation/Unit
Jonas HEIRMAN	Head of Impact Evaluation Unit	World Food Programme (OEV)
Simone LOMBARDINI	Evaluation Officer	World Food Programme (OEV)
Minh Phuong LA	Monitoring and Evaluation Officer	World Food Programme (OEV)
Florence Kondylis	Research Manager	World Bank (DIME)
Benedetta Lerva	Economist	World Bank (DIME)
Astrid Zwager	Research Officer	World Bank (DIME)
Roshni Khincha	Research Analyst	World Bank (DIME)
Cox Bogaards	Research Assistant	World Bank (DIME)
Takondwa Moyo	Field coordinator	World Food Programme (Malawi)

6.2. INTERNAL REFERENCE GROUP

100. The **Internal Reference Group** (IRG) is programme- or country-specific- management group serving as a key interlocutor during the impact evaluation. The IRG is chaired by the CD (or their designee who can steer the programme implementation, e.g., DCD or Head of Programme), with the EM serving as secretary. It is composed of country office and regional bureau staff, who have a key interest in the evaluation. The IRG is responsible for the co-design of the evaluation, identifying priority questions and

feasible implementation options together with the IET. It reviews key outputs during each phase of the IE. It is expected to meet at the end of each phase and no less than once a year.

Table 6: Internal Reference Group

Name	Role	Organisation/Unit
Simon Denhere	Deputy Country Director	World Food Programme (Malawi)
Nicole CARN	Head of Programme	World Food Programme (Malawi)
Daniel SVANLUND	Head of VAM, M&E and Gender	World Food Programme (Malawi)
Martin MPHANGWE	Programme Policy Officer	World Food Programme (Malawi)
Jason NYIRENDA	Monitoring and Evaluation Officer	World Food Programme (Malawi)
Sandra KAMVAZINA	Programme Associate	World Food Programme (Malawi)
Madalo THOMBOZI	Senior Programme Associate	World Food Programme (Malawi)
Jean Providence NZABONIMPA	Regional Evaluation Officer	World Food Programme (RBJ)
Rosalyn FORD	Regional Programme Policy Officer	World Food Programme (RBJ)
Niamh OGRADY	Head of MEAL	World Food Programme (SBP)

6.3. WORKING GROUP

101. The Working Group (WG) is composed by representatives from the EMG and IET and is responsible for ensuring that programme intervention(s) is implemented as outlined in IE design. The WG serves as the day-to-day key interlocutor between the IET and EMG during the impact evaluation process. It ensures that programme implementation is in line with the evaluation design. The FC or RA coordinates the WG. The working group is expected to engage regularly, depending on the phase, this can be from a weekly base to a monthly base. It is suggested that a member of the CO RAM team and relevant RB focal points also be appointed to the WG.

6.3. EXTERNAL REFERENCE GROUP

102. The **External Reference Group** (ERG) is composed of internal and external stakeholders with the aim to build ownership in the evaluation process and maximise evidence use and uptake. The ERG for this evaluation includes donor representatives and government officials at the national and district level, as described in Table 7. ERG members are expected to be consulted and kept updated on any major developments in the evaluation process.

Table 7: External Reference Group

Name	Role	Organisation/Unit
Grant Angus Dansie	Senior Advisor	Norad - Department for Climate and Cooperation
Glory Mwanyongo	Economist (Ministry of Education)	Malawi National Government - Ministry of Education
George Mtengowadula	Chief Planning Officer	Malawi Kasungu District Council

7. Timeline

103. Figure 6 provides a visual representation of the data collection processes described in Section 4. Figure 7 provides a timeline of the expected activities including school feeding, biding processes, and information campaigns for selected FOs. Finally, preliminary results are expected to take place 4 months from the relevant data collection round (December 2025), and draft reports are expected to take place 6 months from the last data collection round (February 2026).

104.

Figure 6: Timeline data collection

	Mar-23	Apr-23	May-23	Jun-23	Jul-23	Aug-23	Sep-23	Oct-23	Nov-23	Dec-23	Jan-24	Feb-24	Mar-24	Apr-24	May-24	Jun-24	Jul-24	Aug-24	Sep-24	Oct-24	Nov-24	Dec-24	Jan-25	Feb-25	Mar-25	Apr-25	May-25	Jun-25	Jul-25	Aug-25
Data collection																														
Child Survey (1760)																														
Anthropometrics (1760)																														
School survey (176)																														
Farmers' survey (2,200)																														
Coop survey (275)																														

Figure 7: Timeline activities

	Mar-23	Apr-23	May-23	Jun-23	Jul-23	Aug-23	Sep-23	Oct-23	Nov-23	Dec-23	Jan-24	Feb-24	Mar-24	Apr-24	May-24	Jun-24	Jul-24	Aug-24	Sep-24	Oct-24	Nov-24	Dec-24	Jan-25	Feb-25	Mar-25	Apr-25	May-25	Jun-25	Jul-25	Aug-25
Programme Implementation																														
School feeding																														
Bidding process																														
Information campaing to selected coops/FAs																														
Information reminder to selected coops/FAs																														

Acronyms

AEA	American Economic Association
CEA	Cost-effectiveness Analysis
CO	Country Office
DCE	Discrete Choice Experiment
DHS	Demographic and Health Survey
DIME	Development Impact Evaluation
EM	Evaluation Manager
EMG	Evaluation Management Group
EGRA	Early Grade Reading Assessment
EGMA	Early Grade Mathematics Assessment
FGD	Focus Group Discussion
FA	Farmer Association
FC	Field Coordinator
FCS	Food Consumption Score
FO	Farmer Organization
HGSF	Home-Grown School Feeding
ICER	Incremental Cost-Effectiveness Ratio
IE	Impact Evaluation
IET	Impact Evaluation Team
IRB	Institutional Review Board
IRG	Internal Reference Group
ITT	Intend to Treat
IV	Instrumental variable
LATE	Local average treatment effect
LSMS	Living Standard Measurement Survey

M&E	Monitoring and Evaluation
MDE	Minimum Detectable Effect
Norad	Norwegian Agency for Development Cooperation
OEV	Office Of Evaluation
RA	Research Assistant
RCT	Randomised Control Trial
RDD	Regression Discontinuity Design
SCWT	Stroop Colour and Word Test
SD	Standard Deviation
SBP	School-based Programmes
TOR	Terms of Reference
TL	Technical Lead
UNEG	United Nations Evaluation Group
USDA	United States Department of Agriculture
WB	World Bank
WFP	World Food Programme
WG	Working Group

Office of Evaluation

World Food Programme

Via Cesare Giulio Viola 68/70
00148 Rome, Italy
T +39 06 65131 wfp.org