



WFP EVALUATION



World Food Programme

SAVING
LIVES
CHANGING
LIVES

Evaluation Synthesis of WFP's Performance Measurement and Monitoring 2018–2021

Evaluation synthesis report

OEV/2021/022
Office of evaluation

January 2023

Acknowledgements

The synthesis team wish to thank all those who have contributed to this synthesis report. We would like to thank staff from the World Food Programme (WFP) – especially the Internal Reference Group, comprising staff at headquarters and regional bureau – for taking the time to feed into this process. The team would also like to extend a special thank you to Federica Zelada, Francesca Bonino, Michele Gerli, Deborah McWhinney and Anne-Claire Luzot in the Office of Evaluation for their camaraderie, guidance and support.

Disclaimer

The opinions expressed in this report are those of the synthesis team, and do not necessarily reflect those of the World Food Programme. Responsibility for the opinions expressed in this report rests solely with the authors. Publication of this document does not imply endorsement by WFP of the opinions expressed.

The designation employed and the presentation of material in maps do not imply the expression of any opinion whatsoever on the part of WFP concerning the legal or constitutional status of any country, territory or sea area, or concerning the delimitation of frontiers.

Key personnel for the synthesis

OFFICE OF EVALUATION

Anne-Claire Luzot	Deputy Director of Evaluation
Deborah McWhinney	Senior Evaluation Officer
Federica Zelada	Evaluation Synthesis Manager (December 2021 – April 2022)
Francesca Bonino	Evaluation Synthesis Manager (May 2022 – February 2023)
Michele Gerli	Evaluation Synthesis Analyst

EXTERNAL SYNTHESIS TEAM

Team Leader	Elizabeth Mason
Senior Evaluator	Liz Turner
Senior Evaluator	Achim Engelhardt
Project Manager & Data Analyst	Alexander Scales (December 2021 – August 2022)
Project Manager	Ifeanyi Okekearu (September 2022 – November 2022)
Quality Assurance	Chris Perry
Project Officer	Sarah Owen

Photocredits

Cover photo: WFP/Shehzad Noorani

Contents

Executive summary	i
1. Introduction	1
1.1 Synthesis features	1
1.2 Context	3
1.3 Methodology	6
2. Synthesis findings	10
Synthesis Question 1 – <i>Do corporate outcome, output and cross-cutting indicators allow for effective measurement of intervention achievements at the country level?</i>	10
Synthesis Question 2 – <i>To what extent have WFP's monitoring systems generated credible information? How has information generated by WFP monitoring systems been used, and by whom?</i>	12
Synthesis Question 3 – <i>To what extent has WFP's Normative Framework enabled WFP to track programme effectiveness and inform corporate performance reporting?</i>	19
Synthesis Question 4 – <i>To what extent does the evaluative evidence provide learning on the three outcomes of the WFP Corporate Monitoring Strategy?</i>	21
Synthesis Question 5 – <i>What factors contributed to or hindered implementation of performance measurement and monitoring systems? Are there particular activity areas and contexts in which WFP's approach and systems have worked better and why?</i>	24
Synthesis Question 6 – <i>To what extent is WFP's performance measurement system aligned with national monitoring systems? How has WFP pursued opportunities to strengthen national monitoring systems?</i> ..	31
Synthesis Question 7 – <i>To what extent were cross-cutting issues (gender equality and women's empowerment, protection, accountability to affected populations and the environment) reflected in monitoring practices, guidance and systems?</i>	32
3. Conclusions and recommendations	37
3.1 Conclusions	37
3.2 Recommendations	39
Annexes	
Annex I: Summary terms of reference	46
Annex II: Detailed timeline	48
Annex III: Methodology	51
Annex IV: Scoping note	64
Annex V: Analytical framework	79
Annex VI: Findings, conclusions, and recommendations matrix	86
Annex VII: List of component evaluations	90
Annex VIII: List of persons interviewed	94
Annex IX: Bibliography/evaluation library	95
Annex X: Acronyms and abbreviations	101

List of tables

Table 1: Selected list of guidance and initiatives beyond the scope of the evaluations synthesized.....	3
Table 2: Final synthesis sample by type and references used	7
Table 3: Summary of evidence related to standard operating procedures.....	20
Table 4: Evidence of adequate monitoring expertise.....	22
Table 5: Table of recommendations.....	39
Table 6: Detailed timeline	48
Table 7: Evaluations reviewed during scoping by type	51
Table 8: Final sample of evaluations selected for the synthesis by type	53
Table 9: Activity coverage in evaluations included in the synthesis	55
Table 10: Synthesis questions and sub-questions.....	56
Table 11: Data sources by synthesis question	58
Table 12: Risk matrix (high (H), medium (M), low (L))	62
Table 13: Quality assurance steps that have been applied in this synthesis	63
Table 14: Performance measurement and monitoring coverage by evaluation type.....	65
Table 15: Evaluation ranking for scoping.....	67
Table 16: Indicative coverage of activities by report for evaluations included in the final sample	74
Table 17: Synthesis question revisions from TOR to inception report.....	79
Table 18: Analytical framework.....	80
Table 19: Findings, conclusions, and recommendations matrix.....	86
Table 20: List of component evaluations included (53 in total).....	90
Table 21: List of persons interviewed.....	94

List of figures

Figure 1: Synthesis objectives.....	1
Figure 2: Prioritized internal and external stakeholders.....	2
Figure 3: Timeline of key contextual developments related to WFP's performance measurement and monitoring.....	5
Figure 4: Geographic distribution of evaluations included in the synthesis.....	7
Figure 5: Distribution of synthesis evaluations by year	8
Figure 6: Challenges to credibility of monitoring data	12
Figure 7: Number of evaluation recommendations related to monitoring (n=53).....	15
Figure 8: Coverage of monitoring topics in evaluation recommendations	15
Figure 9: Evidence of use of monitoring data in evaluations	17
Figure 10: WFP's Normative Framework for monitoring	19
Figure 11: The eight standard process steps of the CSP monitoring cycle	20
Figure 12: Defined corporate monitoring strategy outcomes.....	21
Figure 13: Factors that contributed to or hindered monitoring.....	24
Figure 14: Monitoring data systems and processes in Democratic Republic of the Congo in 2020.....	27
Figure 15: Evaluations included by type.....	53
Figure 16: Geographic distribution of synthesis evaluations	54
Figure 17: Distribution of synthesis evaluations by year	54

Executive summary

INTRODUCTION

1.1 Synthesis features

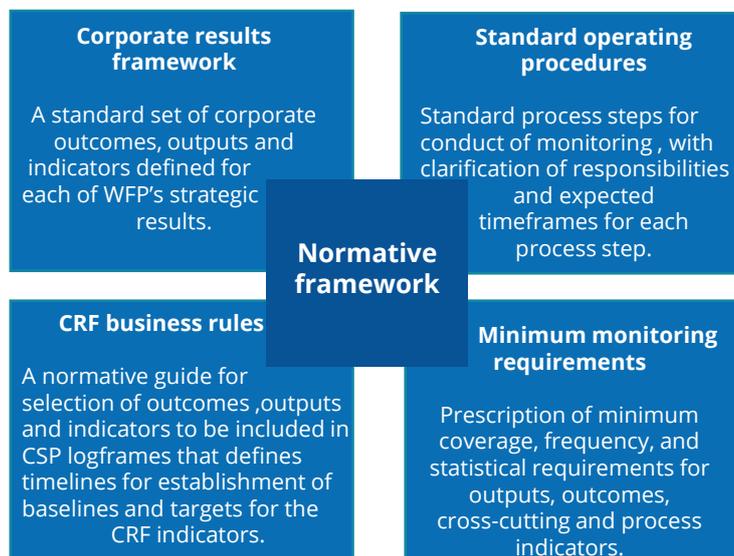
1. The synthesis of evidence and learning from evaluations of WFP's performance measurement and monitoring between 2018 and 2021 was included in the WFP Office of Evaluation workplan for 2021–2023 and was conducted by an external team working between November 2021 and October 2022, with planned submission to the Executive Board for consideration at its first regular session in February 2023.
2. Evaluation syntheses entail the combination and integration of findings from quality-assessed evaluations aimed at developing higher-level or more comprehensive knowledge and informing policy and strategic decisions. The purpose of this synthesis is to contribute to WFP's global and regional evidence base and support key corporate decision making in the short and medium terms. The specific objectives include:
 - identifying recurrent findings and stimulating discussion of performance measurement and monitoring with a view to deriving lessons on WFP's achievements and contributing to evidence-based, strategic and operational decision making; and
 - providing evidence and insights on the credibility, relevance and use of monitoring data and systems in order to inform technical and normative improvements.
3. For the purpose of this synthesis, the term "system" is understood in the broad sense of the entire "ecosystem" surrounding monitoring at both the corporate and country levels. The term "information" is also interpreted broadly, as evaluation reports often refer to "data", "information" and "evidence" interchangeably. "Credibility" is used to refer to monitoring data that were identified as being of high quality, reliable and/or consistent in the evaluation reports reviewed.
4. The intended users of the synthesis include primarily WFP's Corporate Planning and Performance Division (CPP) and Research, Assessment and Monitoring Division (RAM), but also programme and policy owners, regional bureaux and country offices.
5. The synthesis asked seven questions that examine the extent to which:
 - 1) corporate indicators allowed the effective measurement of intervention achievements at the country level;
 - 2) WFP's monitoring systems generated credible information that has been used;
 - 3) WFP's normative framework for monitoring enabled the tracking of programme effectiveness and the informing of corporate performance reporting;
 - 4) evidence from the evaluations provided learning on the outcome of WFP's corporate monitoring strategy;
 - 5) specific factors contributed to or hindered the implementation of performance measurement and monitoring systems;
 - 6) WFP's performance measurement system aligns with national monitoring systems; and
 - 7) cross-cutting priorities are reflected in monitoring practices.
6. The synthesis team examined 53 centralized evaluations (CEs) and decentralized evaluations (DEs) completed between 2018 and 2021. Following a comprehensive document review and internal consultation with selected stakeholders, the team considered some of the most recent or ongoing changes in WFP in order to help target and better situate the conclusions and recommendations.

1.2 Context

7. There is growing demand for evidence generation across WFP, and evaluation syntheses are part of the WFP “toolkit” for supporting evidence-based decision making and responding to growing interest in and demand for succinct and actionable analysis.

8. Performance measurement and monitoring are guided by the WFP normative framework for monitoring, first established in the WFP corporate monitoring strategy for 2015–2017¹ and updated for the 2018–2021 strategy.² The original framework included four components (figure 1): the corporate results framework (CRF),³ the CRF business rules,⁴ standard operating procedures for country strategic plan (CSP) monitoring⁵ and minimum monitoring requirements.⁶ Some of those components have been updated in subsequent years (see figure 2) in response to a range of emerging issues and demands, including those related to organizational restructuring – such as the 2016 launch of the Integrated Road Map and alignment aimed at contributing to the 2030 Agenda, and the 2019 establishment of RAM, which bring together WFP’s field monitoring and vulnerability assessment and mapping functions – or to the need to update and expand (in 2018 and 2022) WFP’s corporate indicators in order to better capture new priorities and areas of focus.

Figure 1: WFP's normative framework for monitoring



Source: [WFP Corporate Monitoring Strategy 2018–2021](#).

9. Performance measurement and monitoring are dynamic functions in WFP where continuous optimization is sought. Although the evaluations reviewed in the synthesis were finalized between 2018 and 2021, the corporate context and frameworks relating to performance measurement and monitoring have continued to evolve. While it is beyond the scope of this synthesis to assess all the actions taken and the newest developments, the synthesis team did review more recent documentation (see selected list in annex II) and consulted stakeholders in order to develop an understanding of the current circumstances in which the synthesis recommendations could be situated.

¹ WFP. 2015. *Corporate Monitoring Strategy 2015–2017*.

² WFP. 2018. *Corporate Monitoring Strategy 2018–2021*.

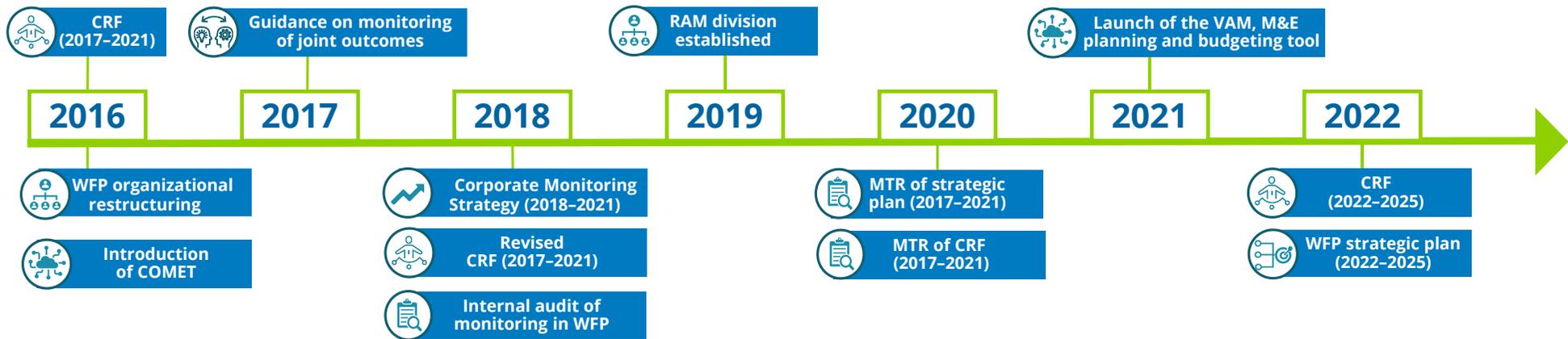
³ “[Corporate Results Framework \(2017–2021\)](#)” (WFP/EB.2/2016/4-B/1/Rev.1).

⁴ WFP. 2019. *Logframe Business Rules*.

⁵ WFP. 2017. *Standard Operating Procedures For CSP Monitoring*.

⁶ WFP. 2020. *Minimum Monitoring Requirements*.

Figure 2: Timeline of key contextual developments related to WFP's performance measurement and monitoring



Source: Evaluation synthesis team.

Abbreviations: COMET = country office tool for managing effectively; MTR = mid-term review; M&E = monitoring and evaluation; VAM = vulnerability analysis and mapping.

1.3 Methodology

10. The primary data for the synthesis came from the reports on 21 CEs⁷ and 32 DEs⁸ issued between 2018 and 2021 (table 1 and annex I). Inception reports for country strategic plan evaluations (CSPEs) and the related management responses have also been considered. All evaluations met the quality threshold of 60 percent (satisfactory) in the Office of Evaluation's outsourced independent post hoc quality assessment system.

TABLE 1: FINAL SYNTHESIS SAMPLE BY TYPE AND REFERENCES USED						
Centralized evaluations			Decentralized evaluations			Total
Country strategic plan	Policy	Strategic	Activity	Thematic	Transfer modality	53 (21 CEs) (32 DEs)
12	3	6	27	3	2	
References and abbreviations used in the synthesis: <ul style="list-style-type: none"> Country strategic plan evaluations and related inception reports – [year] [country] CSPE – e.g., 2020 Indonesia CSPE. Policy evaluations – [year] [theme] PE – e.g., 2020 gender PE. Strategy evaluations – [year] [theme] SE – e.g., 2020 resilience SE. Decentralized evaluations – [year] [country] DE – e.g., 2020 Lebanon DE. 						

11. The synthesis team used an analytical framework and coding structure to guide data extraction using the MAXQDA qualitative data analysis tool to retrieve, transcribe and visualize data. Desk analysis, interviews and a dedicated workshop were conducted with key stakeholders to discuss and validate the emerging findings and, crucially, to provide context for the changes that occurred after the evaluations were completed.

12. The synthesis was affected by great variability across the sample in terms of the availability and depth of evidence related to some of the evaluation questions and themes. As a mitigation measure, desk analysis and key informant interviews were used to supplement the information extracted from evaluation reports. To ensure the utility of the synthesis, the recommendations are based on the current corporate framework and systems, while drawing on findings from evaluations completed in the past. In other words, the synthesis reflects current circumstances in the framing of the recommendations, drawing from the desk analysis of secondary sources and inputs from stakeholders, including those shared during a stakeholder workshop focused on discussing the emerging results from the synthesis.

⁷ CEs are commissioned and managed by the Office of Evaluation and presented to the Executive Board for consideration. Evaluations ongoing at the time of the synthesis are out of the scope of this report.

⁸ DEs are commissioned and managed by country offices, regional bureaux or headquarters-based divisions other than the Office of Evaluation. DE reports are not presented to the Board.

EVALUATION SYNTHESIS FINDINGS

Do corporate outcome, output and cross-cutting indicators allow the effective measurement of intervention achievements at the country level?

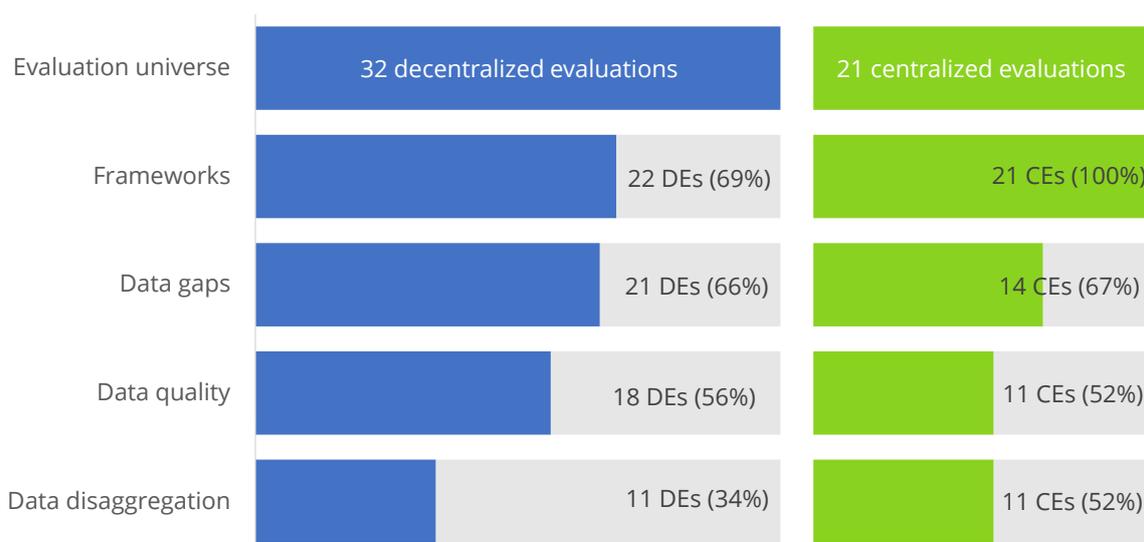
13. While corporate outcome and output indicators allow WFP to aggregate data at the corporate level, they often fall short of enabling country offices to effectively measure and report on the full depth of intervention achievements at the country level over time.

14. In the evaluation sample, all the policy evaluations (PEs) and strategic evaluations (SEs), 67 percent of CSPEs, and 9 percent of the DEs found that corporate indicators were not effective in fully measuring intervention achievements at the country level, mainly owing to being inappropriate for the particular context and to the absence of corporate indicators for certain areas and the changes made to some corporate indicators, which decreased their effective measurement of achievements over time, as highlighted in 50 percent of the CSPEs.

To what extent have WFP’s monitoring systems generated credible information? How has that information been used, and by whom?

15. Evaluations tended to address the credibility of monitoring data only when evaluation teams found a shortcoming or challenges, particularly in relation to monitoring frameworks, data gaps, data quality and data disaggregation (figure 3).

Figure 3: Challenges to the credibility of monitoring data



- *Monitoring frameworks* – Sixty-nine percent of DEs and all CEs raised concerns regarding aspects of the monitoring framework, which influenced the credibility of the data generated; 42 percent of evaluations found poor target setting, weak assumptions and/or missing indicator definitions.

Another challenge to the attribution of achievements was the bundling of activities at the outcome level in a way that made it unclear to the evaluation teams what each of the individual activities were contributing to the outcomes. As noted in the 2020 Indonesia CSPE: “[...] the scale of WFP programming in Indonesia is quite small in comparison to the size of the country and the capacity of the Government. As such, WFP contributions are aligned to the observed changes, but there are many other actors and forces contributing to contextual changes.”⁹

⁹ WFP. 2020. *Evaluation of Indonesia WFP Country Strategic Plan 2017–2020*.

- *Data gaps* – More than 60 percent of evaluations noted gaps in data collection or reporting, such as irregular, infrequent or discontinuous collection of monitoring data, including as a result of funding and access constraints.
- *Data quality* – More than 50 percent of evaluations stressed concerns regarding the quality of monitoring data. Examples included insufficient sample sizes for baselines, double counting of beneficiaries, the use of a single indicator for household and community measures, inconsistencies in the reporting of data among activities or countries, and discrepancies among activity and monitoring reports, COMET and other monitoring and evaluation systems.
- *Data disaggregation* – More than 30 percent of DEs and half of the CEs noted insufficient disaggregation of data by sex, status (such as refugee versus host country national), disability or age, as discussed further in paragraphs 27–30.

16. More than 90 percent of all evaluations recommended improvements to monitoring systems or practice, mostly focused on improving monitoring frameworks, addressing data gaps, data quality, disaggregation and data use. Overall, in respect of monitoring systems, the evaluations reviewed tended to focus more on identifying and explaining the reasons for shortcomings and the areas for improvement, rather than documenting good practices. However, some positive examples are provided.

- 2018 Türkiye DE:¹⁰ Monitoring mechanisms performed strongly, which underpinned the ability of the programme team to learn and adjust interventions. The evaluation attributed the success of the monitoring system to the establishment of a monitoring and evaluation plan and the presence of detailed standard operating procedures.
- 2020 Burkina Faso DE:¹¹ The evaluation noted improvement in the quality of sex-disaggregated data.
- 2021 Libya DE:¹² The evaluation found that data collection and analysis had improved: “WFP is agile in dealing with [third-party monitoring] feedback on partners and timely addresses shortcomings. As such, the monitoring and reporting system is adequate to capture and respond to operational challenges and ensure proper measures are taken in due course.”

17. In terms of use, evaluations provided examples relating to accountability, learning and improvement objectives. Monitoring data used for reporting within WFP and to donors served an accountability objective, while – to a lesser extent – their use by management to inform the adjustment of current activities and activity design and in the sharing of lessons served a learning objective.

18. More than 50 percent of evaluations documented the use of monitoring data for internal and external reporting. However, only 32 percent documented a learning use. This is also consistent with the 2018 internal audit of monitoring in WFP,¹³ which found that indicators were tracked for reporting compliance rather than for learning purposes.

19. Evaluations highlighted a need to expand qualitative data collection, analysis and reporting in order to contextualize WFP’s achievements and support WFP’s ability to learn and adapt using monitoring information.

20. More than 40 percent of the evaluations noted the need to either start new or expand existing qualitative data collection and reporting practices to better contextualize WFP’s achievements. Only 15 percent mentioned that qualitative data collection was occurring. The 2021 technology SE cautioned against the “[...] over-reliance on quantitative and remote approaches, which are not a good substitute for the richness of qualitative information and feedback collected in person.”¹⁴

¹⁰ WFP. 2018. [Evaluation of the DG ECHO funded Emergency Social Safety Net \(ESSN\) in Turkey November 2016–February 2018](#).

¹¹ WFP. 2020. [Evaluation thématique sur les questions de genre dans les interventions du PAM au Burkina Faso \(2016–2018\)](#). (Thematic evaluation on gender in WFP interventions in Burkina Faso (2016–2018)).

¹² WFP. 2021. [General Food Assistance and School Feeding Programmes, Libya 2017–2019](#).

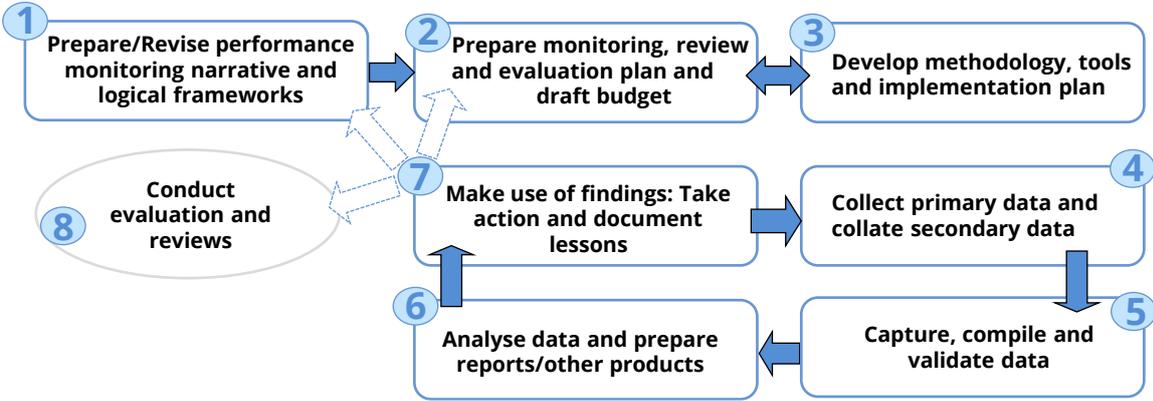
¹³ WFP. 2018. [Internal Audit of Monitoring in WFP – Office of the Inspector General, Internal Audit Report AR/18/11](#).

¹⁴ WFP. 2022. [Strategic Evaluation of WFP’s Use of Technology in Constrained Environments](#).

To what extent has WFP’s normative framework enabled WFP to track programme effectiveness and inform corporate performance reporting?

21. While some evaluations in the sample referred to the CRF (as discussed in paragraph 8), the other components of WFP’s normative framework were not directly discussed. However, some evidence relating to the eight steps of the standard operating procedures for monitoring CSPs was provided (figure 4).

Figure 4: The eight standard process steps of the country strategic plan monitoring cycle



Source: *Standard operating procedures for CSP monitoring 2017*.

22. Overall, the evaluations noted that when the standard operating procedures were followed, the monitoring systems performed well. The following are selected highlights:

- More than 75 percent of the CSPEs reviewed mentioned a logical framework or other monitoring framework for the CSP concerned. In cases where the logical framework was assessed as robust, the monitoring systems tended to perform well. Conversely, where a logical framework was assessed as weak, the monitoring systems did not perform well.
- Fewer than half of the evaluations mentioned at least some components of a monitoring, review and evaluation plan. Details were rarely included in the evaluation reports. Financial resources and budgeting for monitoring were also rarely discussed, and monitoring and evaluation budgets were not directly addressed.
- About two thirds of the evaluations touched on the collection of monitoring data.
- About half of the evaluations addressed data analysis and reporting and nearly all addressed some form of use of the monitoring data.

To what extent does the evaluative evidence provide learning on the three outcomes of the WFP corporate monitoring strategy?

23. While no evaluations directly referenced the corporate monitoring strategy, elements related to the strategy’s three outcomes (figure 5) were documented in more than half of the evaluations in the sample. The following are selected highlights:

Figure 5: Corporate monitoring strategy outcomes



Source: [WFP Corporate Monitoring Strategy 2018–2021](#).

- *Monitoring staffing levels and capacity* – Thirty percent of evaluations identified shortcomings in staffing levels and capacity, with only 6 percent assessing the staffing levels allocated to the monitoring function as sufficient. Evaluations also noted concerns related to the excessive number and diversity of the responsibilities assigned to monitoring staff. Positive examples included the 2021 Libya DE, which found staffing levels to be adequate, and the 2018 Türkiye DE, which noted that, after initial delays in staffing, the monitoring and evaluation function was “strongly staffed” at the country and local office levels. Conversely, the 2020 Cameroon CSPE¹⁵ noted that understaffing in the monitoring and evaluation team, despite recruitment efforts, affected the ability to monitor CSP implementation, and the 2021 Sri Lanka DE¹⁶ cited poor handover of responsibilities, low levels of experience and lack of training among monitoring and evaluation officers as problematic.
- With regard to *staff duties*, evaluations cited fragmentation and overburden, with the 2018 Philippines DE¹⁷ noting a shift in the role of monitoring assistants to more administrative and less “hands-on” activities, and the 2018 Algeria DE¹⁸ stressing how monitoring staff were spread too thinly and were asked to cover more than one position at a time.
- *Financial commitment* – Evaluations rarely covered the financial requirements for monitoring. When references were made, they pointed to a lack of funding as a hindering factor. Only one evaluation (the 2020 Burundi DE¹⁹) noted that funding was sufficient. Key informants consulted for this synthesis echoed the concern raised in the 2018 internal audit of monitoring in WFP, which noted that the resources for monitoring were deprioritized and that there was a tendency to use funding for programming when it was not clearly set aside for monitoring.

What factors contributed to or hindered the implementation of performance measurement and monitoring systems? Are there particular activity areas and contexts in which WFP’s approach and systems have worked better, and why?

24. The evaluations highlighted five factors (figure 6) that could either contribute to or hinder performance measurement and monitoring: government engagement, the use of technology, knowledge management, donor reporting requirements, and staff and financial resources:

¹⁵ WFP. 2020. [Evaluation of Cameroon WFP Country Strategic Plan 2018–2020](#).

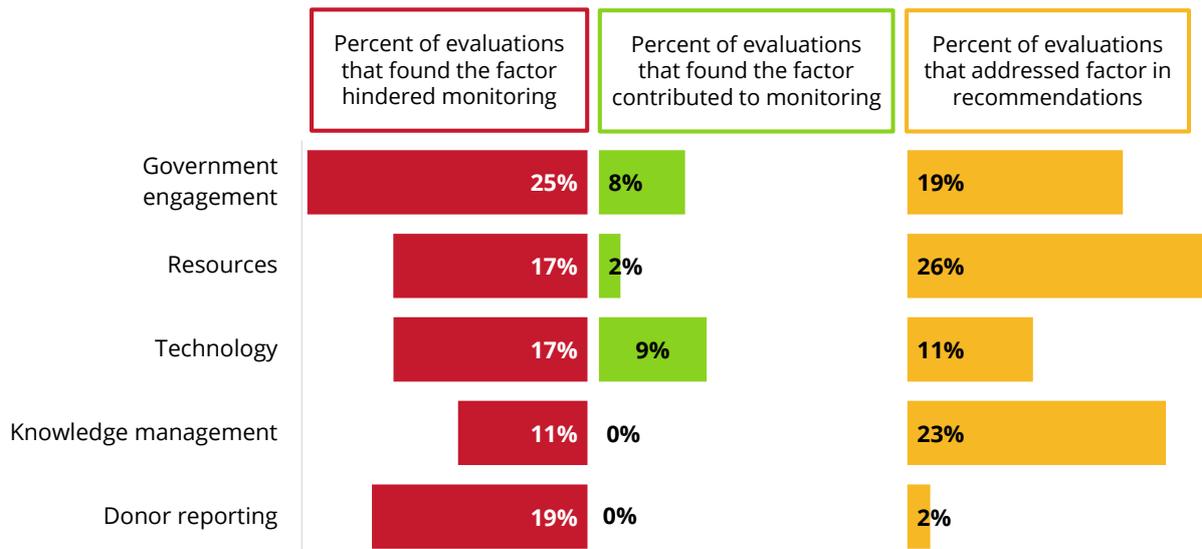
¹⁶ WFP. 2021. [Addressing Climate Change Impacts on Marginalized Agricultural Communities Living in the Mahaweli River Basin of Sri Lanka 2013–2020](#).

¹⁷ WFP. 2017. [Final Evaluation of Disaster Preparedness and Response/Climate Change Adaptation Activities under the Office of Foreign Disaster Assistance Fund in the Philippines May 2011 to September 2017](#).

¹⁸ WFP. 2018. [Evaluation of the Nutrition Components of the Algeria PRRO 200301 January 2013–December 2017](#).

¹⁹ WFP. 2020. [Evaluation of the Intervention for the Treatment of Moderate Acute Malnutrition in Ngozi, Kirundo, Cankuzo and Rutana 2016–2019](#).

Figure 6: Factors contributing or hindering monitoring



- *Government engagement* – Evaluations noted that a positive and supportive relationship with the host government contributed to the implementation of performance measurement and monitoring, while challenges with the local government hindered monitoring. Recommendations called for starting or improving joint monitoring, creating minimum monitoring requirements that covered multiple donor requests, integrating data collection practices into government systems and supporting capacity strengthening.
- *Resources* – Evaluations pointed to a lack of funding and/or staff resources as hindering monitoring and recommended making improvements by increasing staffing levels and/or enhancing the capacities of existing staff through training; improving budgeting processes; and making sufficient resources available for staffing and monitoring activities.
- *Technology* – Evaluations highlighted that effective use of technology contributed to the implementation of performance measurement and monitoring and was credited with enhancing the ease and timeliness of reporting, improving efficiency, reducing costs and increasing the scale and/or frequency of data collection. Challenges included the fragmentation of data systems and the underutilization of WFP’s mapping assets. Recommendations called for actions such as the merging of databases or sharing of data among systems, increased digitalization and enhanced visualization to encourage data use.
- *Knowledge management* – Evaluations found that poor knowledge management and/or a lack of mechanisms for sharing monitoring data within a country office resulted in missed opportunities for learning and results-based management. Recommendations called for the establishment of annual learning events, the development of dashboards that can be periodically consulted to inform analysis and decision making, the dissemination of lessons learned, the creation of feedback loops between field monitors and the CSP development process and the improvement of existing knowledge management systems for use in constrained environments.
- *Donor reporting* – While recognizing that meeting donor reporting requirements is necessary, evaluations also stressed the burden placed on country offices, particularly when the information generated for donor reporting was not used beyond that purpose. Evaluations noted a lack of harmonization among donors, but recommendations put forward very little on how to address that challenge.

25. Regarding the question as to whether there are activity areas and contexts in which WFP’s approach and systems have worked better, and why, the evaluations showed no clear pattern of links between activity categories and monitoring challenges and opportunities, other than in country capacity

strengthening and resilience building, where challenges in measurement and use of indicators were clearly mentioned.

To what extent is WFP's performance measurement system aligned with national monitoring systems? How has WFP pursued opportunities to strengthen national monitoring systems?

26. Evaluations documented WFP's overarching alignment with government priorities and plans, but rarely went into detail about the alignment of WFP's performance measurement systems with national monitoring systems and noted that where insufficient attention is given to the strengthening of national monitoring systems, efforts towards a more sustainable transition and handover of activities to national counterparts may be undermined. The 2020 school feeding SE²⁰ is one of the few evaluations that explored alignment with national monitoring systems, noting that "poor alignment of WFP monitoring and reporting systems with those of national governments is a challenge for the sustainable handover of school feeding programmes to national institutions".

To what extent were cross-cutting issues (gender equality and women's empowerment, protection, accountability to affected populations and environmental sustainability) reflected in monitoring practices, guidance and systems?

27. *Gender equality and women's empowerment* – Of the four cross-cutting issues included in the scope of the synthesis,²¹ gender equality and women's empowerment was addressed most frequently in evaluations. Evaluations found limited integration of gender equality and women's empowerment indicators into monitoring frameworks and an over-reliance on quantitative data. Evaluations also noted that shortcomings in qualitative data collection and analysis, including at the intra- and inter-household levels, have limited the ability to measure and analyse achievements in gender equality and women's empowerment, especially at the outcome level given the need to consider perceptions and other aspects of a sensitive and intangible nature such as issues relating to personal safety and security, discrimination and intra-household dynamics. Various issues were explored in the evaluations:

- Thirty-two percent of the evaluations referred to limited analysis and use of gender-related data, limiting the ability to understand how change occurs and informs potential programme adaptations. Only three evaluations discussed examples of the analysis and use of gender-related data,²² mainly to report on beneficiary numbers,²³ with limited attention beyond an assessment of whether participation is equal.²⁴ The 2020 school feeding SE characterized the approach as "hitting the target (of equal numbers)" but "missing the point". The 2020 gender PE noted some improvements in the collection and use of sex-disaggregated data in corporate reporting and in CSPEs (such as the 2020 Cameroon CSPE).
- To understand the effects of interventions on gender equality and women's empowerment, 15 percent of evaluations mentioned the importance of collecting and analysing qualitative evidence, including at the inter- and intra-household levels, and exploring household dynamics and gender-related socioeconomic and power relationships between men and women that influence access to resources and opportunities.

²⁰ WFP. 2020. [Strategic Evaluation of the Contribution of School Feeding Activities to the Achievement of the Sustainable Development Goals](#).

²¹ Disability and inclusion appeared explicitly in the corporate results framework for 2022–2025 but were not prioritized for inclusion in the present synthesis considering the timeframe of the evaluations included in the sample.

²² WFP. 2020. [Evaluation of the Gender Policy \(2015–2020\)](#); WFP. 2022. [Evaluation of Zimbabwe WFP Country Strategic Plan 2017-2021](#); and WFP. 2018. [An evaluation of the effects and a cost benefit analysis of the GFD Cash Modality scale up \(Cash-Based Transfers for PRRO 200737\) for refugees and host communities in Kenya August 2015–November 2017](#).

²³ For example: WFP. 2020. [Evaluation of the Gender Policy \(2015–2020\)](#); and WFP. 2020. [Strategic Evaluation of the Contribution of School Feeding Activities to the Achievement of the Sustainable Development Goals](#).

²⁴ For example: WFP. 2019. [Strategic Evaluation of WFP Support for Enhanced Resilience](#); and WFP. 2020. [Strategic Evaluation of WFP's Capacity to Respond to Emergencies](#).

- The 2021 technology SE found that technology is often seen as “gender-neutral”. However, gender may interact with technology, for example, where women have less access than men to mobile devices. The evaluation recommended providing gender training to technology teams and improving the gender balance of such teams, which were found to comprise mostly men. The 2021 Lebanon CSPE²⁵ reflected on recent experience with remote data collection during the 2019 coronavirus disease pandemic, raising concerns about gender equitable participation in feedback processes, given that the people responding were usually men.

28. *Protection of affected people* – There is limited evidence from evaluations on how protection of affected people is covered through monitoring practices, guidance and systems, beyond an examination of the need for improved indicators and additional data collection. However, the examples identified included the following:

- The 2021 Lebanon CSPE noted that WFP’s focus on measuring corporate indicators by using quantitative methods is not suited to capturing protection risks. Coupled with more remote data collection and “a trend in reduced interaction among WFP, cooperating partners and beneficiaries,” this has the potential to reduce the visibility of protection challenges on the ground and to limit the voice of affected communities in programme design.
- The 2021 Gambia DE²⁶ discussed the use of questionnaires on protection and accountability issues as part of monitoring efforts and how the findings led to remedial measures for overcoming the issues identified.

29. *Accountability to affected populations* – Evaluations provided limited evidence on accountability to affected populations in monitoring practices, beyond a few mentions of data availability and use. Relating to accountability to affected populations, several evaluations discussed complaints and feedback mechanisms, but only a few directly related those mechanisms to monitoring practices and adjustments made potentially as a result of the feedback received:

- The 2021 El Salvador CSPE²⁷ and the 2021 Zimbabwe CSPE²⁸ noted the availability of data on accountability to affected populations indicators. The 2021 Gambia CSPE²⁹ found good tracking of accountability indicators.
- The 2019 safety nets PE³⁰ found that in Türkiye “WFP has used evidence from monitoring and accountability to affected population mechanisms to advocate with the government for measures to promote the inclusion of households of refugees who lacked official addresses and adequate transfer values in the face of inflation.”

30. *Environmental sustainability* – Overall, evaluations made little reference to the monitoring of environmental issues, and where they did they typically referred to gaps. It is unclear whether those gaps are acceptable (because the indicators are not mandatory for the interventions) or represent deficiencies in monitoring.

CONCLUSIONS

31. The synthesis provides a snapshot in time of WFP’s performance management and monitoring, as seen through the lens of the evaluation questions and the evaluation team’s findings, conclusions and recommendations from each of the evaluations reviewed.

32. The findings identify trends and opportunities for improvement, related mainly to strengthening the use of data for learning and improving the way in which WFP captures its achievements by enhancing

²⁵ WFP. 2021. [Evaluation of Lebanon WFP Country Strategic Plan 2018–2021](#).

²⁶ WFP. 2021. [Decentralized Evaluation: Midterm Evaluation of Nutrition Activities in The Gambia 2016–2019](#).

²⁷ WFP. 2022. [Evaluación del plan estratégico para El Salvador 2017–2021](#) (Evaluation of El Salvador WFP Country Strategic Plan 2017–2021).

²⁸ WFP. 2022. [Evaluation of Zimbabwe WFP Country Strategic Plan 2017–2021](#).

²⁹ WFP. 2021. [Evaluation of The Gambia WFP Country Strategic Plan 2019–2021](#).

³⁰ WFP. 2019. [Update of WFP’s Safety Nets Policy](#).

data credibility and increasing the generation and use of qualitative data, particularly in relation to cross-cutting issues.

33. Evaluations included in the synthesis noted that there are margins for improving the use of monitoring data for programme adjustment and learning, provided that the capacities and resourcing of monitoring functions are also enhanced.

34. Overall, the three outcomes of the corporate monitoring strategy remain relevant and the normative framework continues to provide a structure that supports effective performance management and monitoring but does not encourage WFP to capture the breadth of its achievements, especially at the country level, or to track them over time. Where the framework has been adhered to, monitoring has been conducted effectively. The concern that “what gets measured matters” (and consequently gets funded) was reflected in countries where country office efforts were not well aligned with the CRF indicators.

35. While evaluations raised concerns regarding the feasibility of tracking progress over time owing to changes in the normative framework, those changes appear unavoidable in the short term as WFP moves to address many of the issues captured in this synthesis with a view to establishing a better-fitting monitoring framework for the long term.

36. WFP has a strong reporting system that draws from monitoring data. While evaluations noted that some country offices were using monitoring data to inform current or future activities, that finding was not universal. Some staff and units may be too overburdened with accountability requirements to be able to move towards learning.

37. Evaluations also provided strong evidence of the desire and need at both the country and corporate levels to expand qualitative data collection, analysis and reporting in support of learning and adaptation, pointing out that a focus on merely counting beneficiaries will result in “hitting the target” but “missing the point”.

38. Gender equality and women's empowerment is an example of an area where evaluations point to the need to strengthen the practice of disaggregated data collection and the use of qualitative data collection and analysis at the intra- and inter-household levels to enable the measurement and analysis of results, especially at the outcome level.

39. Representing relatively new themes of monitoring for WFP during the period of the evaluations, country capacity strengthening and resilience building emerged as specific activity areas in need of additional development in order to better capture and monitor WFP's achievements.

40. Regarding cross-cutting issues, monitoring of efforts towards gender equality and women's empowerment focus largely on accountability and reporting and less on learning, with evaluations noting that a lack of gender outcome data and qualitative data collection and analysis reduced the ability to measure results effectively and to understand the drivers of change. For issues such as accountability to affected populations, protection and environmental sustainability, evidence is typically limited and fragmented across evaluations, indicating either a lack of available evidence or insufficient use of existing evidence in evaluations.

RECOMMENDATIONS

41. Overall, the synthesis conclusions and recommendations complement and echo the findings from the 2018 internal audit of monitoring in WFP, especially in relation to prioritizing the use of monitoring to inform decision making and learning, investing in staff capacity and skills for monitoring and prioritizing resources for monitoring. The synthesis team has formulated the following recommendations while acknowledging that WFP is working constantly to improve its performance management and monitoring systems and that many of the weaknesses identified in the evaluations included in the synthesis have already been addressed or are the focus of ongoing efforts to identify and implement changes.

No.	Recommendation	Responsibility	Supporting entities	Priority	Deadline for completion
1	Strengthen the resourcing and use of the monitoring function as an integral component of the programme cycle in support of learning objectives.	Director RAM	Programme – Humanitarian and Development Division (PRO); country directors and country office heads of programmes	High	June 2024
1.1	Strengthen communication and advocacy, including with regional and country directors, emphasizing that an effective and adequately resourced monitoring function is an integral part of the programme cycle. Advocate the use of the monitoring function not only for accountability but also in providing fundamental support for learning and programme or operational adjustments. Efforts should consider the importance of country office-level mechanisms for reflection with decision leaders, drawing from best practices, such as regular debriefing meetings, and acting on the results of analysis from the vulnerability analysis and mapping and monitoring and evaluation planning and budgeting tool. (Strategic)	Director RAM	Regional and country directors; Programme Cycle Management Unit (PRO-M); Field Monitoring Service (RAM-M); regional monitoring advisers	High	Starting in June 2023 in line with the design, approval and implementation of second-generation country strategic plans
1.2	Consolidating and expanding on existing initiatives, take further steps to improve the visibility and use of monitoring data in the programme cycle. This may include enhancing existing templates, dashboards and guidance, and encouraging evaluations to capture lessons learned on monitoring, when appropriate. WFP should also track the use of published guidance relating to monitoring and performance measurement in order to examine whether the use of data for learning objectives has improved. (Operational)	RAM-M	Regional monitoring advisers; country office heads of programmes and monitoring and evaluation; Office of Evaluation	High	June 2024

No.	Recommendation	Responsibility	Supporting entities	Priority	Deadline for completion
2	Increase the use of qualitative data collection, analysis and reporting to better capture and enhance understanding of and learning from WFP's achievements.	RAM	CPP	High	February 2024
2.1	WFP should explore how the evidence gathered through qualitative data collection and analysis approaches – including, but not limited to, data on cross-cutting issues – can be better incorporated into corporate reporting and can better complement evidence gathered through more quantitative approaches, and should gather and share examples of instances where this is effectively achieved. (Strategic)	RAM	CPP; convenors and relevant members of the qualitative evidence generation task force	High	February 2024
2.2	Programme and policy monitoring and evaluation leads should build on current efforts in results measurement, including through qualitative evidence generation approaches, to facilitate learning at the country level. Such efforts should build on the tools and guidance that have been developed by the Research, Assessment and Monitoring Division. High priority areas include gender, country capacity strengthening and resilience building. (Operational)	Monitoring and evaluation leads in the Programme and Policy Development Department (PD); convenors of the qualitative evidence generation task force	RAM-M, CPP	Medium	Starting in June 2023, in line with the design, approval and implementation of second-generation country strategic plans
3	Provide enhanced support for improving country office monitoring systems based on the main threats to credibility identified in the evaluation synthesis.	RAM	Country office heads of programmes and monitoring and evaluation; PD monitoring and evaluation leads	High	January 2024
3.1	<i>Frameworks:</i> Regional bureaux should work with country offices to ensure that the indicators in the monitoring, review and evaluation plan are selected based on the logical framework and are relevant for measuring programme objectives and that the plan is implemented. This may include providing technical support on indicators, assumptions and targets, or the development of additional resource documents and training. (Operational)	Regional monitoring advisers	Country office heads of programmes and monitoring and evaluation	High	Starting in June 2023, in line with the design, approval and implementation of second-generation country strategic plans

No.	Recommendation	Responsibility	Supporting entities	Priority	Deadline for completion
3.2	<i>Data gaps:</i> WFP should document the use, and distil the learning from implementation, of existing guidance on addressing the gaps in the monitoring data used for setting baselines and the gaps in routine monitoring data collection activities. (Operational)	RAM-M	Country office heads of programmes and monitoring and evaluation; country office annual country report focal points	High	January 2024
3.3	<i>Data quality:</i> WFP should document the use and distil the learning from implementation of existing guidance on data quality issues such as data consistency among countries and interventions, the frequency of data collection and the double counting of beneficiaries. (Strategic)	RAM-M	Country office heads of programmes and monitoring and evaluation; COMET focal points	High	January 2024
3.4	<i>Data disaggregation:</i> WFP should take steps to document and distil the learning from the use of data disaggregation guidelines and aim to close any remaining gaps through the development of additional guidance or training. This may include monitoring the implementation of the guidance in the corporate results framework for 2022–2025 and the associated indicator compendium, minimum reporting requirements and the Research, Assessment and Monitoring Division guidance note on data stratification and disaggregation. (Strategic)	RAM	CPP; PD policy and programme leads; country office heads of monitoring and evaluation; country office annual country report focal points	Medium	January 2024
4	Provide enhanced support for improving country office monitoring systems based on the enabling factors identified in this synthesis.	RAM	Country Capacity Strengthening Unit (PRO-TC); regional bureau and country office heads of programmes and monitoring and evaluation; Technology Division; Innovation and Knowledge Management Division; regional monitoring advisers; PD policy and programme leads	High	January 2024

No.	Recommendation	Responsibility	Supporting entities	Priority	Deadline for completion
4.1	<i>Government collaboration:</i> Building on existing efforts, guidance and support should be made available to country offices for identifying ways to improve government relationships and build the capacity of government monitoring systems. This may include approaches such as joint monitoring or the inclusion of data collection in an existing government system. (Strategic)	RAM-M; PRO-TC	Country office heads of programmes and monitoring and evaluation; regional monitoring advisers	High	July 2024
4.2	<i>Financial and staff resources:</i> At the headquarters and regional levels, strengthen leadership's use of results and analysis from the vulnerability analysis and mapping and monitoring and evaluation planning and budgeting tool in advocating the allocation of adequate and more consistent human and financial resources to monitoring at the country office level. (Operational)	RAM-M	Regional directors; regional monitoring advisers; country directors	Medium	Starting in June 2023, in line with the design, approval and implementation of second-generation country strategic plans
4.3	<i>Technology:</i> Relevant headquarters divisions and units should continue to support digital data collection and survey platforms such as the Codebook and Survey Designer, focusing on improving data quality and timeliness. The Research, Assessment and Monitoring Division should also continue to provide support for improving inclusiveness in remote monitoring data collection. (Operational)	RAM	Technology Division; regional bureau and country office heads of programmes and monitoring and evaluation	High	January 2024
4.4	<i>Knowledge management:</i> Regional bureaux should work with country offices to develop a plan for knowledge management that incorporates monitoring data and analysis, and templates that can be adapted by country offices, with a focus on supporting the use of monitoring data in decision making. See recommendation 1. (Operational)	Regional bureau knowledge management focal points	Innovation and Knowledge Management Division; regional monitoring advisers; PD policy and programme leads	High	January 2024

ANNEX I

List of evaluations included in the synthesis

The 53 evaluations reviewed for the synthesis are tabled below.

Full title of the report	Abbreviated title [not all of the evaluations are mentioned in the synthesis]	Evaluation type	Evaluation category	Commissioner		Year
Strategic Evaluation of the Pilot Country Strategic Plans	2018 pilot CSP SE	Strategic	Centralized	OEV	OEV	2018
Strategic Evaluation of WFP Support for Enhanced Resilience	2018 resilience SE	Strategic	Centralized	OEV	OEV	2018
Strategic Evaluation of WFP's Capacity to Respond to Emergencies	2019 emergency response SE	Strategic	Centralized	OEV	OEV	2019
Strategic Evaluation of the Contribution of School Feeding Activities to the Achievement of the Sustainable Development Goals	2020 school feeding SE	Strategic	Centralized	OEV	OEV	2020
Joint Evaluation of collaboration among the United Nations Rome-Based Agencies	2021 Rome-based agency collaboration SE	Strategic	Centralized	OEV	OEV	2021
Strategic Evaluation of WFP's Use of Technology in Constrained Environments	2021 technology SE	Strategic	Centralized	OEV	OEV	2021
Update of WFP's Safety Nets Policy	2019 safety nets PE	Policy	Centralized	OEV	OEV	2019
Evaluation of the Gender Policy (2015–2020)	2020 gender PE	Policy	Centralized	OEV	OEV	2020
Evaluation of the WFP South–South and Triangular Cooperation Policy	2021 South–South and triangular cooperation PE	Policy	Centralized	OEV	OEV	2021
Evaluation of Bangladesh WFP Country Strategic Plan 2016–2019	2020 Bangladesh CSPE	CSP	Centralized	OEV	RBB	2020
Evaluation of Cameroon WFP Country Strategic Plan 2018–2020	2020 Cameroon CSPE	CSP	Centralized	OEV	RBD	2020
Evaluation of Democratic Republic of the Congo Interim Country Strategic Plan 2018–2020	2020 Democratic Republic of the Congo CSPE	CSP	Centralized	OEV	RBJ	2020

Full title of the report	Abbreviated title [not all of the evaluations are mentioned in the synthesis]	Evaluation type	Evaluation category	Commissioner		Year
Evaluation of Indonesia WFP Country Strategic Plan 2017–2020	2020 Indonesia CSPE	CSP	Centralized	OEV	RBB	2020
Evaluation of Timor-Leste WFP Country Strategic Plan 2018–2020	2020 Timor-Leste CSPE	CSP	Centralized	OEV	RBB	2020
Evaluation of China WFP Country Strategic Plan 2017–2021	2021 China CSPE	CSP	Centralized	OEV	RBB	2021
Evaluation of El Salvador WFP Country Strategic Plan 2017–2021	2021 El Salvador CPSE	CSP	Centralized	OEV	RBP	2021
Evaluation of Honduras WFP Country Strategic Plan 2018–2021	2021 Honduras CSPE	CSP	Centralized	OEV	RBP	2021
Evaluation of Lao People's Democratic Republic WFP Country Strategic Plan 2017–2021	2021 Lao People's Democratic Republic CSPE	CSP	Centralized	OEV	RBB	2021
Evaluation of Lebanon WFP Country Strategic Plan 2018–2021	2021 Lebanon CSPE	CSP	Centralized	OEV	RBC	2021
Evaluation of The Gambia WFP Country Strategic Plan 2019–2021	2021 Gambia CSPE	CSP	Centralized	OEV	RBD	2021
Evaluation of Zimbabwe WFP Country Strategic Plan 2017–2021	2021 Zimbabwe CSPE	CSP	Centralized	OEV	RBJ	2021
Global End-term Evaluation of the Joint Programme on Accelerating Progress towards the Economic Empowerment of Rural Women in Ethiopia, Guatemala, Kyrgyzstan, Liberia, Nepal, Niger and Rwanda from 2014 to 2020	2021 economic empowerment of rural women DE	Activity	Decentralized	Gender Unit	OEV	2021
Evaluation of the DG ECHO funded Emergency Social Safety Net (ESSN) in Turkey November 2016–February 2018	2018 Türkiye DE	Activity	Decentralized	Türkiye	RBC	2018
Evaluation of the National School Feeding Programme in Lesotho, in consultation with the	2018 Lesotho DE	Activity	Decentralized	Lesotho	RBJ	2018

Full title of the report	Abbreviated title [not all of the evaluations are mentioned in the synthesis]	Evaluation type	Evaluation category	Commissioner		Year
Lesotho Ministry of Education and Training						
Evaluation of the Nutrition Components of the Algeria PRRO 200301	2018 Algeria DE	Activity	Decentralized	Algeria	RBC	2018
Final Evaluation of Disaster Preparedness and Response/Climate Change Adaptation Activities under the Office of Foreign Disaster Assistance Fund in the Philippines May 2011 to September 2017	2018 Philippines DE	Activity	Decentralized	Philippines	RBB	2018
Final Evaluation of McGovern-Dole-supported School Feeding Programme in Bangladesh (FFE-388-2014/048-00) March 2015 to December 2017	2018 Bangladesh DE	Activity	Decentralized	Bangladesh	RBB	2018
Final Evaluation of the School Meals Programme in Malawi with support from United States Department of Agriculture, and the Governments of Brazil and the United Kingdom 2013 to 2015	2018 Malawi DE	Activity	Decentralized	Malawi	RBJ	2018
Final Evaluation of WFP'S USDA McGovern-Dole International Food for Education and Child Nutrition Programme's Support in Afar and Somali Regions in Ethiopia 2013-2017	2018 Ethiopia DE	Activity	Decentralized	Ethiopia	RBN	2018
Evaluation of National School Feeding Programme in Eswatini 2010-2018	2019 Eswatini DE	Activity	Decentralized	Eswatini	RBJ	2019
Mid-Term Evaluation of Integrated Risk Management and Climate Services Programme in Malawi from 2017-2019	2019 Malawi DE	Activity	Decentralized	Malawi	RBJ	2019
WFP's USDA McGovern-Dole International Food for Education and Child Nutrition Program's Support in Rwanda 2016-2020 -	2019 Rwanda DE	Activity	Decentralized	Rwanda	RBN	2019

Full title of the report	Abbreviated title [not all of the evaluations are mentioned in the synthesis]	Evaluation type	Evaluation category	Commissioner		Year
Evaluation Report: Midterm Evaluation						
Contribution des cantines scolaires aux résultats de l'éducation dans le sud de Madagascar (2015 à 2019): Une analyse de la contribution – De janvier 2015 à juin 2019	2020 Madagascar DE	Activity	Decentralized	Madagascar	RBJ	2020
Évaluation conjointe à mi-parcours du Programme National d'Alimentation Scolaire Intégré (PNASI) Août 2017–Mai 2019	2020 Benin DE	Activity	Decentralized	Benin	RBD	2020
Evaluation of Namibia National School Feeding Programme 2012–2018	2020 Namibia DE	Activity	Decentralized	Namibia	RBJ	2020
Evaluation of the Joint Programme for Girls Education (JPGE) with financial support from the Norwegian Government July 2014–October 2017	2020 Malawi DE	Activity	Decentralized	Malawi	RBJ	2020
Evaluation Series on Emergency School Feeding in the Democratic Republic of Congo, Lebanon, Niger and Syria 2015–2019 – Niger Evaluation Report	2020 Niger DE	Activity	Decentralized	Niger	RBD	2020
Evaluation Series on Emergency School Feeding in the Democratic Republic of Congo, Lebanon, Niger and Syria 2015–2019 – Syria Evaluation Report	2020 Syrian Arab Republic DE	Activity	Decentralized	Syrian Arab Republic	RBC	2020
Final Evaluation of the Programme “Accelerate Progress Towards Millennium Development Goal 1C (MDG1.C Programme)”	2020 Mozambique DE	Activity	Decentralized	Mozambique	RBJ	2020
Midterm Evaluation of McGovern-Dole Funded School Feeding Project in Guinea-Bissau (January 2016–June 2018)	2020 Guinea-Bissau DE	Activity	Decentralized	Guinea-Bissau	RBD	2020

Full title of the report	Abbreviated title [not all of the evaluations are mentioned in the synthesis]	Evaluation type	Evaluation category	Commissioner		Year
WFP Livelihoods and Resilience Activities in Lebanon 2016–2019	2020 Lebanon DE	Activity	Decentralized	Lebanon	RBC	2020
Addressing Climate Change Impacts on Marginalized Agricultural Communities Living in the Mahaweli River Basin of Sri Lanka 2013–2020	2021 Sri Lanka DE	Activity	Decentralized	Sri Lanka	RBB	2021
End line Evaluation of USDA Local Regional Procurement project in Nalae District, Luang Namtha Province in Lao PDR [FY 16-19]	2021 Lao People's Democratic Republic DE	Activity	Decentralized	Lao People's Democratic Republic	RBB	2021
Evaluación del modelo de descentralización del Programa Nacional de Alimentación Escolar (PNAE) 2016–2019	2021 Honduras DE	Activity	Decentralized	Honduras	RBP	2021
Final Evaluation of Enhanced Nutrition and Value Chains (ENVAC) Project 2016–2021	2021 Ghana DE	Activity	Decentralized	Ghana	RBD	2021
Final Evaluation of McGovern-Dole International Food for Education and Child Nutrition Program in Guinea-Bissau 2016–2019	2021 Guinea-Bissau DE	Activity	Decentralized	Guinea-Bissau	RBD	2021
General Food Assistance and School Feeding Programmes, Libya 2017–2019	2021 Libya DE	Activity	Decentralized	Libya	RBC	2021
Midterm Evaluation of Nutrition Activities in The Gambia 2016–2019	2021 Gambia DE	Activity	Decentralized	The Gambia	RBD	2021
Evaluación de género del Plan Estratégico de País de El Salvador (2017–2021)	2020 El Salvador DE	Thematic	Decentralized	El Salvador	RBP	2020
Evaluation of the Intervention for the Treatment of Moderate Acute Malnutrition in Ngozi, Kirundo, Cankuzo and Rutana 2016–2019	2020 Burundi DE	Thematic	Decentralized	Burundi	RBN	2020
Évaluation thématique sur les questions de genre dans les interventions du PAM au Burkina Faso (2016–2018)	2020 Burkina Faso DE	Thematic	Decentralized	Burkina Faso	RBD	2020

Full title of the report	Abbreviated title [not all of the evaluations are mentioned in the synthesis]	Evaluation type	Evaluation category	Commissioner		Year
An evaluation of the effects and a cost benefit analysis of the GFD Cash Modality scale up (Cash Based Transfers for PRRO 200737) for refugees and host communities in Kenya August 2015–November 2017	2018 Kenya DE	Transfer modality	Decentralized	Kenya	RBN	2018
Évaluation décentralisée de la modalité transfert monétaire utilisée dans le programme de cantines scolaires appuyé par le PAM au Sénégal	2018 Senegal DE	Transfer modality	Decentralized	Senegal	RBD	2018

Source: OEV Management Information System.

ANNEX II

List of selected policy changes, guidance and initiatives linked to monitoring and performance measurement

Year	Examples
2022	<ul style="list-style-type: none"> • Guidance on data stratification and disaggregation was published, covering the role of planning and the budget for a stratified sampling frame aimed at enabling representative collection of data disaggregated by critical dimensions, and explaining why the analysis and interpretation of disaggregated data to generate evidence are important. • The qualitative research task force was launched as a collaboration between RAM and programme units tasked with strengthening and developing methods for qualitative data collection and analysis. • Detailed resilience monitoring and measurement guidance was issued. • <i>The country capacity strengthening policy update</i> confirms the commitment to measuring progress in country capacity strengthening and developing materials related to its implementation, taking into account the workforce planning and budgetary considerations and including updated guidance and tools for embedding country capacity strengthening in CSPs and a robust monitoring framework that builds on the CRF. • Survey Designer, an online platform that facilitates standardized data collection, was launched. • A draft resilience toolkit was launched, including a five-step approach to monitoring and measuring resilience.
2021	<ul style="list-style-type: none"> • The CRF and strategic plan for 2022-2025 were released with the aim of improving monitoring systems and addressing some of the challenges identified in earlier versions of the normative framework. The CRF for 2022–2025 reflects WFP’s increased focus on disability inclusion by introducing dedicated indicators on disability, mainstreamed in all the strategic outcomes, and the disaggregation of data by disability status, where possible. • The vulnerability analysis and mapping and monitoring and evaluation planning and budgeting tool was launched to support budgeting and resource allocation for monitoring and evaluation needs.
2020	<ul style="list-style-type: none"> • Data quality guidance was published to address the need for systematic and consistent practices for ensuring that WFP monitoring systems produce high-quality data that measure the outputs, outcomes, cross-cutting priorities and processes of WFP’s programmes at the country office level. • E-learning on qualitative data was developed by RAM, along with related guidance materials and technical support.

Source: Compiled by the synthesis team.

1 Introduction

1. The Evaluation Synthesis of WFP’s Performance Measurement and Monitoring 2018–2021 was included in the WFP Office of Evaluation (OEV) Work Plan 2021–2023, presented to the Executive Board (EB) as part of the WFP Management Plan at the Second Regular session in November 2020. It was conducted by an external synthesis team contracted by Itad between November 2021 and November 2022, with planned submission to the EB for consideration in February 2023.

1.1 SYNTHESIS FEATURES

2. This synthesis aggregates and analyses evidence related to WFP performance measurement and monitoring from 53 centralized evaluations (CEs) and decentralized evaluations (DEs) conducted between 2018 and 2021.³¹ Evaluation synthesis are: “[a] combination and integration of findings from quality-assessed evaluations to develop higher-level or more comprehensive knowledge and inform policy and strategic decisions.”³²

3. While backward-looking, this synthesis also considered some of the most recent or ongoing changes in WFP frameworks and guidance, to help target and better situate the synthesis results. The overarching purpose is to contribute to WFP’s global and regional evidence base and to support corporate decision-making in the short and medium terms. Figure 1 outlines the synthesis’ specific objectives.

Figure 1: Synthesis objectives



³¹ See section 1.3 for an overview of the methodology for report selection, and Annex III for more details.

³² WFP. 2021. Evaluation Synthesis of WFP’s Performance Measurement and Monitoring 2018–2021. Terms of Reference. The definition is adapted from: Wyburn et al (2018) Understanding the Impacts of Research Synthesis: Environmental Science and Policy Journal, Volume 86, August 2018, pp 72–84.

4. Figure 2 below outlines the stakeholders to whom this synthesis is primarily addressed.

Figure 2: Prioritized internal and external stakeholders

Corporate Planning and Performance Division (CPP)	Responsible for producing key corporate planning and reporting documents to inform data-driven decision-making and providing oversight on the entire resource management process to ensure optimized fund utilization.
Research, Assessment and Monitoring Division (RAM)	Responsible for providing credible, relevant, and timely evidence forming the basis for operation design, and supporting the monitoring function in the field.
Programme and policy owners	Responsible for setting WFP strategic and operational direction and developing normative guidance.
Regional bureaux (RB) and CO staff	Primarily implementing monitoring activities, including collecting and analysing data and reporting on progress.
WFP senior management and EB members	Responsible for considering, approving (in part or full) or rejecting the recommendations of this synthesis, as well as overseeing consistent performance improvements.
External stakeholders	Such as national governments, donor government agencies, other United Nations agencies, non-governmental organizations, implementing partners and civil society organizations.

5. The synthesis addressed seven questions, which formed the basis of the analytical framework presented in detail in Annex V:

- SQ1. Do corporate outcome, output and cross-cutting indicators allow for effective measurement of intervention achievements at the country level?
- SQ2. To what extent have WFP's monitoring systems generated credible information? How has information generated by WFP monitoring systems been used, and by whom?
- SQ3. To what extent has WFP's Normative Framework enabled WFP to track programme effectiveness and inform corporate performance reporting?
- SQ4. To what extent does the evaluative evidence provide learning on the three outcomes of the WFP Corporate Monitoring Strategy?
- SQ5. What factors contributed to or hindered implementation of performance measurement and monitoring systems? Are there particular activity areas and contexts in which WFP's approach and systems have worked better and why?
- SQ6. To what extent is WFP's performance measurement system aligned with national monitoring systems? How has WFP pursued opportunities to strengthen national monitoring systems?
- SQ7. To what extent were cross-cutting priorities reflected in monitoring practices, guidance and systems (gender equality and women's empowerment [GEWE], protection, accountability to affected population, and environment)?

1.2 CONTEXT

6. There is growing demand for evidence generation across WFP, and a commitment to further systematic use of evidence to inform strategic directions, policies and programmes. Evaluation syntheses are part of the WFP 'toolkit' in support of its commitment to evidence-based decision-making and to respond to growing interest in and demand for succinct and actionable analysis.

7. Monitoring encompasses the tracking, collection and analysis of activities' achievements and the overall performance of WFP's programmes. Performance measurement informs operational decision-making, including the design of WFP's activities, and maintains an operational focus on results through the measurement of outcomes, outputs and processes for programmes worldwide.

8. Performance measurement and monitoring is currently guided by the WFP Normative Framework for Monitoring, which was first established in the 2015–2017 WFP Corporate Monitoring Strategy³³ and was updated for 2018–2021.³⁴ The original framework included four key documents: the Corporate Result Framework (CRF),³⁵ the CRF Business Rules,³⁶ the Standard Operating Procedures (SOPs) for CSP monitoring³⁷ and the Minimum Monitoring Requirements (MMRs),³⁸ some of which have been updated in subsequent years (as shown in Table 1). In response to a range of emerging issues and demands – some relating to organizational restructuring (such as in 2016 with the Integrated Road Map and measures taken to align with Agenda 2030, and in 2019 with the establishment of the Research, Assessment and Monitoring Division (RAM) to bring together WFP's field monitoring function and vulnerability assessment and mapping), and others relating to the need to update and expand the WFP's corporate indicators to better capture new priorities and areas of focus (in 2018 and 2022).

9. Performance measurement and monitoring are dynamic functions in WFP, where continual optimization is sought. Whereas the evaluations included in the synthesis were finalized between 2018 and 2021, the corporate context and frameworks relating to performance measurement and monitoring have continued to evolve. While it is beyond the synthesis' scope to document and assess all the actions taken and the latest developments, the synthesis team did review current documentation and consulted with stakeholders to develop an understanding of the current context against which the synthesis' recommendations could be situated. To illustrate some of these changes, a non-exhaustive list of tools, initiatives, policy changes, and so on is presented in Table 1.

Table 1: Selected list of guidance and initiatives beyond the scope of the evaluations synthesized

Year	Key guidance and initiatives
2022	<ul style="list-style-type: none"> • Release of <i>CRF 2022–2027</i> and <i>Strategic Plan 2022–2027</i>, which aim to improve monitoring systems and address certain challenges in relation to earlier versions of the Normative Framework identified in the evaluations covered in this synthesis sample. • Release of Guidance on Data Stratification and Disaggregation on the role of planning and budgeting for a stratified sampling frame, which enables representative, disaggregated data collection by critical disaggregation dimensions, and clarifies why analysing and interpreting disaggregated data to generate evidence matters • Launch of Qualitative Research Task Force as a collaboration between RAM and Programme units to strengthen and develop standardized methods for qualitative data collection and analysis • Release of detailed resilience monitoring and measurement guidance • <i>CCS Policy Update</i> confirms commitment to measuring country capacity strengthening (CCS) through developing materials related to its implementation, including workforce planning and

³³ WFP. 2015. Corporate Monitoring Strategy 2015–2017.

³⁴ WFP. 2018. Corporate Monitoring Strategy 2018–2021.

³⁵ "Corporate Results Framework (2017–2021)" (WFP/EB.2/2016/4-B/1/Rev.1).

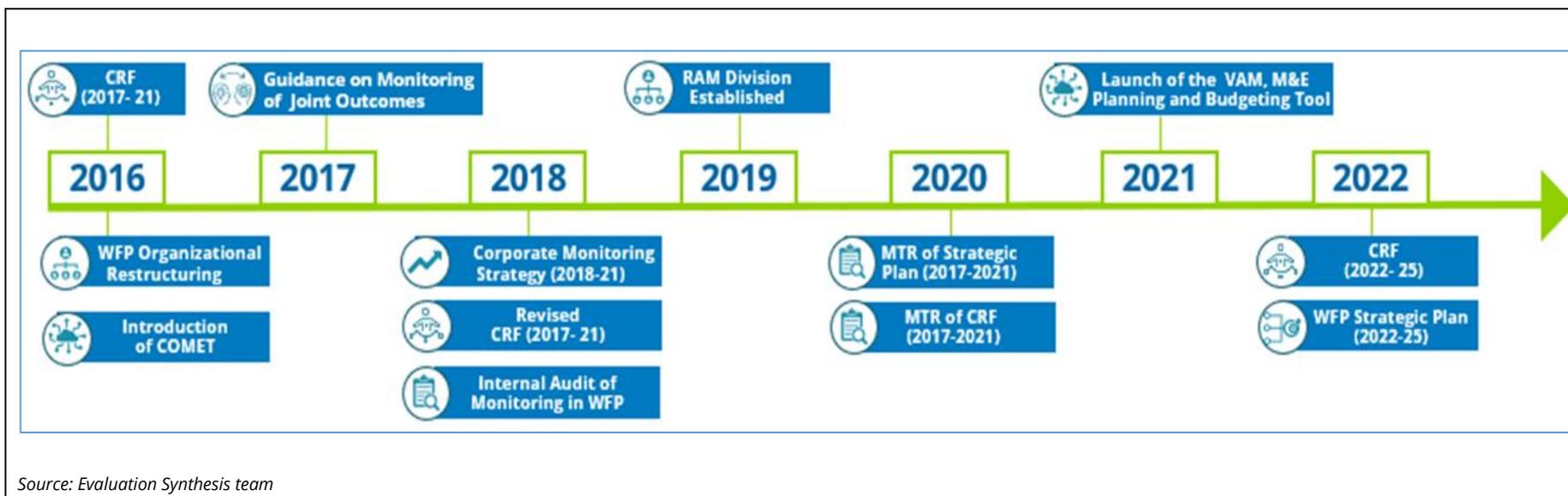
³⁶ WFP. 2017. CRF Logframe Business Rules 2017–2021.

³⁷ WFP. 2017. Standard Operating Procedures for CSP Monitoring.

³⁸ WFP. 2017. Minimum Monitoring Requirements.

Year	Key guidance and initiatives
	<p>budgetary considerations, updated guidance and tools for embedding CCS in country strategic plans and a robust monitoring framework that builds on the CRF</p> <ul style="list-style-type: none"> • Launch of Survey Designer, an online platform that facilitates standardized data collection • Launch of Draft resilience toolkit, which includes a five-step approach to monitoring and measuring resilience
2021	<ul style="list-style-type: none"> • Release of CRF 2022–2027 and Strategic Plan 2022–2027, which aim to improve monitoring systems and address certain challenges identified in earlier versions of the Normative Framework • Launch of VAM, Monitoring and Evaluation Planning and Budgeting Tool to support budgeting and resource allocation for monitoring and evaluation (M&E) needs
2020	<ul style="list-style-type: none"> • Publishing of Data quality guidance to address the need for systematic and consistent practices for ensuring that WFP monitoring systems produce high quality data that measures outputs, outcomes, cross-cutting priorities and processes of WFP’s programmes at Country Office level • Development by RAM of three e-learning on qualitative research, the Qualitative Research Facilitated Course, and additional guidance material and technical support on qualitative data

Figure 3: Timeline of key contextual developments related to WFP's performance measurement and monitoring



1.3 METHODOLOGY

1.3.1 Data sources

10. The primary data source for this synthesis are WFP evaluation reports, CSPE inception reports, and the management responses. Evidence gathered from these sources was validated through key informant interviews (KIIs), a stakeholder meeting, and a review of supplementary documents.

11. A preliminary list of 75 evaluations was identified by OEV. It included all evaluations that met the following three criteria:

- **Evaluation type:**
 - **CEs³⁹ – Policy,⁴⁰ Strategic,⁴¹ and CSP evaluations⁴²**
 - **DEs⁴³ covering activities, pilots,⁴⁴ themes, and transfer modalities⁴⁵**
- **Time period:** evaluations completed between 2018 and 2021
- **Quality of the evaluation:** evaluations assessed by OEV's post hoc quality assessment (PHQA) system⁴⁶ above the 60 per cent threshold (satisfactory)

12. During the inception phase, the synthesis team reviewed all reports meeting those criteria, and identified 53 evaluations that directly addressed performance measurement and monitoring topics, detailed by type in Table 2 and listed in Annex VII. A description of the scoping methodology and process is included in Annex IV.

³⁹ Centralized Evaluations are commissioned and managed by OEV and presented to the Executive Board for consideration. Emerging results from some of the Strategic and Policy evaluations underway at the same time as the synthesis (for example the Strategic Evaluation of Nutrition and HIV (2017-2021)) could not be considered.

⁴⁰ The Evaluation of the WFP People Strategy (2014–2017) and the Evaluation of the WFP Corporate Partnership Strategy (2014–2017) were not included in this synthesis as they were considered not relevant by OEV.

⁴¹ The Strategic Evaluation on Funding WFP's work (2020) was not included in this synthesis as it was considered not relevant for the purpose of the synthesis.

⁴² Country Portfolio Evaluations are not included in this synthesis because the operations evaluated will have been designed at minimum two years before the Integrated Road Map period, and because WFP no longer operates under that framework.

⁴³ DEs are commissioned and managed by COs, RBs or HQ Divisions other than OEV. They are not presented to the Board.

⁴⁴ No evaluations of pilots met the other two qualifying criteria for this synthesis, so these are not included in the analysis.

⁴⁵ Operation Evaluations have not been included in the synthesis, as WFP no longer operates under that framework.

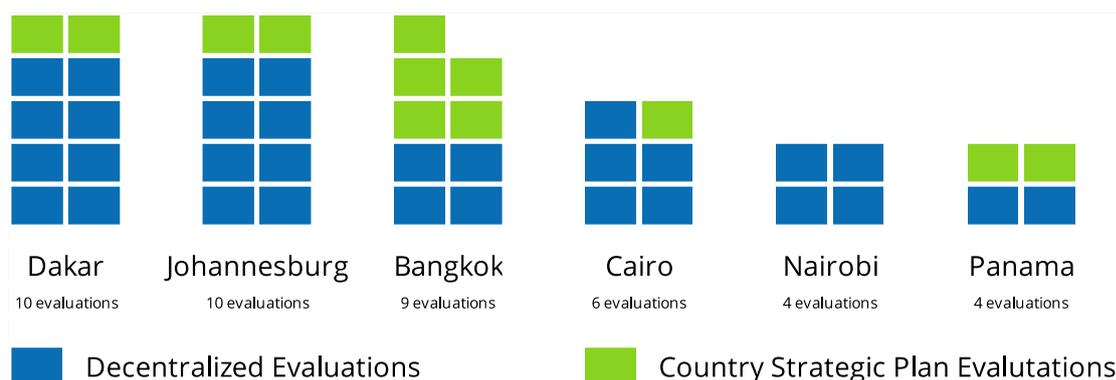
⁴⁶ Since 2016, OEV has used an outsourced PHQA mechanism, through which independent assessors rate the quality of all completed WFP evaluations against WFP's own evaluation quality standards.

Table 2: Final synthesis sample by type and references used

Centralized evaluations (CEs)			Decentralized evaluations (DEs)			Total
Country Strategic Plan	Policy	Strategic	Activity	Thematic	Transfer modality	53 21 CEs 32 DEs
12	3	6	27	3	2	
SHORTHAND REFERENCES USED:						
<ul style="list-style-type: none"> Country Strategic Plan Evaluations and related inception reports – [Country] CSPE [year] – for example Indonesia CSPE 2020 Policy evaluations – [Descriptive words] PE [year] – for example Gender PE 2020 Strategy evaluations – [Descriptive words] SE [year] – for example Resilience SE 2018 Decentralized evaluations – [Country] DE [year] – for example Lebanon DE 2020 						

13. Figure 4 below shows the geographic coverage of the 12 CSP evaluations and 31 DEs across WFP’s six regional bureaux (RBs). The highest numbers concern RB Dakar (10) and RB Johannesburg (10), followed by RB Bangkok (9), RB Cairo (6), RB Nairobi (4) and RB Panama (4).⁴⁷ Figure 5 shows the distribution of evaluations by year of completion. Fewer evaluations were completed in the first half of the sample (11 in 2018 and 5 in 2019) than in the second half (19 in 2020 and 18 in 2021).

Figure 4: Geographic distribution of evaluations included in the synthesis



Source: OEV management information system

⁴⁷ The nine centralized policy and strategic evaluations and one DE commissioned by an HQ-based division (the Global End-term Evaluation of the Joint Programme on Accelerating Progress towards the Economic Empowerment of Rural Women in Ethiopia, Guatemala, Kyrgyzstan, Liberia, Nepal, Niger and Rwanda from 2014 to 2020) involved the participation of multiple countries through desk reviews and country missions, and are therefore not included in Figure 4

Figure 5: Distribution of synthesis evaluations by year



Source: OEV management information system

1.3.2 Data collection

14. The synthesis team developed an analytical framework and coding structure based on sub-questions to guide the data extraction process. MAXQDA, qualitative analysis software, was used to extract and organize relevant information from the evaluations around the synthesis sub-questions. The coding structure included one or more codes per sub-question, as well as cross-cutting codes for topics across sub-questions. The analytical framework and codes are detailed in Annex V.

15. Data extraction included two pilot tests, a revision of the coding structure, data extraction from reports and quality assurance (QA) checks by the Team Leader. Using an iterative approach to coding and QA enabled the team to systematically build a credible and consistent evidence base against the predefined themes linked to the synthesis questions (deductive), while also allowing for emerging themes to be identified (inductive).

16. After extracting data from the evaluation reports, the synthesis team conducted a series of KIs and a desk review of supplementary documents. Interviews were conducted with the evaluation Internal Reference Group (IRG) and other relevant stakeholders, as detailed in Annex VIII. The purpose of the interviews was to validate findings from the evaluation reports and to provide context for any changes that occurred after the evaluations were completed. The synthesis team also conducted a document review of relevant WFP documents, including internal WFP policy, strategy, guidance, and audit documents.

1.3.3 Data analysis and reporting

17. The team used MAXQDA and Microsoft Excel to identify and quantify themes in the coded data across evaluation reports and CSPE inception reports. These themes were used to generate the findings, which in turn have informed the synthesis conclusions and recommendations. Data sets within the themes were analysed to identify trends and exceptions. Examples were selected to be included in the report. While the synthesis team did take representation of each evaluation into consideration for drafting this report, some evaluations addressed monitoring topics more thoroughly than others, and therefore are featured more often in the synthesis.

18. The synthesis followed the OEV QA requirements. Stakeholder inputs were sought through selected key KIs at multiple stages of the process, as well as through a broader stakeholder meeting that brought together RBs and selected HQ divisions and units, with the objective of ensuring greater accuracy in the analysis, and deeper understanding of the broader context against which the synthesis' results will be presented.

1.3.4 Limitations, risks and mitigations

19. The synthesis had several limitations related to scarcer evidence around some synthesis questions. To mitigate this, the synthesis team supplemented the evaluations with information from supporting documents and KIIs where possible. These limitations are discussed within each synthesis question if applicable. They should also be considered in light of developments that occurred after the evaluations' completion – such as the issuance of new policies, or guidance intended to tackle some of the shortcomings in the evaluations included in the synthesis.⁴⁸

20. The synthesis team also found a great variability across the sample in the depth of evidence related to the synthesis themes. Hence, some evaluation reports are referenced more than others as sources of insights and examples that speak to the synthesis questions. Annex III contains information on how the team applied proactive risk mitigation throughout the assignment, along with further detail on the methodology.

21. The main risk to the utility of this synthesis was ensuring that the recommendations would be forward looking while still reflecting the synthesis findings and conclusions drawn from evaluations completed between 2018 and 2021. While the analysis and findings focused on the evaluations themselves, the synthesis team also needed to consider that the corporate frameworks, systems and process relating to monitoring have evolved since their completion. To mitigate this risk, the synthesis team worked to reflect the current context in the framing of the recommendations, based on the key elements emerging from the desk analysis of secondary sources, as well as inputs from stakeholders.

⁴⁸ For a selected listing, see Table 1.

2 Synthesis findings

SYNTHESIS QUESTION 1 – Do corporate outcome, output and cross-cutting indicators allow for effective measurement of intervention achievements at the country level?

22. The Corporate Result Framework (CRF) is intended to be the primary tool for guiding Country Offices (COs) in monitoring and measuring whether interventions are achieving intended outputs, outcomes and cross-cutting priorities. It defines accountability for programme and management, and guides planning, monitoring and reporting at all levels of the organization, by laying out the expected results and indicators that WFP will use to monitor and report on its programmes and management performance.

23. The following findings focus on the effectiveness of corporate outcome and output indicators in general. Cross-cutting indicators are discussed in detail under synthesis question 7.

FINDING 1: While corporate outcome and output indicators allow WFP to aggregate data at corporate level, they often fall short of enabling country offices to effectively measure and report on the full depth of intervention achievements at country level.

24. Corporate indicators were referenced in all but one⁴⁹ of the CEs (95 per cent) and in 22 per cent⁵⁰ of the DEs. Of these, the [Pilot CSP SE 2018](#) and 33 per cent of the CSPEs⁵¹ found that at least some indicators were effective at measuring achievements at country level and facilitated corporate reporting.

25. The [Pilot CSP SE 2018](#) noted that there may be a risk to countries of decreased funding if they do not have appropriate indicators to show results; and the [Indonesia CSPE 2020](#) (part of the first generation of WFP CSPs) noted that a lack of corporate guidance on how to measure progress and report on key emerging areas such as CCS, led to a decreased visibility for the country office to showcase their work.

26. All of the policy and strategic evaluations, 67 per cent of the CSPEs,⁵² and 9 per cent of the DEs⁵³ found that corporate indicators were not effective at fully measuring intervention achievements at country level. This was because either the indicators were not appropriate for the contexts or no corporate indicators were available for certain areas. Major areas of concern included country capacity building and resilience, as well as gender and other cross-cutting themes (see sections 2.5.2 and 2.7).

27. Fifteen per cent of the evaluations⁵⁴ noted challenges with the ability of the corporate indicators to fully capture the achievements of WFP's interventions. For example, the [Timor-Leste CSPE 2020](#) recognized the need for standardized indicators for global reporting, but found that when all COs need to report on same indicators, it is "bound to be an incomplete representation of what each [CO] has actually achieved."⁵⁵ The [Cameroon CSPE 2020](#) noted that while its M&E framework was aligned with the CRF and included CRF indicators, it was not well adapted to the country context.

FINDING 2: Changes to some corporate indicators decreased the effective measurement of achievements at country level over time. However, these changes were made with the intent of improving WFP's ability to measure results more effectively.

28. The period of the synthesis (evaluations completed between 2018 and 2021) spanned a period of change for corporate monitoring. Some interventions started before the first CRF (2017–2021) and had to

⁴⁹ RBA collaboration SE 2021.

⁵⁰ Algeria DE 2018, Kenya DE 2018, Turkey DE 2018, Syria DE 2020, Gambia DE 2021, Ghana DE 2021, Libya DE 2021.

⁵¹ Bangladesh CSPE 2020, Timor-Leste CSPE 2020, Gambia CSPE 2021, Lebanon CSPE 2021.

⁵² Bangladesh CSPE 2020, Cameroon CSPE 2020, Indonesia CSPE 2020, Timor-Leste CSPE 2020, Gambia CSPE 2021, Laos CSPE 2021, Lebanon CSPE 2021, Zimbabwe CSPE 2021.

⁵³ Malawi DE 2019, El Salvador DE 2020, Syria DE 2020.

⁵⁴ Pilot CSP SE 2018, School Feeding SE 2020, Technology SE 2021, SSTC PE 2021, Bangladesh CSPE 2020, Indonesia CSPE 2020, Timor-Leste 2020, Laos DE 2021.

⁵⁵ Timor-Leste CSPE 2020.

bring their monitoring systems into line with the CRF. Others experienced the changes between the first CRF (2017–2021) and the revised CRF (2019). Later evaluations also referred to the then-forthcoming CRF (2022–2025).

29. Fifty per cent of the CSP evaluations⁵⁶ and one DE evaluation⁵⁷ specifically noted that changes in the CRF indicators caused challenges with measuring intervention achievements over time. For example, the original CSP logframe in Timor-Leste included nine corporate indicators, but only five could be reported on over time. One indicator was dropped and three were revised to align with the 2019 revised CRF. “As in all [COs] implementing CSPs, comprehensive monitoring was constrained until the end of 2018 by the ongoing revisions to the CRF. CRF methodology for measuring [capacity strengthening] was not available until then, and – like other COs – WFP Timor-Leste did not report on this.”⁵⁸ The Zimbabwe CO developed three consecutive logframes for the CSP in response to changes in the CRF indicators, noting that “[c]ontinuous CRF adaptations posed difficulties for data consistency, completeness and comparability across years.”⁵⁹

30. The [Syria DE 2020](#) evaluation concluded that, “M&E system tracked corporate indicators as required but due to changing requirements and approaches over the years, these indicators are not sufficient to properly assess the [emergency school feeding activity’s] efficiency/effectiveness.”⁶⁰ The evaluation went on to recommend that WFP “revise the corporate indicators and expand them with indicators that are ‘fit-for-purpose’ to the specific country contexts”.

31. Evaluations also cited issues related to the absence of corporate indicators in specific areas that prevented the measurement of progress at country level. The following evaluations documented how shortcomings in some of these areas were addressed through revisions to the CRF.

- **Social protection:** The [Safety Nets PE 2019](#) concluded that, “Some corporate systems and processes still continue to hinder work in safety nets and social protection. The strategic results frameworks and CRF in place during the period covered by this evaluation did not adequately provide a way for COs to record and monitor their contributions to safety nets or social protection. This continues to reduce the ability to systematically analyse the extent to which WFP is contributing to social protection. It may also create some disincentives if the axiom that ‘what gets measured matters’ is true.” The evaluation also noted that new indicators would be included in the 2019 revised CRF.
- **School feeding:** The [School Feeding SE 2020](#) found that “opportunities were missed to include school feeding indicators in the CRF.” The evaluation noted that while the original CRF lacked an indicator for school feeding, relevant indicators were included in the indicator compendium in 2018 and CRF indicators were added to the revised CRF in 2019.
- **South-South Triangular Cooperation (SSTC):** The [SSTC PE 2021](#) highlighted the inclusion of only one SSTC-relevant indicator in the revised CRF.⁶¹ Reporting on that indicator “tended to be fragmented and incomplete ... which may be due in part to the fact that it is still comparatively new.” The evaluation noted the progress made in the CRF 2022–2025, with three additional output indicators for institutional capacity strengthening that will facilitate reporting on SSTC.
- **Technology:** The [Technology SE 2021](#) noted that as there was no CRF indicator related to the use of technology, it was not being measured across WFP operations. However, this was addressed in the CRF for 2022–2025, with two relevant management results to leverage technology (management result 6) and leverage innovation (management result 7).

⁵⁶ Indonesia CSPE 2020, Timor-Leste CSPE 2020, China CSPE 2021, Laos CSPE 2021, Lebanon CSPE 2021, Zimbabwe CSPE 2021, Syria DE 2020.

⁵⁷ Syria DE 2020.

⁵⁸ Timor-Leste CSPE 2020.

⁵⁹ Zimbabwe CSPE 2021.

⁶⁰ Syria DE 2020.

⁶¹ Number of programmes benefiting from WFP-facilitated SSTC.

SYNTHESIS QUESTION 2 – To what extent have WFP’s monitoring systems generated credible information? How has information generated by WFP monitoring systems been used, and by whom?

32. In answering these questions, the synthesis team defined ‘system’ in the broad sense of the entire ecosystem surrounding monitoring at both corporate and country levels. The term ‘information’ is also interpreted broadly, as evaluation reports often refer to ‘data’, ‘information’ and ‘evidence’ interchangeably. ‘Credibility’ referred to monitoring data that was identified as high quality, reliable and/or consistent by the evaluation reports.

2.2.1 Credibility of information generated through WFP’s monitoring systems

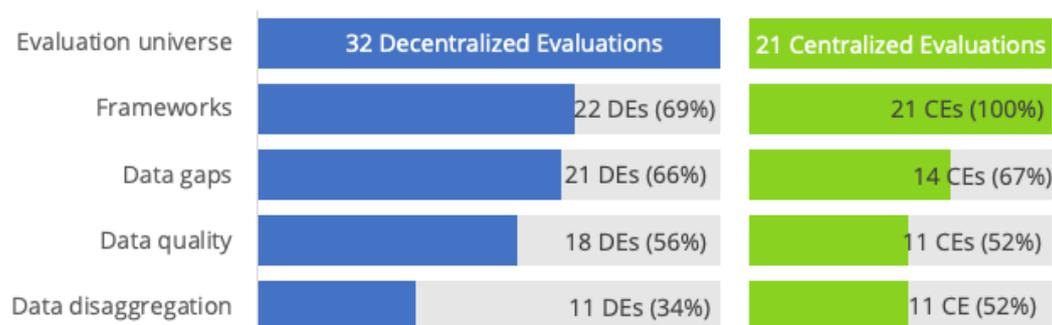
FINDING 3: Evaluations tended to only address the credibility of monitoring data if they found a shortcoming or problem. Four sets of challenges to credibility identified in the evaluations relate to monitoring frameworks, data gaps, data quality and data disaggregation.

33. In general, the credibility of monitoring data was only specifically addressed in the evaluations when the evaluation teams found a shortcoming or problem. Overall, when it comes to monitoring systems, the evaluations reviewed tend to focus their analysis more on identifying and explaining the reasons for shortcomings and areas for improvement, rather than documenting good practices. However, 38 per cent of DEs and 25 per cent of CSPEs⁶² provided some positive examples, particularly around generating credible monitoring data. For example:

- [Turkey DE 2018](#): Monitoring mechanisms performed strongly, which underpinned the ability of the programme to learn and adjust. The evaluation attributed the success of the monitoring system to the establishment of an M&E plan and having detailed SOPs.
- [Burkina Faso DE 2020](#): The evaluation noted improvement in the quality of sex-disaggregated data.
- [Libya DE 2021](#): The evaluation found that data collection and analysis had improved. “WFP is agile in dealing with [third-party monitoring] feedback on partners and timely addresses shortcomings. As such, the monitoring and reporting system is adequate to capture and respond to operational challenges and ensure proper measures are taken in due course.”

34. The synthesis identified four areas that threatened the credibility of WFP’s monitoring data. As summarized in figure 6, challenges with data gaps and data quality were identified more often in DEs than CEs. Conversely, issues with monitoring frameworks and data disaggregation were identified more often in CEs than DEs.

Figure 6: Challenges to credibility of monitoring data



Source: Synthesis team analysis

⁶² Kenya DE 2018, Turkey DE 2018, Burkina Faso DE 2020, Guinea Bissau DE 2020, Mozambique DE 2020, Namibia DE 2020, Guinea Bissau DE 2021, Libya DE 2021, Cameroon CSPE 2020, DRC CSPE 2020, Lebanon CSPE 2021.

i. Monitoring frameworks

35. Some 69 per cent of DEs and all CEs⁶³ raised concerns with aspects of the monitoring framework that influenced the credibility of the data generated; 40 per cent of evaluations⁶⁴ noted that the indicators were insufficient to track WFP's contributions, particularly in capacity strengthening and resilience (discussed further in question 5); 42 per cent of evaluations⁶⁵ found poor target setting, weak assumptions, and/or missing indicator definitions; and 15 per cent of evaluations⁶⁶ noted a challenge to measuring change over time due to changes in monitoring frameworks, and specifically revisions to the CRF (as discussed in question 1).

36. Fifteen per cent of the evaluations⁶⁷ raised concerns about attributing the results captured by the monitoring systems to WFP's interventions. One concern was whether indicators and monitoring data were measuring WFP's achievements or external factors. These concerns included indicators that were outside of WFP's control and the influence of other actors' contributions to the changes observed. Another challenge to attribution was the bundling of activities at outcome level in a way that made it unclear to the evaluation teams what each of the individual activities was contributing to the outcomes. Examples related to attribution included:

- [Indonesia CSPE 2020](#): "The scale of WFP programming in Indonesia is quite small in comparison to the size of the country and the capacity of the Government. As such, WFP contributions are aligned to the observed changes, but there are many other actors and forces contributing to contextual changes."
- [Lebanon CSPE 2021](#): The inception report noted that: "One of the performance indicators is the 'retention' rate. While school feeding has shown to have an effect on retention, factors such as livelihood opportunities of the family and support measures to address the cost of schooling – for example, transport costs donated to the school by other agencies – must also be considered." The evaluation noted that the limitation was overcome through triangulation with external data sources.

ii. Data gaps

37. Sixty-six per cent of Des and 67 per cent of CEs⁶⁸ noted concerns about gaps in data collection and/or reporting, such as irregular, infrequent or discontinuity in the collection of monitoring data. Causes of these gaps are discussed further under synthesis question 5, including challenges with funding and

⁶³ Bangladesh DE 2018, Kenya DE 2018, Lesotho DE 2018, Malawi DE 2018, Eswatini DE 2019, Malawi DE 2019, El Salvador DE 2020, Guinea Bissau DE 2020, Lebanon DE 2020, Madagascar DE 2020, Malawi DE 2020, Mali DE 2020, Mozambique DE 2020, Namibia DE 2020, Niger DE 2020, Syria DE 2020, Gambia DE 2021, Ghana DE 2021, Guinea Bissau DE 2021, Economic Empowerment of Rural Women DE 2021, Libya DE 2021, Sri Lanka DE 2021, Safety Nets PE 2019, Gender PE 2020, SSTC PE 2021, Pilot CSP SE 2018, Resilience SE 2018, Emergency Response SE 2019, School Feeding SE 2020, UN Cooperation SE 2021, Technology SE 2021, El Salvador CSPE 2021, Bangladesh CSPE 2020, Cameroon CSPE 2020, DRC CSPE 2020, Indonesia CSPE 2020, Timor-Leste CSPE 2020, China CSPE 2021, Gambia CSPE 2021, Honduras CSPE 2021, Laos CSPE 2021, Lebanon CSPE 2021, Zimbabwe CSPE 2021.

⁶⁴ El Salvador DE 2020, Madagascar DE 2020, Namibia DE 2020, Niger DE 2020, Syria DE 2020, Gambia DE 2021, Ghana DE 2021, Guinea Bissau DE 2021, Safety Nets PE 2019, Gender PE 2020, SSTC PE 2021, Resilience SE 2018, Emergency Response SE 2019, Technology SE 2021, Cameroon CSPE 2020, DRC CSPE 2020, Timor-Leste CSPE 2020, China CSPE 2021, Laos CSPE 2021, Lebanon CSPE 2021, Zimbabwe CSPE 2021.

⁶⁵ Malawi DE 2018, Eswatini DE 2019, Malawi DE 2019, Guinea Bissau DE 2020, Lebanon DE 2020, Malawi DE 2020, Mozambique DE 2020, Guinea Bissau DE 2021, Economic Empowerment of Rural Women DE 2021, Libya DE 2021, Sri Lanka DE 2021, Safety Nets PE 2019, Gender PE 2020, SSTC PE 2021, Resilience SE 2018, UN Cooperation SE 2021, Technology SE 2021, Cameroon CSPE 2020, Timor-Leste CSPE 2020, Gambia CSPE 2021, Honduras CSPE 2021, Zimbabwe CSPE 2021.

⁶⁶ Bangladesh DE 2018, Malawi DE 2020, Pilot CSP SE 2018, Resilience SE 2018, Emergency Response SE 2019, School Feeding SE 2020, Bangladesh CSPE 2020, Timor-Leste CSPE 2020.

⁶⁷ Bangladesh DE 2018, Gambia DE 2021, Cameroon CSPE 2020, DRC CSPE 2020, Indonesia CSPE 2020, Gambia CSPE 2021, Lebanon CSPE 2021, Emergency Response SE 2019.

⁶⁸ Algeria DE 2018, Bangladesh DE 2018, Malawi DE 2018, Philippines DE 2018, Senegal DE 2018, Somalia DE 2018, Turkey DE 2018, Eswatini DE 2019, Malawi DE 2019, Burundi DE 2020, Lebanon DE 2020, Mozambique DE 2020, Namibia DE 2020, Niger DE 2020, Syria DE 2020, Gambia DE 2021, Guinea Bissau DE 2021, Economic Empowerment of Rural Women DE 2021, Laos DE 2021, Libya DE 2021, Sri Lanka DE 2021, Gender PE 2020, SSTC 2021, Technology SE 2021, Cameroon CSPE 2020, DRC CSPE 2020, Indonesia CSPE 2020, Timor-Leste CSPE 2020, China CSPE 2021, Gambia CSPE 2021, Honduras CSPE 2021, Laos CSPE 2021, Lebanon CSPE 2021, Zimbabwe CSPE 2021.

physical access. Twelve evaluations⁶⁹ reported that a baseline was either not established, was partially developed for some indicators or regions but not all, or was established after the activity commenced.

iii. Data quality

38. Fifty-six per cent of DEs and 62 per cent of CEs⁷⁰ noted general concerns with the quality of monitoring data. Examples included insufficient sample sizes for baselines, double counting of beneficiaries, use of the same indicator for household and community measures, and inconsistencies regarding how data were reported between activities or countries. Timing was also an issue with, for example, seasonal data being collected at different points during the year in certain cases. In some cases, the limited data quality meant that evaluation teams made limited or no use of certain data sets.

39. Eleven per cent of evaluations⁷¹ questioned the credibility of the monitoring data due to discrepancies found by the evaluation teams between data sources. This included discrepancies between activity and monitoring reports, the Country Office Monitoring and Evaluation Tool (COMET), other M&E systems, activity databases and/or internal communications.

iv. Data disaggregation

40. Thirty-four per cent of DEs and 52 per cent of CEs⁷² noted insufficient data disaggregation in data collection, analysis and/or reporting. This included disaggregation for sex, status (refugee versus national), disability and/or age, as discussed further under SQ 7 related to cross-cutting issues.

2.2.2 Recommended improvements to monitoring systems or practices

FINDING 4: Over 90 per cent of all evaluations in the sample recommended that improvements be made to monitoring systems or practices. Most of these recommendations focused on improving monitoring frameworks, but some also addressing data gaps, data quality, and disaggregation.

41. Over 90 per cent of the evaluations (95 per cent of CEs⁷³ and 88 per cent of DEs⁷⁴) included one or more recommendations to improve aspects of the monitoring system. Fifty-three per cent included more than one monitoring-related recommendation.

⁶⁹ Algeria DE 2018, Bangladesh DE 2018, Malawi DE 2018, Philippines DE 2018, Eswatini DE 2019, Malawi DE 2019, Mozambique DE 2020, Syria DE 2020, Economic Empowerment of Rural Women DE 2021, Sri Lanka DE 2021, Laos CSPE 2021, Zimbabwe CSPE 2021.

⁷⁰ Algeria DE 2018, Bangladesh DE 2018, Kenya DE 2018, Philippines DE 2018, Senegal DE 2018, Malawi DE 2019, Rwanda DE 2019, Burundi DE 2020, Guinea Bissau DE 2020, Lebanon DE 2020, Madagascar DE 2020, Mali DE 2020, Niger DE 2020, Syria DE 2021, Economic Empowerment of Rural Women DE 2021, Sri Lanka DE 2021, Gender PE 2020, SSTC PE 2021, Emergency Response SE 2019, School Feeding SE 2020, Cameroon CPSE 2020, DRC CSPE 2020, Indonesia CSPE 2020, Timor-Leste CSPE 2020, Gambia CSPE 2021, Laos CSPE 2021, Zimbabwe CSPE 2021.

⁷¹ Algeria DE 2018, Bangladesh DE 2018, Rwanda DE 2019, Timor-Leste CSPE 2020, Laos CSPE 2021, Zimbabwe CSPE 2021.

⁷² Algeria DE 2018, Kenya DE 2018, Philippines DE 2018, Malawi DE 2019, El Salvador DE 2020, Lebanon DE 2020, Mozambique DE 2020, Namibia DE 2020, Gambia DE 2021, Guinea Bissau DE 2021, Economic Empowerment of Rural Women DE 2021, Safety Nets PE 2019, Gender PE 2020, Resilience SE 2018, School Feeding SE 2020, El Salvador CSPE 2021, Bangladesh CSPE 2020, Cameroon CSPE 2020, Timor-Leste CSPE 2020, China CSPE 2021, Laos CSPE 2021, Lebanon CSPE 2021.

⁷³ Bangladesh CSPE 2020, Cameroon CSPE 2020, China CSPE 2021, DRC CSPE 2020, El Salvador CPSE 2021, Gambia CSPE 2021, Honduras CSPE 2021, Indonesia CSPE 2020, Laos CSPE 2021, Lebanon CSPE 2021, Timor-Leste CSPE 2020, Zimbabwe CSPE 2021, Gender PE 2020, Safety Nets PE 2019, SSTC PE 2021, Emergency Response SE 2019, Pilot CSP SE 2018, Resilience SE 2018, School Feeding SE 2020, Technology SE 2021.

⁷⁴ Algeria DE 2018, Bangladesh DE 2018, Economic Empowerment of Rural Women DE 2021, Eswatini DE 2019, Gambia DE 2021, Ghana DE 2021, Guinea-Bissau DE 2021, Honduras DE 2021, Laos DE 2021, Lebanon DE 2020, Libya DE 2021, Madagascar DE 2020, Malawi DE 2018, Malawi DE 2019, Malawi DE 2020, Mozambique DE 2020, Namibia DE 2020, Niger DE 2020, Philippines DE 2018, Somalia DE 2018, Sri Lanka DE 2021, Syria DE 2020, Bangladesh DE 2020, Mali DE 2020, Lesotho DE 2018, El Salvador DE 2020, Kenya DE 2018, Senegal DE 2018.

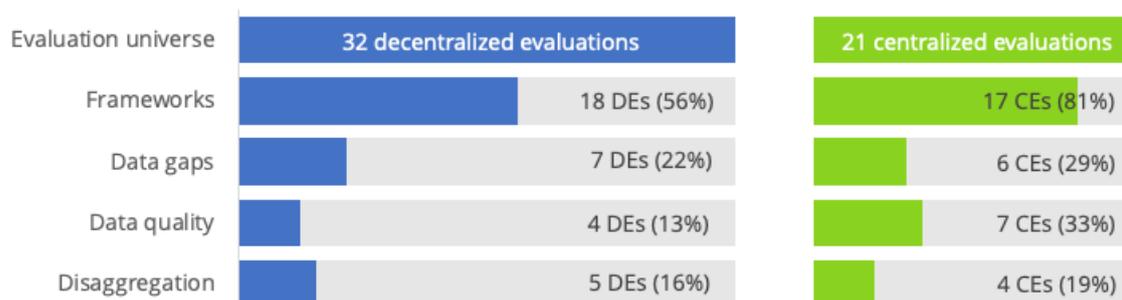
Figure 7: Number of evaluation recommendations related to monitoring (n=53)



Source: Synthesis team analysis

42. The numbers and proportions of evaluations that made recommendations to address credibility issue areas identified in section 2.1.1 are shown below in Figure 8.

Figure 8: Coverage of monitoring topics in evaluation recommendations



Source: Synthesis team analysis

i. Monitoring frameworks

43. Fifty-six per cent of the DEs and 81 per cent of the CEs⁷⁵ included at least one recommendation that specifically addressed improvements to the monitoring framework. These recommendations included developing or improving a theory of change; adding or changing indicators; setting targets; creating or improving assumptions; and expanding the use of qualitative data. This was most notable in CEs, where 81 per cent of evaluations included at least one recommendation in this area (compared to 56 per cent of DEs).

⁷⁵ Economic Empowerment of Rural Women DE 2021, Guinea-Bissau DE 2021, Libya DE 2021, Gambia DE 2021, Namibia DE 2020, Syria DE 2020, Malawi DE 2020, Lebanon DE 2020, Benin DE 2020, Madagascar DE 2020, Malawi DE 2019, Eswatini DE 2019, Niger DE 2020, Sri Lanka DE 2021, Mozambique DE 2020, El Salvador DE 2020, Kenya DE 2018, Bangladesh DE 2018, Cameroon CSPE 2021, Indonesia CSPE 2020, Bangladesh CSPE 2020, El Salvador CPSE 2021, Honduras CSPE 2021, Lebanon CSPE 2021, Gambia CSPE 2021, Laos CSPE 2021, Zimbabwe CSPE 2021, Gender PE 2020, Safety Nets PE 2019, SSTC PE 2021, School Feeding SE 2020, Emergency Response SE 2019, Pilot CSP SE 2018, Resilience SE 2018, Technology SE 2021.

ii. *Data gaps*

44. Twenty-two per cent of DEs and 29 per cent of CEs⁷⁶ recommended improving or expanding data collection. These recommendations were made to address gaps in baselines, expand into new areas, or to recommend additional analysis to improve monitoring performance.

iii. *Data quality*

45. Thirteen per cent of DEs and 33 per cent of CEs⁷⁷ recommended improvements to data quality. This included clarifying what and how data were collected, strengthening data collection systems, and changes to data collection tools.

iv. *Data disaggregation*

46. Sixteen per cent of DEs and 19 per cent of CEs⁷⁸ recommended improving the disaggregation of data by modifying or adding data collection fields, adding disaggregation categories to data storage systems, conducting additional analysis of data already collected, or providing clear guidance on specific categories such as “start monitoring the output indicators that refer specifically to girl pupils in grades 4, 5 and 6 separately.”⁷⁹ Some recommendations include how data should be used such as:

“GEWE-relevant effects of climate service provision should be monitored and analysed through sex-disaggregated data at both the individual and household levels ... and existing sex-disaggregated quantitative and qualitative data should be used to explore any apparent differences between men and women in their perceptions and use of [services] and their access to [technology].”⁸⁰

2.2.3 Use of monitoring information

FINDING 5: Evaluations provided examples of monitoring information used for both accountability and learning. Monitoring information used for reporting within WFP and to donors served an accountability objective, while – to a lesser extent – its use by management to adjust current activities, inform activity design and share lessons with stakeholders served a learning objective.

47. Evaluations provided examples of the use of monitoring data for accountability and, to a lesser extent, learning purposes. Fifty-seven per cent of the evaluations⁸¹ documented the use of monitoring data for reporting within WFP or to donors. However, only 32 per cent of the evaluations⁸² documented a learning use outside of the evaluation itself, such as data use for management of current activities, informing activity design and sharing lessons with stakeholders.

48. As shown below in Figure 9, evidence of use was more widespread in CEs, with 81 per cent documenting accountability uses and 48 per cent documenting learning uses. Within this group, all the

⁷⁶ Economic Empowerment of Rural Women DE 2021, Namibia DE 2020, Lebanon DE 2020, Laos DE 2021, Malawi DE 2019, Sri Lanka DE 2021, Algeria DE 2018, Indonesia CSPE 2020, Gambia CSPE 2021, Zimbabwe CSPE 2021, Timor-Leste CSPE 2020, Resilience SE 2018, Technology SE 2021

⁷⁷ Economic Empowerment of Rural Women DE 2021, Lebanon DE 2020, Benin DE 2020, Philippines DE 2018, Malawi DE 2019, Niger DE 2020, Indonesia CSPE 2020, Gambia CSPE 2021, SSTC PE 2021, Emergency Response SE 2019, Senegal DE 2018.

⁷⁸ Lebanon DE 2020, Malawi DE 2019, Guinea-Bissau DE 2021, Namibia DE 2020, Lesotho DE 2018, DRC CSPE 2020, El Salvador CSPE 2021, Laos CSPE 2021, Safety Nets PE 2021.

⁷⁹ Guinea-Bissau DE 2021.

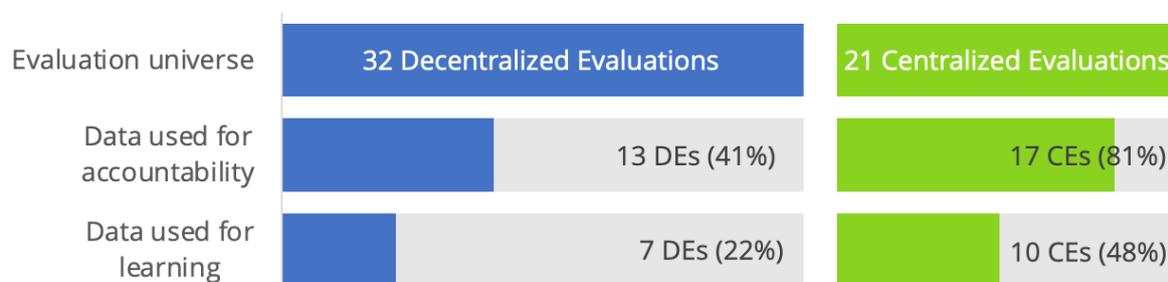
⁸⁰ Malawi DE 2019.

⁸¹ Algeria DE 2018, Kenya DE 2018, Senegal DE 2018, Somalia DE 2018, Turkey DE 2018, Malawi DE 2019, El Salvador DE 2020, Lebanon DE 2020, Syria DE 2020, Gambia DE 2021, Ghana DE 2021, Guinea Bissau DE 2021, Libya DE 2021; Safety Nets PE 2019, Gender PE 2020, SSTC PE 2021; Pilot CSP SE 2018, School Feeding SE 2020; El Salvador CSPE 2021, Bangladesh CSPE 2020, Cameroon CSPE 2020, DRC CSPE 2020, Indonesia CSPE 2020, Timor-Leste CSPE 2020, China CSPE 2021, Gambia CSPE 2021, Honduras CSPE 2021, Laos CSPE 2021, Lebanon CSPE 2021, Zimbabwe CSPE 2021.

⁸² Algeria DE 2018, Kenya DE 2018, Turkey DE 2018, Syria DE 2020, Gambia DE 2021, Libya DE 2021, Safety Nets PE 2019, SSTC PE 2021, Bangladesh CSPE 2020, DRC CSPE 2020, Indonesia CSPE 2020, Timor-Leste CSPE 2020, Honduras CSPE 2021, Laos CSPE.

CSPEs addressed accountability and 67 per cent addressed learning. DEs were less likely to address use, documenting 41 per cent and 22 per cent of accountability and learning uses respectively. Higher documentation of the use of monitoring data in the CEs reflects the CSP process, which is designed to both report on corporate indicators (accountability) and incorporate monitoring data into the CSP development process (learning).

Figure 9: Evidence of use of monitoring data in evaluations



Source: Synthesis team analysis

49. Interviewees indicated that the use of monitoring information did not go far enough, focusing mainly on corporate reporting and accountability with insufficient attention to learning lessons, and use in decision-making and programme adaptation. This finding is also consistent with the Internal Audit of Monitoring in WFP (2018),⁸³ which found that indicators were tracked for reporting compliance rather than for learning purposes.

50. Examples of the use of monitoring data for **accountability** included:

- [Somalia DE 2018](#): Monitoring data were used to produce reports; however, the underlying data were assessed as too weak to be used for management purposes.
- [School Feeding SE 2020](#): “The burden of [corporate and donor] reporting reduces the attention paid to analysis and use of the data gathered ... Reporting on corporate indicators is an unavoidable chore, but not one that adds much to performance management at country office level (with little use and analysis of data collected).”
- [Democratic Republic of Congo \(DRC\) CSPE 2020](#): “Monitoring and evaluation focal points described writing an average of 12 post-distribution monitoring reports per quarter of 15–20 pages in length. This was considered by some programme managers to be a very large volume of information that was not systematically analysed and therefore of limited use for programming. A significant proportion of quantitative data was collected for donor reporting rather than for programme management.”
- [Zimbabwe CSPE 2021](#): “Overall the use of the available monitoring information was heavily oriented towards external reporting and accountability, rather than towards learning.”

51. Good practices that led to data use for learning included regular monitoring feedback sessions with management, internal and external sharing of information, data visualization and the use of dashboards. Examples of the use of monitoring data for **learning** by WFP, governments and other partners included:

- [Turkey DE 2018](#): “Monitoring mechanisms performed strongly under WFP leadership which underpinned the programme’s ability to learn and adjust, with WFP channelling relevant findings to an inclusive and responsive Governing Board. M&E systems provided critical feedback to partners and inform programme improvements. Overall, monitoring systems allowed clear and positive conclusions to be drawn on beneficiary level results, as well as adjusting programme implementation.”

⁸³ WFP. 2018. Office of the Inspector General Internal Audit Internal Audit of Monitoring in WFP - Report No. AR/18/11.

- [Indonesia CSPE 2020](#): Monitoring data and vulnerability analysis and mapping (VAM) products were used by the government to inform current and future activities.
- [Lebanon CSPE 2021](#): Monitoring information was used to adapt the CSP and to expand and adapt programmes. Monitoring data was also used to increase the e-voucher transfer amounts to reflect inflation.
- [Zimbabwe CSPE 2021](#): Evidence on gender-specific outcome monitoring was used to inform programming adaptations: "The refugee baseline survey identified that women of reproductive age suffered poor nutrition, prompting targeted initiatives to improve their nutrition."

52. The [School Feeding SE 2020](#) provided a detailed description of the issue, noting that, "Chronic weaknesses in M&E limit knowledge generation of WFP's own operational effectiveness and efficiency ... and this hampers advocacy and resource mobilization." The evaluation also found that: "Learning from programme implementation has been insufficiently systematic and comprehensive, reducing learning and ability to demonstrate results." However, the evaluation also recognized that WFP was aware of continued evidence gaps on school feeding, and was developing a research consortium and knowledge platform for school feeding to improve access and knowledge-sharing. It also found that WFP has made efforts to collect stories, good practices and lessons learned, with 59 per cent of their survey respondents agreeing that "lessons learned and good practices have been shared in the region."

FINDING 6: Evaluations highlighted a need to expand qualitative data collection, analysis and reporting to contextualize WFP's achievements and to support WFP's ability to learn and adapt using its monitoring information.

53. Forty-three per cent of evaluations⁸⁴ noted the need to either start or expand qualitative data collection and reporting to be able to contextualize WFP's achievements. This need was primarily identified in the CEs – including all the PEs, 83 per cent of strategic evaluations and 50 per cent of CSP evaluations – and supported by 28 per cent of DEs. These evaluations noted the difficulty of relying solely on quantitative measures, especially for data related to gender, country capacity strengthening, and resilience. Only 15 per cent of evaluations⁸⁵ mentioned that qualitative data collection was taking place, including by the DRC CO, which was noted as a leader in their "qualitative context and data analysis."⁸⁶ The [Technology SE 2021](#) specifically cautioned against the "over-reliance on quantitative and remote approaches, which are not a good substitute for the richness of qualitative information and feedback collected in person."⁸⁷

54. Nineteen per cent of evaluations⁸⁸ recommended starting or improving qualitative data collection, analysis and/or reporting. Specific recommendations from the CEs, which were all agreed to in the management responses, included:

- [Safety Nets PE 2019](#): "Standardize monitoring of and reporting on WFP's contributions to social protection in order to establish a reliable base of quantitative and qualitative evidence on WFP's role and added value and enhance learning."
- [Emergency Response SE 2019](#): "Define success in emergency response and establish monitoring and evaluation processes and guidance to track success across emergencies. Include qualitative data questions to facilitate assessment of issues like relevance and responsiveness of operations to various needs of populations."

⁸⁴ Pilot CSP SE 2018, Resilience SE 2018, Emergency Response SE 2019, School Feeding SE 2020, Technology SE 2021, Safety Nets PE 2019, Gender PE 2020, SSTC PE 2021, Bangladesh CSPE 2020, Cameroon CSPE 2020, DRC CSPE 2020, Indonesia CSPE 2020, Gambia CSPE 2021, Zimbabwe CSPE 2021, Bangladesh DE 2018, Kenya DE 2018, Pakistan DE 2018, Malawi DE 2019, Mali DE 2019, Burundi DE 2020, Madagascar DE 2020, Namibia DE 2020, Libya DE 2021.

⁸⁵ Resilience SE 2018, DRC CSPE 2020, Gambia CSPE 2021, Bangladesh CSPE 2020, Kenya DE 2018, Somalia DE 2018, Malawi DE 2019, Malawi DE 2020.

⁸⁶ DRC CSPE 2020.

⁸⁷ Technology SE 2021.

⁸⁸ Libya DE 2021, Namibia DE 2020, Benin DE 2020, Madagascar DE 2020, Malawi DE 2019, Bangladesh CSPE 2020, Gender PE 2020, Safety Nets PE 2019, SSTC PE 2021, Emergency Response SE 2019.

- [Gender PE 2020](#): “WFP should ensure that the framework and guidance for midterm reviews and evaluations of first generation CSPs incorporate quantitative and qualitative analyses of WFP’s gender equality-related activities.”
- [SSTC PE 2021](#): “Ensure that approaches to generating evidence and fostering learning on SSTC draw from both qualitative and quantitative analysis and reporting.”

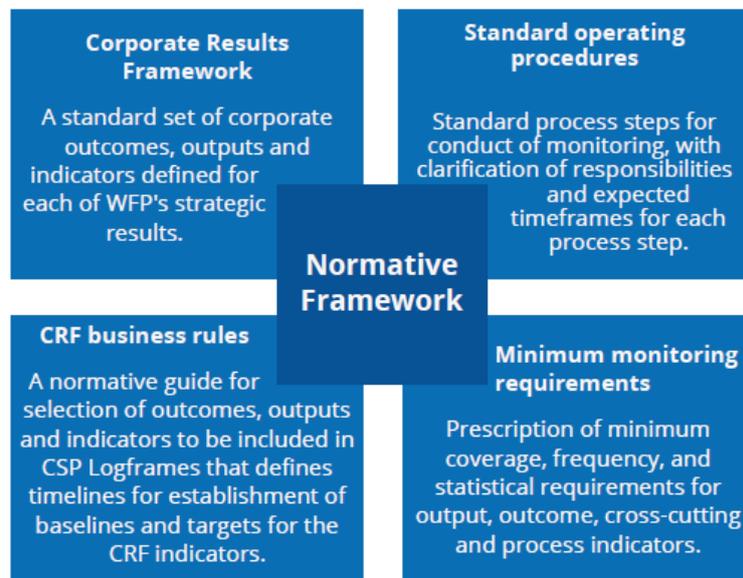
SYNTHESIS QUESTION 3 – To what extent has WFP’s Normative Framework enabled WFP to track programme effectiveness and inform corporate performance reporting?

FINDING 7: While some evaluations refer to the CRF (see questions 1 and 2), the other components of WFP’s Normative Framework were not directly discussed. However, some evidence was provided around the steps of the monitoring SOPs that linked the use of the SOPs to tracking programme effectiveness and corporate performance reporting.

55. Performance measurement and monitoring is guided by the WFP Normative Framework for Monitoring, which includes the CRF, CRF Business Rules, SOPs and MMRs. A summary of each component’s expected contributions to the Normative Framework is shown below in Figure 10. Of these four component elements, the evaluations included in the synthesis:

- provide some evidence related to the **CRF** (see synthesis questions 1 and 2)
- do not directly discuss the **SOPs**, but present evidence around its underlying processes in the context of the CSP monitoring cycle (see 2.3.1 below)
- contain only one direct reference to the **MMRs** (see 2.3.2 below)
- do not refer to any aspect pertaining to the **CRF Business Rules**.

Figure 10: WFP's Normative Framework for monitoring

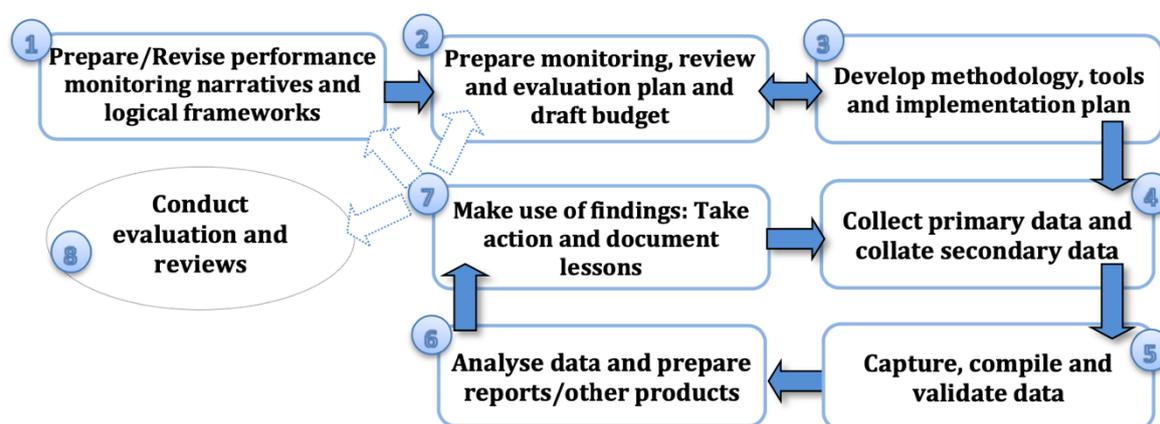


Source: WFP Corporate Monitoring Strategy 2018–2021

2.3.1 Standard operating procedures

56. While the evaluations in the synthesis did not directly reference the SOPs, elements of the eight standard process steps of the CSP monitoring cycle (Figure 11) were documented.

Figure 11: The eight standard process steps of the CSP monitoring cycle



Source: WFP SOPs for CSP Monitoring 2017

57. Each step was assessed according to the number of evaluations in which reference is made to that component of the process. High coverage signifies that over 75 per cent of all evaluations in the sample include evidence on that topic; medium coverage is 50–75 per cent; and low coverage is when evidence appears in fewer than 50 per cent of the evaluations. In general, where the SOP steps were being followed, the monitoring systems performed well.

Table 3: Summary of evidence related to standard operating procedures

SOP	Sample coverage	Examples of issues highlighted
Step 1: Prepare/revise performance monitoring narratives and logical frameworks	High	More than 75 per cent of the evaluations mentioned a logframe or other monitoring framework. In cases where logframes were assessed as robust, monitoring systems tended to perform well. Conversely, if logframe were assessed as weak, monitoring systems did not perform well. See also SQ2: evaluations noted difficulties developing logframes when the CRF indicators were changing.
Step 2: Prepare monitoring, review, and evaluation plan and draft budget	Low	Fewer than half of the evaluations mentioned having some component of a monitoring, review and evaluation (MRE) plan. Details were rarely given other than to state that there either was or was not a plan. Financial resources or budgeting for monitoring were not widely discussed in the evaluations, and drafting an M&E budget was not directly addressed. See also SQ4.
Step 3: Develop methodology, tools and implementation plan	Low	No evaluations made direct references to developing a monitoring methodology, tools or implementation plan. Some evaluations did note that development of a monitoring database was important for tracking programming effectiveness and informing corporate reporting, and that having data split over multiple databases could be detrimental to data use. See also SQ5.
Step 4: Collect primary data and collate secondary data	Medium	About two thirds of the evaluations addressed monitoring data collection. Data credibility and challenges in data collection are addressed in SQ 2, and the factors affecting data collection are addressed in SQ 5.
Step 5:	Medium	About half of the evaluations addressed capture of monitoring data. These findings are addressed in detail in the section on technology under question 5.

Capture, compile, and validate data		
Step 6: Analyse data and prepare reports or other products	Medium	About half of the evaluations addressed data analysis and reporting.
Step 7: Make use of findings	High	Nearly all of the evaluations addressed use of monitoring data. See question 2 for these findings.
Step 8: Conduct evaluation and reviews	High	All of the evaluations addressed monitoring data in evaluation by nature of either its inclusion or its exclusion from the reports. See also SQ1 and 2

2.3.2 Minimum monitoring requirements

58. Only one evaluation directly referenced the MMRs. The evaluation’s findings “[suggest] that WFP should revise its MMRs to provide further flexibility and not require the use of corporate indicators that are not suitable,”⁸⁹ noting a perceived burden on COs to collect additional country office specific data to support learning.

SYNTHESIS QUESTION 4 – To what extent does the evaluative evidence provide learning on the three outcomes of the WFP Corporate Monitoring Strategy?

59. The WFP Corporate Monitoring Strategy has three outcomes, as shown in Figure 12. While no evaluations directly referenced the Strategy, elements related to the three outcomes were documented in 76 per cent of the CEs and 56 per cent of the Des.⁹⁰ Findings by outcome are discussed below.

Figure 12: Defined corporate monitoring strategy outcomes



Source: WFP Corporate Monitoring Strategy 2018–2021

2.4.1 Strategic monitoring outcome 1: Adequate monitoring expertise

FINDING 8: Thirty per cent of the evaluations included in the synthesis identified shortcomings in staffing levels and capacity, with only 6 per cent assessing as sufficient the staffing levels allocated to the

⁸⁹ Laos CSPE 2021.

⁹⁰ Algeria DE 2018, Bangladesh DE 2018, Lesotho DE 2018, Philippines DE 2018, Senegal DE 2018, Somalia DE 2018, Turkey DE 2018, Burundi DE 2020, El Salvador DE 2020, Malawi DE 2020, Mozambique DE 2020, Namibia DE 2020, Gambia DE 2021, Guinea-Bissau DE 2021, Economic Empowerment of Rural Women DE 2021, Laos DE 2021, Libya DE 2021, Sri Lanka DE 2021, Gender PE 2020, Pilot CSP SE 2018, Resilience SE 2018, Emergency Response SE 2019, School Feeding SE 2020, Technology SE 2021, Bangladesh CSPE 2020, Cameroon CSPE 2020, DRC CSPE 2020, Timor-Leste CSPE 2020, China CSPE 2021, Gambia CSPE 2021, Honduras CSPE 2021, Laos CSPE 2021, Lebanon CSPE 2021, Zimbabwe CSPE 2021.

monitoring function. The evaluations also noted concerns about monitoring staff being assigned too many varied responsibilities.

60. The evaluations gathered evidence on three aspects relating to monitoring expertise: staffing levels, staff capacity, and staff duties. References to training on monitoring were found in relation to WFP staff providing training to government or other partners, but not to WFP internal training for staff.

61. As with the findings for synthesis question 2 around the credibility of monitoring data, the evaluations tended to address deficiencies (30 per cent of evaluations⁹¹) more often than successes (6 per cent⁹²). Table 4 features several examples.

Table 4: Evidence of adequate monitoring expertise

Evidence category	Found to be sufficient (+)	Found to be deficient (-)
Staffing levels	<p>6 per cent of evaluations,⁹³ for example:</p> <ul style="list-style-type: none"> - Libya DE 2021: By enhancing its recruitment of national staff, the CO had better access to the field and stakeholders and improved activity-monitoring. - Turkey DE 2018: After initial delays in staffing, the M&E function was “strongly staffed” at the country and local office levels. 	<p>15 per cent of evaluations,⁹⁴ for example:</p> <ul style="list-style-type: none"> - Emergency Response SE 2019: Lack of personnel to support knowledge management. - Cameroon CSPE 2020: The M&E team was understaffed and could not effectively monitor the CSP. Despite recruitment efforts, many posts remained unfilled. - DRC CSPE 2020: “[Monitoring] staff shortages have a negative impact on timeliness and monitoring of operations.” - Lebanon CSPE 2021: Insufficient staffing in the VAM and M&E unit, which “impacted the capacity of the country office to analyse progress and results at outcome and impact levels.” - Sri Lanka DE 2021: Cited poor handover between M&E officers, low levels of experience, and lack of training as problematic.
Staff capacity	<p>No evidence</p>	<p>13 per cent of evaluations,⁹⁵ for example:</p> <ul style="list-style-type: none"> - Emergency Response SE 2019: Monitoring was the weakest skill set in WFP as identified during its KIIs. - School Feeding SE 2020: Monitoring staff capacity varied from one office to another, but overall WFP would benefit from focusing on strengthening staff capacity. - Gender PE 2020 and Gambia CSPE 2021: Noted a need to increase monitoring capacity specifically around gender. - Technology SE 2021: Successful use of technology is a key factor for the success of monitoring, but staff capacity in digital literacy had not been sufficiently supported.

⁹¹ Algeria DE 2018, Philippines DE 2018, Gambia DE 2021, Sri Lanka DE 2021, Economic Empowerment of Rural Women DE 2021, Gender PE 2020, Emergency Response SE 2019, Technology SE 2021, Cameroon CSPE 2020, DRC CSPE 2020, China CSPE 2021, Gambia CSPE 2021, Honduras CSPE 2021, Laos CSPE 2021, Lebanon CSPE 2021, Zimbabwe CSPE 2021.

⁹² Turkey DE 2018, Burundi DE 2020, Libya DE 2021.

⁹³ Turkey DE 2018, Burundi DE 2020, Libya DE 2021.

⁹⁴ Algeria DE 2018, Economic Empowerment of Rural Women DE 2021, Emergency Response SE 2019, Cameroon CSPE 2020, DRC CSPE 2020, Honduras CSPE 2021, Lebanon CSPE 2021, Zimbabwe CSPE 2021.

⁹⁵ Gambia DE 2021, Sri Lanka DE 2021, Gender Policy 2020, Emergency Response SE 2019, Technology SE 2021, China CSPE 2021, Gambia CSPE 2021.

Staff duties	No evidence	<p>9 per cent of evaluations,⁹⁶ for example:</p> <ul style="list-style-type: none"> - Philippines DE 2018: The Monitoring Assistant position had changed over the life of the activity to be more administrative and less hands-on. - Algeria DE 2018: Staff were spread too thin and were asked to cover more than one position at a time. - Laos CSPE 2021: The Monitoring Assistant position was changed to be a District Community Facilitator, increasing the role's responsibilities to cover both community mobilization and monitoring, resulting in a lack of monitoring expertise in those hired. - Zimbabwe CSPE 2021: The CO monitoring unit was "overburdened with diverse responsibilities (...) consequently insufficient information on performance and results was available during implementation to support adaptive management."
---------------------	-------------	---

2.4.2 Strategic monitoring outcome 2: Financial commitment

FINDING 9: The evaluations included in the synthesis rarely addressed the financial requirements for monitoring. When references are made, they pointed to a lack of funding as a hindering factor.

62. Financial resources or budgeting for monitoring was not widely discussed, with evidence identified in only 13 per cent of evaluations. Eleven per cent of evaluations⁹⁷ found that a lack of funding was hindering monitoring, while one evaluation⁹⁸ noted that funding was sufficient. Interviews echoed the concern raised by the 2018 Internal Audit of Monitoring in WFP of the deprioritization of resources for monitoring. Key informants at regional level expressed that when funding is not clearly set aside for monitoring, there is a tendency to use it for programming instead of measurement.

2.4.3 Strategic monitoring outcome 3: Functional capacity

63. Thirty-two per cent of evaluations⁹⁹ discussed the use of technology and knowledge management. These findings are discussed in detail in response to question 5 below.

⁹⁶ Philippines DE 2018, Algeria DE 2018, Laos CSPE 2021, Zimbabwe CSPE 2021, Resilience SE 2018.

⁹⁷ Somalia DE 2019, Eswatini DE 2019, Madagascar DE 2020, Sri Lanka DE 2021, DRC CSPE 2021, Emergency Response SE 2019.

⁹⁸ Burundi DE 2020.

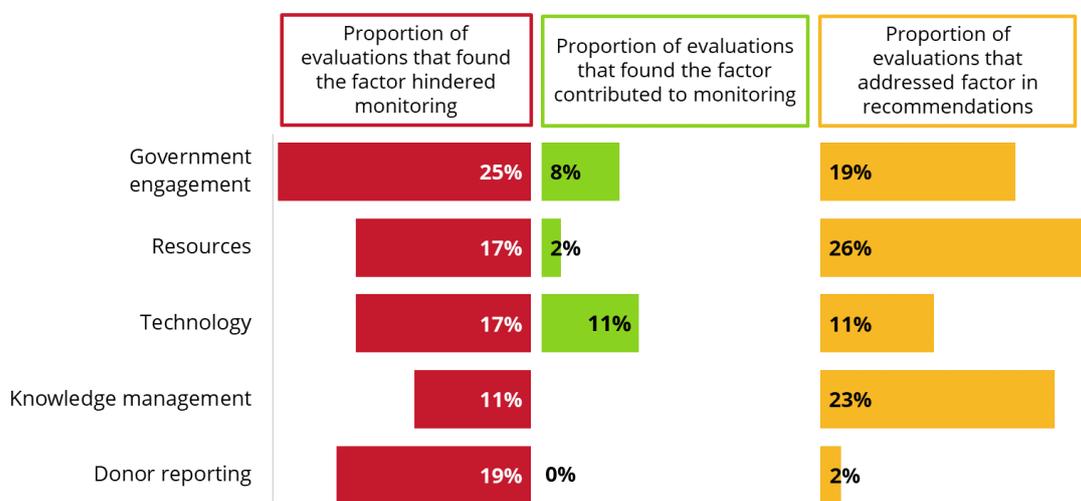
⁹⁹ Technology SE 2021, Bangladesh DE 2018, Senegal DE 2018, Somalia DE 2018, El Salvador DE 2020, Guinea Bissau DE 2021, Honduras DE 2021, Burundi DE 2020, Namibia DE 2020, Gambia DE 2021, Algeria DE 2018, Kenya DE 2018, Emergency Response SE 2019, Bangladesh CSPE 2020, DRC CSPE 2020, Indonesia CSPE 2020, El Salvador CSPE 2021.

SYNTHESIS QUESTION 5 – What factors contributed to or hindered implementation of performance measurement and monitoring systems? Are there particular activity areas and contexts in which WFP’s approach and systems have worked better and why?

2.5.1 Factors that contributed to or hindered monitoring

64. The synthesis identified five set of factors in the evaluations that contributed to or hindered implementation, and associated recommendations: government engagement, use of technology, knowledge management, donor reporting requirements, and staff and financial resources (Figure 13). Some of these factors clearly hindered or contributed to monitoring, but some were found to do both. As with the findings on other synthesis questions, evaluations were more likely to address negative factors than positive factors.

Figure 13: Factors that contributed to or hindered monitoring



65. In addition to these five factors, three evaluations¹⁰⁰ noted factors that relate to WFP’s culture of monitoring that had a positive influence. The [Gambia DE 2021](#) noted that the uptake and responses to monitoring findings reflect WFP’s interest in the views and feedback from the affected population. The [Turkey DE 2018](#) noted that WFP played a strong role in establishing the monitoring system, and that the CO had good and regular communications with external stakeholders. The [Timor-Leste CSPE 2020](#) noted that WFP’s presence on the ground helped data verification efforts and that WFP staff’s high level of commitment to gender issues was key to the success of its monitoring systems.

66. Two evaluations¹⁰¹ noted that having a strong local partner or non-governmental organization (NGO) for data collection was beneficial. The [Bangladesh CSPE 2020](#) found that “Local non-governmental organizations claim their partnership with WFP is mutually beneficial – non-governmental organizations receive support to develop their capacity and image, WFP benefits from their local knowledge and connections.”

¹⁰⁰ Gambia DE 2021, Turkey DE 2018, Timor-Leste CSP 2020.

¹⁰¹ Turkey DE 2018, Bangladesh CSP 2020.

i. *Government relationship and support*

FINDING 10: The evaluations note that where engagement with the host government is positive, this contributed to the implementation of performance measurement and monitoring, while challenges with the host government hindered monitoring. Recommendations to strengthen government engagement included starting or improving joint monitoring, creating minimum monitoring requirements to cover multiple donor requests, integrating data collection practices into government systems, and capacity strengthening.

67. Eight per cent of the evaluations¹⁰² credited WFP's relationship with and support from the government as supporting monitoring. In the case of [Bangladesh DE 2018](#), "Close work and alignment between the WFP and the education local authorities enhanced their involvement in monitoring and supporting activities." [Namibia DE 2020](#) noted that: "[Technical assistance, including monitoring] can be said to be client owned and demand driven. This is due to the fact that the bulk of [technical assistance] costs have been paid for by the Government of Namibia." The evaluation also noted that WFP has staff collocated in the government's offices, which is "especially useful for forging close working relationships and developing insight on context including opportunities, constraints and timing/sequencing of interventions." Two evaluations note that use of government monitoring systems or data collection have been useful.¹⁰³

68. In comparison, 25 per cent of evaluations¹⁰⁴ cited challenges with the host government as hindering monitoring. The most common issues were challenges accessing data, or gaps in data collected by the government. Another issue was the government's lack of staff, staff capacity or financial resources for monitoring. Further detail was not provided in the reports.

69. For example, the [Guinea-Bissau DE 2020](#) found that the government did not have the capacity to set its own monitoring system, so it relied on organizations to do so. This resulted in the monitoring system changing with each agency or implementing partner coming in and out, hindering implementation.

70. Nineteen per cent of the evaluations¹⁰⁵ included at least one recommendation for improvements to government relationships and support. These included starting or improving joint monitoring, creating MMRs across donors to allow governments to collect the same data regardless of donors, integrating data collection into government systems, and capacity strengthening. Six were agreed to and four were partially agreed to.

ii. *Financial and staff resources*

FINDING 11: The evaluations pointed to a lack of funding and/or staff resources as hindering monitoring and recommended improvements on resourcing aspects including through: increasing staffing levels and/or improving current staff capacity through training; improving budgeting processes; and making sufficient resources available for staffing and monitoring activities.

71. Nine per cent of the evaluations¹⁰⁶ found that a lack of funding was hindering monitoring while one evaluation¹⁰⁷ noted that funding was sufficient. Fifteen per cent of the evaluations¹⁰⁸ noted challenges to staffing or staff capacity that hindered monitoring. Issues included high turnover, low staffing levels, monitoring capacity gaps and high workloads. The [DRC CSPE 2020](#) noted a particular challenge with staffing levels in sub-offices in areas of greater risk.

¹⁰² Bangladesh DE 2018, Burundi DE 2020, Namibia DE 2020, China CSPE 2021.

¹⁰³ Burundi DE 2020, China CSPE 2021.

¹⁰⁴ Senegal DE 2018, Somalia DE 2018, Turkey DE 2018, Eswatini DE 2019, Burundi DE 2020, Guinea Bissau DE 2020, Malawi DE 2020, Guinea Bissau DE 2021, Laos DE 2021, Bangladesh CSPE 2020, Indonesia CSPE 2020, Laos CSPE 2021, Lebanon CSPE 2021.

¹⁰⁵ Namibia DE 2020, Laos DE 2021, Honduras DE 2021, Eswatini DE 2019, Lesotho DE 2019, Malawi DE 2018, Bangladesh DE 2018, Zimbabwe CSPE 2021, Senegal DE 2018, Bangladesh DE 2020.

¹⁰⁶ Somalia DE 2019, Eswatini DE 2019, Madagascar DE 2020, Sri Lanka DE 2021, DRC CSP 2021.

¹⁰⁷ Burundi DE 2020.

¹⁰⁸ Somalia DE 2018, Eswatini DE 2019, Sri Lanka DE 2021, Cameroon CSP 2020, DRC CSP 2020, Gambia CSP 2021, Honduras CSP 2021, Lebanon CSP 2021.

72. The Internal Audit of Monitoring in WFP (2018) also observed difficulties in sustaining dedicated monitoring staff at the sub-office level, alongside the movement of staff to priority areas following CO reorganizations.

73. Twenty-six per cent of the evaluations¹⁰⁹ included at least one recommendation related to resources. Recommendations on staffing issues focused on increasing staffing levels and/or improving current staff capacity through training. Financial recommendations focused on budgeting and making sufficient resources available for staffing and monitoring activities. Of the 15 individual recommendations, 11 were agreed to and 4 were partially agreed to.

iii. Use of technology

FINDING 12: Effective use of technology – such as mobile data collection, mapping, dashboards, and data visualization – contributed to the implementation of performance measurement and monitoring, and was credited with enhanced ease and timeliness of reporting, improved efficiency, cost reductions, increases in the scale and/or frequency of data collection, and standardization of data collection.

Challenges that hindered monitoring included the lack of a useful centralized database for monitoring data at country level, the fragmentation of data systems, and underutilization of WFP’s mapping assets.

Evaluation recommendations included merging databases or sharing data between systems; increasing use of digital data collection; and creating dashboards and visualizations to encourage data use.

74. Nine per cent of the evaluations¹¹⁰ credited the appropriate use of technology as supporting monitoring. The benefits of technology use included enhanced ease and timeliness of reporting, improved efficiency, cost reduction, increases in the scale and/or frequency of data collection, and standardization of data collection. The *El Salvador CSPE 2021* credited switching from in-person to telephone interviews, with an 84 per cent reduction in costs.

75. Examples of successful uses of technology included integrating monitoring within government systems, mobile data collection, mapping, dashboards and visualization tools. Remote data collection was also credited for allowing data collection to continue through the COVID-19 pandemic. One interviewee highlighted that when the COVID-19 pandemic hit, WFP was able to quickly adapt systems and guidance and employ remote monitoring.

76. The *Technology SE 2021* found that WFP is moving towards better use of technology to support monitoring. However, it also found that low levels of digital literacy and lack of training were leaving a gap between the technology available to use and the capacity to use it, noting that “cooperating partners, implementing partners and field-level and local staff capacities need to be further developed to ensure the effective, safe and sustainable use of technologies.”

77. Seventeen per cent of the evaluations¹¹¹ noted areas where lack of technology or inefficient use of technology were hindering monitoring. Common issues were the lack of a database for storing monitoring data and data being split over multiple database systems. This was supported by the Internal Audit of Monitoring in WFP (2018), which observed multiple instances of ‘shadow systems’. Access to systems was another issue raised in the evaluations, such as mention of a sub-office unit reported to have no access to COMET and that needed CO support to access the data they required to perform their jobs. Moreover, in June 2021, the External Auditor’s report on the management of information on beneficiaries documented persistent challenges to data sharing and stressed that “progress needs to be made on sharing data with the main partners, for example by enabling them to access COMET.”¹¹² However, the *Technology SE 2021* also provided a nuanced analysis of the need to balance the costs and benefits of technology and the need

¹⁰⁹ Economic Empowerment of Rural Women DE 2021, Benin DE 2020, Sri Lanka DE 2021, Malawi DE 2018, Somalia DE 2018, Bangladesh CSPE 2020, Honduras CSPE 2021, Lebanon CSPE 2021, Zimbabwe CSPE 2021, SSTC PE 2021, Pilot CSP SE 2018, Ghana DE 2021, Namibia DE 2020, Malawi DE 2019.

¹¹⁰ Technology SE 2021, Burundi DE 2020, Namibia DE 2020, Gambia DE 2021, El Salvador CSPE 2021.

¹¹¹ Technology SE 2021, Bangladesh DE 2018, Senegal DE 2018, Somalia DE 2018, El Salvador DE 2020, Guinea Bissau DE 2021, Honduras DE 2021, Bangladesh CSPE 2020, DRC CSPE 2020.

¹¹² Report of the External Auditor on the management of information on beneficiaries. EB. 21-25 June 2021.

81. Eleven per cent of the evaluations¹¹⁵ found poor knowledge management and/or a lack of formal mechanisms to share monitoring data within the country office resulting in missed opportunities for learning and results-based management. Examples include:

- [Emergency Response SE 2019](#): Poor knowledge management at the organizational level constrains organizational learning and application of lessons across departments. There has been corporate acknowledgement in WFP of the need to improve knowledge management. Steps taken include: formalizing the requirement to undertake a lesson-learning exercise; expanding the OEV; and setting up a new tracking system for monitoring the implementation of evaluation and audit recommendations. The evaluation also noted that some individual divisions – such as nutrition and logistics – have well-developed knowledge management functions.
- [Zimbabwe CSPE 2020](#): Knowledge management systems were inadequately developed to support results-based management. The use of monitoring information was “heavily oriented towards external reporting and accountability, rather than towards learning,” and “[a]ctivity managers were only responsible for collection of output data and displayed an incomplete knowledge of performance at higher levels, which was ‘owned’ by the monitoring unit.”
- [Indonesia CSPE 2020](#): “Limitations in WFP knowledge management hamper the potential for reporting on achievements in analysis, communications, discourse and relationship building for policy-level discussions. The existing corporate monitoring and knowledge management mechanisms (country briefs, annual reports, logframes, and so forth) are not constructed to reflect the time and energy required for affecting policy.”
- [El Salvador CSPE 2021](#): “The lack of a knowledge management strategy was correlated with a low use of results-based management. Management (design, development, monitoring, exit strategy, and evaluation) has not yet optimized the systematic use of evidence and analysis. There seems to be a gap between the level at which data is generated and systematized and the managerial decision that determines programmatic progress. Lack of consolidation of information and decision flows can interfere with managing the [CSP] as an integrated whole.”

82. Twenty-three per cent of the evaluations¹¹⁶ recommended improvements to knowledge management to improve learning, including establishing annual learning events, developing dashboards that can be periodically consulted to inform analysis and decision, disseminating lessons learned internally and externally, creating feedback loops between field monitors and the CSP development process, and improving existing knowledge management systems for use in constrained environments. Eleven of the 12 recommendations were agreed to, and one was partially agreed to.

v. *Donor reporting requirements*

FINDING 14: While recognizing that meeting donor monitoring requirements is necessary, the evaluations found that these requirements placed burdens on country offices, and that the data generated were not used beyond reporting. They also noted a lack of harmonization between donors.

83. Nineteen per cent of the evaluations¹¹⁷ noted challenges with the donor reporting requirements. Eight per cent¹¹⁸ noted that the monitoring data collected for donors was being reported on, but was not useful for any other purpose. The [School Feeding SE 2020](#) and [Lebanon CSPE 2021](#) pointed out the additional burden placed on monitoring systems of donor reporting requirements. The Pilot CSP SE 2018 and the [Rome-based agency \(RBA\) collaboration JE 2021](#) noted that the lack of monitoring harmonization between donors and United Nations agencies led to efficiency challenges. The [Malawi DE 2018](#)

¹¹⁵ Algeria DE 2018, Kenya DE 2018, Emergency Response SE 2019, Bangladesh CSPE 2020, Indonesia CSPE 2020, El Salvador CSPE 2021.

¹¹⁶ Namibia DE 2020, Lebanon DE 2020, Honduras DE 2021, Mozambique DE 2020, DRC CSPE 2020, Bangladesh CSPE 2020, China CSPE 2021, El Salvador CPSE 2021, Lebanon CSPE 2021, Gender PE 2020, Safety Nets PE 2021, Technology SE 2021.

¹¹⁷ Malawi DE 2018, Malawi DE 2019, Syria DE 2020, Pilot CSP SE 2018, School Feeding SE 2018, RBA collaboration 2021, Bangladesh CSPE 2020, DRC CSPE 2020, Laos CSPE 2021, Lebanon CSPE 2021.

¹¹⁸ Malawi DE 2019, Syria DE 2020, DRC CSPE 2020, Laos CSPE 2021.

recommended that data be collected on all possible indicators so that they would have data over time when donors changed.

2.5.2 WFP's approach and systems by activity category

FINDING 15: Except for the areas of country capacity strengthening and resilience – domains where challenges in measurement and use of indicators are clearly mentioned – the evaluations showed no clear patterns linking activity categories and monitoring challenges and opportunities.

84. The synthesis team reviewed the evaluations for links between the 13 activity types defined by the revised CRF and monitoring challenges and opportunities, but found little evidence that was specific to the activity categories.¹¹⁹ However, the evaluations raised concerns about the effectiveness of output and outcome indicators related to the domains of CCS¹²⁰ and resilience.¹²¹

i. Capacity strengthening

85. Thirteen per cent of the evaluations¹²² noted historical shortcomings for the capacity strengthening indicators. The [Gambia CSPE 2021](#) found that: “While most of [the CRF] indicators are appropriate for measuring outcomes for crisis response, school feeding and nutrition activities, the indicators proposed for measuring progress supported by capacity strengthening activities are limited,” and noted that CCS-specific indicators were under review at the time. Two evaluations¹²³ (4 per cent) noted that they did not monitor capacity strengthening, because the indicators did not exist when the CSPs were developed.

86. The 2021 WFP *Synthesis of evidence and lessons on country capacity strengthening from DEs* found that only a third of the evaluations in their sample recorded any data for CCS-related indicators or targets. Of these 13 evaluations, only 4 found the monitoring to be satisfactory. The synthesis recommended that CCS indicators be refined to improve performance measurement and reporting in line with the 2022–2026 Strategic Plan and the CRF. The CCS policy update recently approved by the Board reflects a full uptake of the synthesis recommendation, with the view to update guidance and tools for embedding CCS in country strategic plans and a robust monitoring framework that builds on the CRF.¹²⁴

ii. Resilience

87. Eleven per cent of evaluations¹²⁵ noted historical shortcomings in measuring resilience. [The Resilience SE 2018](#) notes that as CRF indicator guidance is developed by programme focal points and does not address overlaps or synergies, each strategic outcome can only be tagged with one of three focus areas (crisis response, resilience-building and root causes). “The issue of focus-area tagging is relevant due to its role in funding alignment, since a resilience approach requires multi-sectoral, multi-year funding. The fact that one of the focus areas is called ‘resilience-building’ might give the impression that the others are not related to resilience, which would be unfortunate.” The evaluation also notes that: “Programme and monitoring staff see the value in gathering information on resilience, but current corporate tools do not enable them to do so systematically or effectively” and highlight a desire to use more qualitative data.

¹¹⁹ Activity categories are defined in the Revised CRF 2017–2021 as: 1) unconditional resource transfers to support access to food; 2) asset creation and livelihood support activities; 3) climate adaptation and risk management activities; 4) school meal activities; 5) nutrition treatment activities; 6) malnutrition prevention activities; 7) smallholder agricultural market support activities; 8) individual capacity strengthening activities; 9) institutional capacity strengthening activities; 10) service provision and platform activities; 11) emergency preparedness activities; 12) analysis, assessment and monitoring activities; and 13) other.

¹²⁰ CRF activity category 9 – institutional capacity strengthening activities.

¹²¹ While resilience is not its own activity category, it is widely used as such and is related to asset creation and livelihood support activities, activity category 2.

¹²² Pilot CSP SE 2018, Gambia CSPE, China CSPE 2021, Lebanon CSPE 2021, Timor-Leste CSPE 2020, Indonesia CSPE 2020, Laos CSPE 2021.

¹²³ China CSPE 2021, Lebanon CSPE 2021.

¹²⁴ WFP Country capacity-strengthening policy update WFP/EB.A/2022/5-A.

¹²⁵ Resilience SE 2018, Bangladesh CSPE 2020, Cameroon CSPE 2020, Lebanon CSPE 2021, Malawi DE 2019, Lebanon DE 2020.

88. The [Resilience SE 2018](#) found that two types of resilience assessment were under development that may help address this issue: the Resilience Index Measurement and Analysis II initiative and the resilience context analysis.

89. The [Malawi DE 2019](#) noted that Malawi was “ahead of the curve” on implementing resilience activities, and recommended that they share their lessons learned around monitoring with headquarters.

2.5.3 WFP’s approach and systems for the country context

FINDING 16: Few evaluations provided evidence on country contexts in which WFP’s approach and systems for monitoring have worked better and why.

90. While little evidence was found in the evaluations on this synthesis question, a few examples regarding country context can be found below.

91. The [Bangladesh CSPE 2020](#) provides a wealth of information related to adjustments that it made for monitoring and to its CSP in response to the influx of Rohingya refugees. While the evaluation found that the CSP structure was not initially well adapted to respond to a crisis, WFP was able to revise its plan by adding a new strategic outcome and succeeded in adapting to the new circumstances. Additionally, “many respondents across United Nations agencies, NGOs and the government applaud the scale-up and speed of the WFP response.”¹²⁶

92. While the evaluation did not characterize it as contributing or hindering, the [China CSPE 2021](#) notes that WFP’s operations in China are unique in that it only uses government data and reports directly to WFP headquarters instead of a regional bureau (RB).

93. The use of technology in monitoring can be hindered in low resource environments. Some challenges identified by the [Technology SE 2021](#) include weak physical infrastructure (such as low electricity and Internet coverage) and human and financial resource constraints (such as low digital literacy, lack of appropriate training, insufficient support staff).

94. Two evaluations – [El Salvador CSPE 2021](#) and [School Feeding SE 2020](#) – documented the effects of the COVID-19 pandemic on monitoring data collection. In both cases, country offices were able to successfully switch to remote data collection. The [School Feeding SE 2020](#) found that COVID-19 response showed “promising signs that digital data collection could be an efficient way to improve the quality of management information” and that the “introduction of digital methods of data collection offers possibilities for making monitoring more efficient and timely.”¹²⁷

¹²⁶ Bangladesh CSPE 2020.

¹²⁷ School Feeding SE 2020.

SYNTHESIS QUESTION 6 – To what extent is WFP’s performance measurement system aligned with national monitoring systems? How has WFP pursued opportunities to strengthen national monitoring systems?

2.6.1 Alignment with national monitoring systems

FINDING 17: Evaluations documented WFP’s overarching alignment with government priorities and plans, but rarely went into detail about the alignment of WFP’s performance measurement systems with national monitoring systems.

95. While evaluations discuss the extent of alignment of programmes and plans with national priorities and plans, they rarely analyse whether WFP’s performance measurement systems align with national monitoring systems. Most evaluations focus on alignment with WFP’s corporate requirements, primarily the CRF, rather than with national systems.

96. The *School Feeding SE 2020* was one of the few evaluations that explored alignment with national monitoring systems. It found that “poor alignment of WFP monitoring and reporting systems with those of national governments is a challenge for the sustainable handover of school feeding programmes to national institutions”. According to a survey conducted as part of the evaluation, only 13 per cent of HQ and RB respondents “agreed or strongly agreed that “WFP monitoring and reporting systems are able to align well with government systems.”

2.6.2 Strengthening national monitoring systems

FINDING 18: Evaluations noted that where insufficient attention is given to strengthening national monitoring systems, this may undermine efforts for transitioning interventions and sustainability.

97. Eleven per cent of the evaluations¹²⁸ noted WFP investments in monitoring systems and the provision of technical assistance and training on monitoring. Three evaluations¹²⁹ highlighted that the primary aim of technical assistance and training is to ensure partners can use WFP’s monitoring systems, tools and reporting, rather than to support governments to develop their own monitoring systems and/or take on monitoring responsibilities, and that this undermines ownership and sustainability.

98. Thirteen per cent of the evaluations¹³⁰ discuss the use of WFP performance measurement systems rather than national systems given the absence of, or inadequacies in, national performance measurement systems. According to the *School Feeding SE 2020* “almost 85 per cent of the countries did not yet have an established functional school feeding M&E system”. Examples of weak capacity and systems include:

- In Lao PDR, “district education officers were responsible for monitoring the activities, but as the quality of data generated by the system was too poor to satisfy donors (and WFP was not capable of addressing this lack of capacity), WFP removed the monitoring function from the officers”.¹³¹ The *Laos CSPE 2018* noted that the government is not yet ready (in terms of capacity and finances) to take over the monitoring function.
- In Togo, the WFP school feeding M&E system “was implemented for less than a year because the principals and school inspectors were not sufficiently trained to use it.”¹³²

¹²⁸ Bangladesh DE 2018, Lesotho DE 2018, Somalia DE 2018, Eswatini DE 2019, Namibia DE 2020, Malawi DE 2018.

¹²⁹ School Feeding SE 2020; Senegal DE 2018, Turkey DE 2018.

¹³⁰ School Feeding SE 2020, Malawi DE 2018, Senegal DE 2018, Guinea Bissau DE 2020, Guinea Bissau DE 2021, Technology SE 2021, Laos CSPE 2018.

¹³¹ School Feeding SE 2020.

¹³² Ibid.

99. Twenty-five per cent of evaluations¹³³ highlighted the need to do more to strengthen country-level monitoring systems and plans. According to the [School Feeding SE 2020](#), a lack of attention to strengthening national M&E systems was found to be hindering WFP efforts towards transitioning, and undermining sustainability:

“WFP efforts to strengthen national M&E capacities and align its school feeding monitoring and reporting systems with those of governments have been limited ... Where efforts were made to introduce national M&E systems that governments can take over, these efforts have in practice faced a range of challenges, and in some cases have been insufficiently owned by other school feeding cooperating partners ... WFP is also insufficiently exploring how it could strengthen and leverage existing government data from various sectors and improve synergies across the different sector data systems.”¹³⁴

SYNTHESIS QUESTION 7 – To what extent were cross-cutting issues (gender equality and women’s empowerment, protection, accountability to affected populations and the environment) reflected in monitoring practices, guidance and systems?

100. The synthesis team looked at the cross-cutting issues from the WFP Strategic Plan (2017–2021) and indicators from the revised CRF (2017–2021).

2.7.1 Accountability to affected populations

FINDING 19: There is limited evidence across the evaluations of accountability to affected populations in monitoring practices, guidance and systems, beyond a few mentions of data availability and use.

101. Accountability to affected populations (AAP) was measured by two cross-cutting corporate indicators. Relating to AAP, several evaluations discuss complaints and feedback mechanisms, but few directly relate these mechanisms to monitoring practices and potential adjustments made as a result of feedback received. Only a few evaluations (typically SEs and CSPEs) noted issues around the availability, quality and use of monitoring indicators and data on AAP.

- The [Bangladesh CSPE 2020](#) noted the limited use of indicators on AAP.
- The [El Salvador CSPE 2021](#) and [Zimbabwe CSPE 2021](#) noted the availability of data against accountability indicators. The [Gambia CSPE 2021](#) found good tracking of accountability indicators. The [Algeria DE 2018](#) found that the revised logframe gave greater attention to AAP.
- The [Safety Net PE 2019](#) found that in Turkey: “WFP has used evidence from monitoring and accountability to affected population mechanisms to advocate with the government for measures to promote the inclusion of households of refugees who lacked official addresses and adequate transfer values in the face of inflation.”
- The [Gambia DE 2021](#) discusses the use of questionnaires on protection and accountability in monitoring and how the findings led to remedial measures to overcome issues identified, reflecting WFP’s interest in views and feedback from the affected population. However, the evaluation includes a recommendation to “include an appropriate indicator for accountability to beneficiaries” implying inadequate coverage by existing indicators.
- The [Technology SE 2021](#) highlighted that technology-based feedback mechanisms have broadened the ways in which beneficiaries can share feedback and any concerns, and that these mechanisms have improved WFP’s ability to monitor, analyse and respond to beneficiary feedback. However, “an over-reliance on technological solutions may unintentionally exclude certain target populations,

¹³³ School Feeding SE 2020, Bangladesh CSPE 2020, Timor-Leste CSPE 2020, Lebanon CSPE 2021; Lesotho DE 2018, Malawi DE 2018, Senegal DE 2018, Somalia DE 2018, Turkey DE 2018, Eswatini DE 2019, Guinea Bissau DE 2020, Guinea-Bissau DE 2021, Namibia DE 2020.

¹³⁴ School Feeding SE 2020.

especially in contexts where there is no systematic effort to accommodate the needs of marginalized groups in the use of technology.”

2.7.2 Protection of affected people

FINDING 20: There is limited evidence across evaluations of how protection of affected populations is covered through monitoring practices, guidance and systems, beyond the need for improved indicators and additional data collection.

102. Not enough evaluations discussed monitoring protection-related issues to be able to identify patterns, but some examples identified included:

- *Capturing results with indicators:* According to the [Lebanon CSPE 2021](#), WFP’s focus on measuring corporate indicators using quantitative methods is not suited to capturing protection risks. Coupled with more remote data collection and “a trend in reduced interaction among WFP, cooperating partners and beneficiaries,” this has the potential to reduce the visibility of protection challenges on the ground and limit the voice of affected communities in programme design.
- *Data availability:* The [Burkina Faso DE 2020](#) found limited information on protection in reports and no evidence of any improvements in indicators to ensure coverage of protection-related issues across all activities. The evaluation also identified inadequate coverage in monitoring of protection-related issues at intra-household level. However, the evaluation noted improved communication between WFP staff managing feedback mechanisms and those managing monitoring. In comparison, the [El Salvador CSPE 2021](#) found that data were available against protection indicators.
- *Use of data:* The [Gambia DE D021](#) evaluation discusses the use of questionnaires on protection and accountability in monitoring, and how the findings led to remedial measures to overcome the issues identified, reflecting WFP’s interest in the views and feedback from the affected population. The [Gambia CSPE 2021](#) found good tracking of protection indicators.
- *Disaggregation:* The [Bangladesh CSPE 2020](#) and [DRC CSPE 2020](#) also found a lack of disaggregation by disability, restricting the ability to adequately monitor inclusion.

2.7.3 Gender equality and women’s empowerment

FINDING 21: The evaluations found limited integration of gender indicators in monitoring frameworks. Also, reporting on GEWE tends to focus on quantitative data, namely reporting beneficiary numbers, with limited analysis and use of qualitative data. The lack of qualitative data collection and analysis, including at the intra- and inter-household level, limits the ability to measure and analyse GEWE achievements, especially at the outcome level, given the need to consider perceptions and other aspects of more sensitive and intangible nature such as issues relating to personal safety and security, discrimination, and intra-household dynamics. Also, the evaluations found limited use and analysis of gender-related data. The evaluations often recommended a need to improve monitoring frameworks and practices, especially the collection of disaggregated data and gender-specific outcomes.

103. Of the four cross-cutting issues included in the scope of the synthesis,¹³⁵ GEWE was addressed most frequently in evaluations. The following sections detail evidence around GEWE related to monitoring frameworks; data availability and quality; data analysis and use; qualitative data; and technology and feedback processes.

¹³⁵ Disability and inclusion appeared explicitly as part of the CRF in 2022 and have not been prioritized for inclusion in the present synthesis, given the timeframe of evaluations included in the sample.

i. Monitoring frameworks

104. GEWE was measured by three mandatory cross-cutting corporate outcome indicators. [The Gender PE 2020](#) found that further progress was needed to align the CRF with the WFP Gender Policy. Reporting against Gender Action Plan Layer 1, which pairs gender policy objectives with selected CRF indicators, provided limited support to gender-based programming. The factors contributing to this weakness included use of the same cross-cutting indicator across different objectives; reliance on quantitative indicators compared to reporting on the results of qualitative gender analysis; and the absence of activities clearly linked to gender equality and transformative change. This exacerbated the overlaps among the Gender Action Plan, the gender policy objectives, and the minimum monitoring standards, creating confusion among country offices and the field office on what data to collect, when, for what purpose, and against which framework when reporting on GEWE results.¹³⁶

105. The [Gender PE 2020](#) found that a key driver behind COs developing 'gender-responsive monitoring strategies' was the country office management team, rather than corporate accountabilities. It also recommended that efforts to improve collection and use of data include capacity building to ensure robust analysis of both qualitative and quantitative data.

106. Forty-two per cent of evaluations¹³⁷ included at least one recommendation on improving monitoring practices, guidance and/or systems around GEWE. This included 58 per cent of CSPEs, 67 per cent of SEs, 67 per cent of PEs, but only 31 per cent of DEs. The recommendations primarily addressed collecting sex-disaggregated data and improving monitoring to capture gender-specific outcomes. A recent WFP review of the implementation of evaluation recommendations highlighted that gender ranks as the least frequently addressed theme, while the themes most frequently addressed by WFP management were those focused on funding and human resources.¹³⁸

ii. Data availability and quality

107. In 2017, the Gender Office published 142 gender equality output indicators for COs to use in their CSPs. The [Gender PE 2020](#) found that "only five were included in the corporate list, indicating a very low level of integration in the CRF and limited uptake and use by country offices." Fifteen evaluations¹³⁹ noted that most gender-related monitoring data collected is sex-disaggregated quantitative data, primarily at output level, largely because this data is mandatory. However, according to the [Gender PE 2020](#), "compulsory inclusion of sex- and age-disaggregated data ... has been a strong internal driver for action, however, it is not yet universally applied" and there are examples of data quality issues. Thirteen evaluations¹⁴⁰ highlighted gaps and deficiencies in sex-disaggregated data at output level.

108. There are three *mandatory* GEWE outcome indicators in the CRF, but evaluations found that these do not adequately reflect GEWE, with one evaluation suggesting these are "easily achievable"¹⁴¹ and another suggesting the indicator on women's decision-making is "difficult to understand."¹⁴² The [Gender PE 2020](#) and 13 other evaluations¹⁴³ (26 per cent) highlighted the lack of sex-disaggregated data at outcome

¹³⁶ Gender PE 2020.

¹³⁷ Gender PE 2020, Safety Nets PE 2019, Emergency Response SE 2019, Resilience SE 2018, Technology SE 2021, Pilot CSP SE 2018, Cameroon CSPE 2020, Zimbabwe CSPE 2021, Bangladesh CSPE 2020, El Salvador CSPE 2021, Honduras CSPE 2021, Lebanon CSPE 2021, Timor-Leste CSPE 2020, Economic Empowerment of Rural Women DE 2021, Namibia DE 2020, Niger DE 2020, Lebanon DE 2020, Malawi DE 2019, Lesotho DE 2019, Guinea-Bissau 2020, Syria DE 2020, Sri Lanka DE 2021.

¹³⁸ WFP Review of the implementation of recommendations from thematic evaluations of a strategic/global nature WFP/EB.A/2022/7-D. The review covered ten global evaluations published between 2016 and June 2020 and the related 65 recommendations with sub-recommendations.

¹³⁹ Gender PE 2020, Resilience SE 2018; China CSPE 2021, Gambia CSPE 2021, Laos CSPE 2021, Lebanon CSPE 2021, Zimbabwe CSPE 2021; Economic Empowerment of Rural Women DE 2021, Algeria DE 2018, Bangladesh DE 2018, Burkina Faso DE 2020, Kenya DE 2018, Malawi DE 2018, Rwanda DE 2019, Turkey DE 2018.

¹⁴⁰ Gender PE 2020; Emergency Response SE 2019; Cameroon CSPE 2020, DRC CSPE 2020, Indonesia CSPE 2020, Timor-Leste CSPE 2020; El Salvador DE 2020, Lebanon DE 2020, Malawi DE 2019, Mozambique DE 2020, Philippines DE 2018, Somalia DE 2018, Sri Lanka DE 2021.

¹⁴¹ El Salvador DE 2020.

¹⁴² Gambia DE 2021.

¹⁴³ Resilience SE 2018, School Feeding SE 2020; Cameroon CSPE 2020, China CSPE 2021, Lebanon CSPE 2021; Economic Empowerment of Rural Women DE 2021, El Salvador DE 2020, Gambia DE 2021, Kenya DE 2018, Lebanon DE 2020, Mozambique DE 2020, Philippines DE 2018, Timor-Leste CSPE 2020.

level, which is *not mandatory*. Key informants consulted for the synthesis noted how WFP has been working to review monitoring and reporting practice in this area with a view to strengthening mandatory outcome indicator reporting.

iii. *Data analysis and use*

109. The [Gender PE 2020](#) and 16 other evaluations¹⁴⁴ (32 per cent) refer to limited analysis and use of gender-related data. This limited the ability to understand how change occurs and informs potential programme adaptations. This included a lack of systematic analysis and use of sex-disaggregated data to inform programme design, implementation and reporting across programmes and COs. Only three evaluations discuss examples of analysis and use of gender-related data.¹⁴⁵

110. The main use of data is to report on beneficiary numbers,¹⁴⁶ with limited attention beyond assessing whether participation is equal.¹⁴⁷ The [School Feeding Strategy 2020](#) characterized the approach as “hitting the target (of equal numbers)” but “missing the point”. Nine evaluations¹⁴⁸ noted that gender indicators are insufficient to measure empowerment and gender relations.

111. Evaluations note that COs are using the Gender Toolkit, which includes guidance on collection and use of disaggregated data. The [Gender PE 2020](#) noted some improvements in collection and use of sex-disaggregated data in corporate reporting, with examples of improved data collection noted in CSP evaluations (such as the [Cameroon CSPE 2020](#)).¹⁴⁹ According to the [Gender PE 2020](#) and interviews, WFP recognizes the need for a more systematic approach to evidence analysis and use, and efforts are currently under way to improve this. For example, in the updated Gender Policy (2022), WFP commits to “the collection, analysis and use of data disaggregated by sex, age, disability and other sociodemographic attributes, including intra-household and qualitative approaches, wherever possible and as appropriate;” in addition to “updating and disseminating a gender capacity development plan” to ensure consistency in application across units and divisions.¹⁵⁰

iv. *Qualitative data*

112. In addition to the need to conduct more analysis of quantitative data, 21 per cent of evaluations¹⁵¹ highlighted a lack of qualitative information and analysis on gender, as WFP’s focus is largely on quantitative data and corporate reporting. According to the [Gender PE 2020](#), “an unintended effect of the CRF has been a focus on collecting quantitative data rather than building a qualitative understanding of GEWE in relation to WFP programming.”¹⁵² The [Gender PE 2020](#) noted that several evaluations found “a focus on reporting gender results against corporate indicators” for corporate reporting purposes, rather than reviewing, analysing and addressing underlying power imbalances. Interviews highlighted the need to strengthen qualitative measurement of performance, including indicators, data collection, analysis and reporting, especially on capacity strengthening and gender where quantitative measurement is often inadequate to explain change.

113. The [Gender PE 2020](#) found that programmes measured change only at household level. This does not allow WFP to understand issues influencing food and nutritional security that operate at an individual and intra-household level, and this “represents a serious shortcoming in WFP monitoring systems.” To understand the effects of interventions on GEWE, 15 per cent of evaluations¹⁵³ mentioned the importance

¹⁴⁴ School Feeding SE 2020, Resilience SE 2018, Emergency Response SE 2019; Cameroon CSPE 2020, China CSPE 2021, DRC CSPE 2020, Lebanon CSPE 2021; Economic Empowerment of Rural Women DE 2021, Algeria DE 2018, Bangladesh DE 2018, Turkey DE 2018, Eswatini DE 2019, El Salvador DE 2020, Lebanon DE 2020, Mozambique DE 2020, Gambia DE 2021.

¹⁴⁵ Gender PE 2020; Zimbabwe CSPE 2021; Kenya DE 2018.

¹⁴⁶ For example, Gender PE 2020 and School Feeding SE 2020.

¹⁴⁷ For example, Resilience SE 2018 and Emergencies SE 2019.

¹⁴⁸ Gender PE 2020; School Feeding SE 2020, Technology SE 2021; Cameroon CSPE 2020, Gambia CSPE 2021, Lebanon CSPE 2021, Zimbabwe CSPE 2021; Mozambique DE 2020, Gambia DE 2021.

¹⁴⁹ Gender PE 2020, Cameroon CSPE 2020.

¹⁵⁰ WFP Gender Policy 2022. (WFP/EB.1/2022/4-B/Rev.1)

¹⁵¹ Gender PE 2020, School Feeding SE 2020, Emergency Response SE 2019, Resilience SE 2018, Bangladesh CSPE 2020, DRC CSPE 2020; Kenya DE 2018, Malawi DE 2019, Madagascar DE 2020, Namibia DE 2020, Libya DE 2021.

¹⁵² Gender PE 2020.

¹⁵³ Gender PE 2020, School Feeding SE 2020, Emergencies SE 2019; Bangladesh CSPE 2020, DRC CSPE 2020; Madagascar DE 2020, Namibia DE 2020, Libya DE 2021.

of collecting and analysing qualitative evidence, including at the inter-/intra-household level, and exploring household dynamics and gender-related socioeconomic and power relationships between men and women that influence access to resources and opportunities.

v. *Technology and feedback processes*

114. [Technology SE 2021](#) found that technology is often seen as “gender-neutral” and gender is not considered during technology development, for example, when developing monitoring data collection applications. However, gender may interact with technology, for example, where women have lower access to mobile devices. The evaluation recommended that providing gender training to technology teams could help improve gender outcomes, as would focusing on improving the gender balance of technology teams which were found to be mostly men.

115. As mentioned in section 2.7.2 on protection and accountability to affected populations, recent experience of collecting data remotely from beneficiaries due to the COVID-19 pandemic has led to concerns regarding gender equitable participation in feedback processes, given that men are usually the ones responding. In addition, data on sensitive issues is difficult to collect remotely.¹⁵⁴

2.7.4 Environmental sustainability

FINDING 22: There is limited evidence across evaluations on environmental issues in monitoring practices, guidance and systems.

116. It is important to note that according to interviews, ‘environment’ and ‘climate change’ are conceptually distinct and monitored differently within WFP, but the terms are often used interchangeably, including in the evaluations reviewed.

117. There was one cross-cutting indicator related to environment in the 2019 revised CRF, which was mandatory for ‘Food Assistance for Assets’ interventions only. The indicator measures “the proportion of activities for which environmental risks have been screened and, as required, mitigation actions identified” in order to assess the extent to which “targeted communities benefit from WFP programmes in a manner that does not harm the environment.”¹⁵⁵

118. Overall, the evaluations make little reference to monitoring environmental issues, and where they do they typically refer to gaps. It is unclear if these gaps are acceptable (that is, the indicators are not mandatory for the interventions) or if they represent deficiencies in monitoring. For instance:

- [DRC CSPE 2020](#): Monitoring and reporting related to environmental risks is “highly fragmented”.
- [Gambia CSPE 2021](#): There is a lack of baselines and targets on environment and mentioned deficiencies in the indicator used (does not cover screening of environmental risks or mitigation actions).
- [Honduras CSPE 2021](#): No indicators found for the management environmental risks and mitigation actions.

¹⁵⁴ Lebanon CSP 2021 and KIIS.

¹⁵⁵ Revised CRF 2017–21.

3 Conclusions and recommendations

“What gets measured matters”

3.1 CONCLUSIONS

119. The evaluations reviewed for this synthesis provide a snapshot in time of WFP’s performance management and monitoring systems, as seen through the lens of each evaluation’s questions and their evaluation team’s findings, conclusions and recommendations. The findings enable the identification of trends and opportunities for improvement, and can be the basis for supporting ongoing and future improvement efforts. The main themes include opportunities to strengthen the use of data for learning and adjustments, improving the way that WFP captures its achievements through enhanced data credibility, and increasing the use of qualitative data (particularly in relation to cross-cutting issues).

120. The synthesis acknowledges that WFP is constantly working to improve its performance measurement and monitoring systems. As highlighted in the introduction (see also Table 1), many of the weaknesses identified in the evaluations included in the synthesis’ sample have already been addressed, or are the subject of ongoing efforts to identify and implement changes. Drawing from stakeholder consultations and secondary document analysis, the synthesis team could better situate the synthesis recommendations against the current context.

Conclusion 1: While WFP is collecting credible monitoring data, and the three outcomes of the Corporate Monitoring Strategy remain relevant overall, there are still opportunities for improvement around resourcing the monitoring function, using monitoring data, and supporting monitoring systems and processes. Focusing on factors that influence the approach and conduct of monitoring can also help to guide improvements to monitoring systems and data.

121. In many instances, WFP is collecting and reporting on credible monitoring data, even if evaluations tend to only highlight shortcomings and gaps in implementation and systems. The evaluations provided clear evidence of where the credibility of monitoring data was threatened related to monitoring frameworks, data collection, data quality, and data disaggregation. The following six factors contributed to or hindered implementation of performance measurement and monitoring and provide areas where WFP could focus efforts:

- Relationships with **government partners** are critical for successful monitoring practices. Instances of close collaboration, joint monitoring, colocation, technical support, training, and capacity strengthening provided support to monitoring practices.
- Sufficient **human and financial resources** remain a critical part of successful monitoring and evaluations. Hindering factors include elements such as high turnover, low staffing levels, monitoring capacity gaps and high workloads.
- Appropriate use of **technology** in support of monitoring goals will be critical going forward, especially while systems are still adjusting to the COVID-19 pandemic. Attention should also be paid to ensuring that technology does no harm and takes into account protection for affected populations.
- The successful implementation of **knowledge management** practices and systems is expected to support a shift towards using monitoring data for learning, as well as supporting a results-oriented culture.
- While recognizing that **donor reporting requirements** are beyond WFP’s control, they should be acknowledged as a factor in the workload of monitoring staff.

Conclusion 2: The Normative Framework continues to provide the structure to support effective performance management and monitoring, but it does not encourage WFP to better capture the breadth of its achievements or track them over time.

122. Where the framework has been adhered to, monitoring has been conducted effectively. For example, the use of logframes was linked with successful monitoring systems. However, evaluations repeatedly expressed concerns that the CRF indicators were not fully capturing WFP's efforts and achievements. The concern that "what gets measured matters" (and consequently gets funded) was reflected in countries where CO efforts were not well aligned to the CRF indicators. This gap presents opportunities for increasing use of country-specific and qualitative indicators to encourage the collection and use of country-level data for learning and adjustments. But it also raises questions about whether WFP is capturing the breadth of their country-level achievements at corporate level.

123. While evaluations raised the concern of tracking progress over time because of the changes to the Normative Framework, these changes appear unavoidable in the short term while WFP is moving towards addressing many of the issues captured in this synthesis to establish a better-fitting monitoring framework for the longer term.

Conclusion 3: WFP uses its monitoring data for reporting, but does not consistently leverage it for learning and adaptation of programmes.

124. WFP has a strong system for reporting that relies on the use of monitoring data. While evaluations noted that some country offices were using monitoring data to inform current or future activities, this finding was not universal. Some staff and units may be overburdened with accountability requirements to be able to move towards learning. Evaluations provided clear examples of good practices, such as debriefing meetings to share monitoring results and making use of data visualization and dashboards as communication and analysis tools.

125. Evaluations also provided strong evidence of the desire and need at both the country and corporate levels to expand qualitative data collection, analysis and reporting to support learning and adaptation, pointing out that a focus on merely counting beneficiaries will result in "hitting the target" but "missing the point".

Conclusion 4: Measurement of country capacity strengthening and resilience are specific activity areas that are in need of additional development to better capture WFP's performance and achievements.

126. CCS and resilience represented relatively new types of monitoring efforts for WFP during the period of the evaluations, and the available corporate indicators were seen as insufficient for measuring WFP's contributions. These areas represent efforts that are highly contextual and are closely aligned to each country's context and government priorities, which may not be easy to measure with a standardized set of indicators. While COs had the option of developing their own indicators to track CCS and resilience, the evaluations did not highlight this as a common practice. Insufficient ability to measure country capacity strengthening and resilience may also have contributed to the lack of evidence. These are also areas for which monitoring would benefit from additional qualitative data to support contextualization.

Conclusion 5: Monitoring efforts for gender equality and women's empowerment largely focus on accountability and reporting, especially quantitative reporting of beneficiary numbers, and less on learning.

127. The lack of gender outcome data and qualitative data collection and analysis reduces ability to measure results effectively and understand what contributes to change. For other cross-cutting issues (accountability to affected populations, protection and environmental sustainability), evidence is typically limited and fragmented across the evaluations, indicating either a lack of available evidence or insufficient use of existing evidence by evaluations.

3.2 RECOMMENDATIONS

128. The following recommendations are based on the synthesis findings and conclusions. A table documenting the pathways between findings, conclusions, and recommendations is found in Annex VI.

129. As discussed in the Introduction,¹⁵⁶ performance measurement and monitoring are dynamic functions in WFP where continual optimization is sought. Efforts to address the issues raised in this set of evaluations have, in many cases, already begun. While an evaluation synthesis by definition looks back at evidence gathered through past evaluations, the synthesis team sought inputs from stakeholders about the current status of performance measurement and monitoring in order to develop the following recommendations.

Table 5: Table of recommendations

No.	Recommendation	Responsibility	Supporting entities	Priority	Deadline for completion
1	Strengthen the resourcing and use of the monitoring function as an integral component of the programme cycle in support of learning objectives.	Director RAM	Programme – Humanitarian and Development Division (PRO); country directors and country office heads of programmes	High	June 2024
1.1	Strengthen communication and advocacy, including with regional and country directors, emphasizing that an effective and adequately resourced monitoring function is an integral part of the programme cycle. Advocate the use of the monitoring function not only for accountability but also in providing fundamental support for learning and programme or operational adjustments. Efforts should consider the importance of country office-level mechanisms for reflection with decision leaders, drawing from best practices, such as regular debriefing meetings, and acting on the results of analysis from the vulnerability analysis and	Director RAM	Regional and country directors; Programme Cycle Management Unit (PRO-M); Field Monitoring Service (RAM-M); regional monitoring advisers	High	Starting in June 2023 in line with the design, approval and implementation of second-generation country strategic plans

¹⁵⁶ See sections 1.2 Context and 1.3.4 Limitations, Risks, and Mitigations.

No.	Recommendation	Responsibility	Supporting entities	Priority	Deadline for completion
	mapping and monitoring and evaluation planning and budgeting tool. (Strategic)				
1.2	Consolidating and expanding on existing initiatives, take further steps to improve the visibility and use of monitoring data in the programme cycle. This may include enhancing existing templates, dashboards and guidance, and encouraging evaluations to capture lessons learned on monitoring, when appropriate. WFP should also track the use of published guidance relating to monitoring and performance measurement in order to examine whether the use of data for learning objectives has improved. (Operational)	RAM-M	Regional monitoring advisers; country office heads of programmes and monitoring and evaluation; Office of Evaluation	High	June 2024

No.	Recommendation	Responsibility	Supporting entities	Priority	Deadline for completion
2	Increase the use of qualitative data collection, analysis and reporting to better capture and enhance understanding of and learning from WFP's achievements.	RAM	CPP	High	February 2024
2.1	WFP should explore how the evidence gathered through qualitative data collection and analysis approaches – including, but not limited to, data on cross-cutting issues – can be better incorporated into corporate reporting and can better complement evidence gathered through more quantitative approaches, and should gather and share examples of instances where this is effectively achieved. (Strategic)	RAM	CPP; convenors and relevant members of the qualitative evidence generation task force	High	February 2024
2.2	Programme and policy monitoring and evaluation leads should build on current efforts in results measurement, including through qualitative evidence generation approaches, to facilitate learning at the country level. Such efforts should build on the tools and guidance that have been developed by the Research, Assessment and Monitoring Division. High priority areas include gender, country capacity strengthening and resilience building. (Operational)	Monitoring and evaluation leads in the Programme and Policy Development Department (PD); convenors of the qualitative evidence generation task force	RAM-M, CPP	Medium	Starting in June 2023, in line with the design, approval and implementation of second-generation country strategic plans
3	Provide enhanced support for improving country office monitoring systems based on the main threats to credibility identified in the evaluation synthesis.	RAM	Country office heads of programmes and monitoring and evaluation; PD monitoring and evaluation leads	High	January 2024
3.1	<i>Frameworks:</i> Regional bureaux should work with country offices to ensure that the indicators in the monitoring, review and evaluation plan are selected based on the logical framework and are relevant for measuring programme objectives and that the plan is implemented. This may include providing technical support on indicators, assumptions and targets, or the	Regional monitoring advisers	Country office heads of programmes and monitoring and evaluation	High	Starting in June 2023, in line with the design, approval and implementation of second-generation country strategic plans

No.	Recommendation	Responsibility	Supporting entities	Priority	Deadline for completion
	development of additional resource documents and training. (Operational)				

No.	Recommendation	Responsibility	Supporting entities	Priority	Deadline for completion
3.2	<i>Data gaps:</i> WFP should document the use, and distil the learning from implementation, of existing guidance on addressing the gaps in the monitoring data used for setting baselines and the gaps in routine monitoring data collection activities. (Operational)	RAM-M	Country office heads of programmes and monitoring and evaluation; country office annual country report focal points	High	January 2024
3.3	<i>Data quality:</i> WFP should document the use and distil the learning from implementation of existing guidance on data quality issues such as data consistency among countries and interventions, the frequency of data collection and the double counting of beneficiaries. (Strategic)	RAM-M	Country office heads of programmes and monitoring and evaluation; COMET focal points	High	January 2024
3.4	<i>Data disaggregation:</i> WFP should take steps to document and distil the learning from the use of data disaggregation guidelines and aim to close any remaining gaps through the development of additional guidance or training. This may include monitoring the implementation of the guidance in the corporate results framework for 2022–2025 and the associated indicator compendium, minimum reporting requirements and the Research, Assessment and Monitoring Division guidance note on data stratification and disaggregation. (Strategic)	RAM	CPP; PD policy and programme leads; country office heads of monitoring and evaluation; country office annual country report focal points	Medium	January 2024
4	Provide enhanced support for improving country office monitoring systems based on the enabling factors identified in this synthesis.	RAM	Country Capacity Strengthening Unit (PRO-TC); regional bureau and country office heads of programmes and monitoring and evaluation; Technology Division; Innovation and Knowledge Management Division; regional monitoring advisers; PD policy and programme leads	High	January 2024

No.	Recommendation	Responsibility	Supporting entities	Priority	Deadline for completion
4.1	<i>Government collaboration:</i> Building on existing efforts, guidance and support should be made available to country offices for identifying ways to improve government relationships and build the capacity of government monitoring systems. This may include approaches such as joint monitoring or the inclusion of data collection in an existing government system. (Strategic)	RAM-M; PRO-TC	Country office heads of programmes and monitoring and evaluation; regional monitoring advisers	High	July 2024
4.2	<i>Financial and staff resources:</i> At the headquarters and regional levels, strengthen leadership's use of results and analysis from the vulnerability analysis and mapping and monitoring and evaluation planning and budgeting tool in advocating the allocation of adequate and more consistent human and financial resources to monitoring at the country office level. (Operational)	RAM-M	Regional directors; regional monitoring advisers; country directors	Medium	Starting in June 2023, in line with the design, approval and implementation of second-generation country strategic plans
4.3	<i>Technology:</i> Relevant headquarters divisions and units should continue to support digital data collection and survey platforms such as the Codebook and Survey Designer, focusing on improving data quality and timeliness. The Research, Assessment and Monitoring Division should also continue to provide support for improving inclusiveness in remote monitoring data collection. (Operational)	RAM	Technology Division; regional bureau and country office heads of programmes and monitoring and evaluation	High	January 2024
4.4	<i>Knowledge management:</i> Regional bureaux should work with country offices to develop a plan for knowledge management that incorporates monitoring data and analysis, and templates that can be adapted by country offices, with a focus on supporting the use of monitoring data in decision making. See recommendation 1. (Operational)	Regional bureau knowledge management focal points	Innovation and Knowledge Management Division; regional monitoring advisers; PD policy and programme leads	High	January 2024

Annexes

Below, please find the following annexes:

- Annex I: Summary terms of reference
- Annex II: Evaluation timeline
- Annex III: Methodology
- Annex IV: Scoping note
- Annex V: Analytical framework
- Annex VI: Findings, conclusions and recommendations matrix
- Annex VII: List of component evaluations
- Annex VIII: List of persons interviewed
- Annex IX: Bibliography/evaluation library
- Annex X: Acronyms

Annex I: Summary terms of reference

Evaluation Synthesis of WFP's Performance Measurement and Monitoring 2018–2021

Summary Terms of Reference



Evaluation syntheses are part of the WFP 'toolkit' in support of its commitment to evidence-based decision-making. They serve the dual objectives of accountability and learning.

Reasons for the evaluation synthesis

There is growing demand for evidence generation across WFP, and a commitment for further systematic use of evidence to inform strategic directions, policies and programmes. Evaluation synthesis is one of the products commissioned by OEV to respond to such growing interest in and demand for succinct and actionable analysis drawing from completed evaluations.

WFP is committed to performance management and determined to strengthen its accountability. With the introduction of the Integrated Road Map, it has become vital for WFP to present robust and evidence-based results against resources deployed. Furthermore, accountability to affected populations as well as to governments, national and international partners and donors is a priority. WFP's performance measurement system also calls for learning. When credible and evidence-based information generated from WFP's monitoring systems is used to ask the right questions and provoke reflection, then sound management decisions can be made on operational design, planning and implementation, which also will influence positive operational effectiveness, impact, relevance, efficiency and sustainability.

The WFP Strategic Plan 2022-2025 states that evidence and lessons-learned from WFP's performance based on research, monitoring and evaluations, as well as knowledge produced by communities, governments and partners, will guide WFP's programmatic approaches. A commitment to leveraging technology and data to become a knowledge and evidence-driven organization, increasing the agility of monitoring systems and demonstrating long-term impact is being made.

Objectives and stakeholders of the evaluation synthesis

This synthesis aims to bring together findings on WFP's performance measurement and monitoring from WFP's centralized and decentralized evaluations completed between 2018 and 2021, provide learning and generate further evidence in this area of work to support WFP to enhance its corporate approach, as well as to inform the formulation of future guidance and strategies.

The synthesis will seek the views of, and be useful to, a range of WFP's internal and external stakeholders. The primary internal stakeholders and intended audience of the synthesis are WFP's Corporate Planning and Performance Division (CPP) and the Research, Assessment and Monitoring Division (RAM).

The synthesis will be presented to WFP's Executive Board First Session of 2023.

Key synthesis questions

This synthesis aims to answer the following questions:

QUESTION 1: Do the corporate outcome, output and cross-cutting indicators collected and analysed at country level allow for effective measurement of intervention achievements?

QUESTION 2: To what extent are the monitoring practices in line with corporate guidance and tools?

QUESTION 3: To what extent have monitoring systems been able to generate credible and evidence-based information (quantitative and qualitative) to inform intervention design, planning and implementation? How has this information been used, by whom and for what purpose?

QUESTION 4: To what extent has WFP's Monitoring Strategy and Normative Framework enabled WFP to track programme effectiveness and inform corporate performance reporting by providing achievement values for the programmatic indicators?

QUESTION 5: What factors contributed to or hindered effective implementation of performance measurement and monitoring systems? Are there particular activity areas and contexts in which the approach and systems have worked better and why?

QUESTION 6: To what extent is WFP performance measurement system aligned with national monitoring systems? How has WFP pursued opportunities to strengthen national monitoring systems through modelling and capacity strengthening interventions?

Scope and methodology

The scope of this synthesis will be identified using a two-phased approach. In phase one, OEV has identified a preliminary list of centralized and decentralized evaluations, completed over the time period 2018-2021 and with a post hoc quality assessment (PHQA) threshold above 60 percent (satisfactory).

Phase two will require the evaluation team to finalize the universe of evaluations to be synthesized, and OEV to validate it. This will be achieved through the development of a scoping note including a protocol for inclusion/exclusion criteria, the screening process and a final selection of the evaluation universe.

The evaluation synthesis team will be then expected to take a rigorous methodological approach in order to maximize the quality and use of the synthesis. The methodology will systematically address the synthesis questions in a way that meets the dual purpose of accountability and learning.

Key features of the methodological design are expected to include:

- Confirmation of final sample of evaluations to be included in the synthesis.
- Development of a comprehensive analytical framework based on the refined synthesis questions and early review of a sample of reports.
- Systematic analysis via (electronic or manual methods) of the inception reports and evaluation reports against the analytical framework, including data extraction and coding.
- Primary data gathering through interviews with key stakeholders,
- Secondary data gathering through structured analysis of additional documentation linked to the synthesis questions.

The combination of deductive approach with a inductive one will allow for a structured guided analysis from the outset but permits flexibility as the process unfolds. The methodology will further reflect the standards for independence and impartiality, in line with WFP's commitments under its Evaluation Policy 2016-2021.

Roles and responsibilities

EVALUATION TEAM: The synthesis will be conducted by an external evaluation team.

OEV EVALUATION MANAGEMENT: The evaluation synthesis will be managed by WFP OEV Evaluation Officers Federica Zelada (until April 2022) and by Francesca Bonino (from April 2022 onwards). They will be the main interlocutors between the synthesis team, represented by the team leader, and WFP counterparts, to ensure a smooth implementation process and compliance with OEV quality standards for evaluation synthesis process and content. Michele Gerli will provide research and data analysis support. Second level quality assurance will be provided by Deborah McWhinney, Head of OEV Global & Synthesis Unit. The Deputy Director of OEV, Anne-Claire Luzot, will approve all draft and final synthesis products.

INTERNAL REFERENCE GROUP (IRG): An IRG composed of selected WFP stakeholders will be established and request to review and comment on the draft evaluation synthesis report, provide feedback during briefings as required; be available for interviews with the evaluation team and attend the stakeholder workshop.

Communication

All synthesis products will be produced in English. As part of the international standards for evaluation, WFP requires that all evaluation syntheses are made publicly available. The summary report along with the management response to the synthesis recommendations will be presented to the WFP Executive Board in November 2022.

OEV will ensure dissemination of lessons through the annual evaluation report. The relevant Headquarter division and the Regional Evaluation Units will be encouraged to circulate the final synthesis report with their staff, with WFP country offices and WFP external stakeholders.

Timeline and key milestones

Inception Phase: January – April 2022

Data extraction and interviews: May – June 2022

Analysis and Reporting: July – August 2022

Stakeholder Workshop: September 2022

Executive Board: February 2023

Annex II: Detailed timeline

Table 6: Detailed timeline

	Planned activities	By whom	Key dates
Phase 1 – Preparation			
	Submission of draft TOR for review to QA2	EM	27 Oct
	Review of draft TOR	QA2	28–29 Oct 2021
	Revision of TOR	EM	2–3 Nov 2021
	Submission of draft TOR for review to DDoE	DDoE	4–7 Nov 2021
	Revision of draft TOR	EM	8–9 Nov 2021
	Draft TORs shared with LTAs to start preparing their proposals and with IRG for comments	DDoE	10 Nov (due 22 Nov)
	Deadline for IRG comments	EM	19 Nov 2021
	Revise TORs following stakeholder comments	EM	22–26 Nov
	Revised TOR submitted to QA2 and DDoE	QA2, DDoE	29 Nov – 1 Dec
	TOR approval	DDoE	3 Dec
	LTA proposal review	EM	22–26 Nov 2021
	Team selection & decision memo submitted	EM	3 Dec 2021
	PO finalization	Procurement	20 Dec 2021
	Final TOR sent to WFP Stakeholders	EM	20 Dec 2021
Phase 2 – Inception			
	Desk review of documents and e-library	Team + EM	10–17 Jan 2022
	Team orientation – Introductory calls synthesis team and OEV	Team	18–19 Jan
	Inception briefings with selected stakeholders	Team + EM	20–21 Jan
	Preparation of Inception Report, including selection of final evaluation universe	Team	24 Jan – 18 Feb
	Submission of scoping note with final universe of evaluation	Team	8 Feb
	Review of scoping note	OEV	8–10 Feb
	IR DO – Submit draft Inception Report (IR) to OEV	Team	24 Feb
	Quality assurance and comments to the ET	EM/ RA	24–28 Feb
	Review of D0 IR	QA2	1–2 Mar
	Feedback to ET		3 Mar
	IR D1 – Submission D1 IR	Team	11 Mar
	Review and clearance of D1 IR prior to submission to DDoE	QA2	15 Mar
	Review D1 IR	DDoE	15–18 Mar
	Revisions to address DDoE comments and submission of IR D3	TL	21–23 Mar
	Quality assurance	EM/ RA	24–25 Mar
	Inception Report approval	DDoE	28 Mar
	EM circulates final IR to key WFP stakeholders for their information	EM	1 Apr
Phase 3 – Desk review, content analysis and interviews			
	In-depth review of relevant information across evaluations; data extraction and coding	Team	2 Apr – 2 May

	Planned activities	By whom	Key dates
	Conduct interviews with stakeholders	Team	2 Apr – 2 May
	Content analysis	Team	2 Apr – 2 May
Phase 4 – Reporting			
Draft 0	Submission of draft synthesis report (D0) to OEV	Team	27 May
	OEV EM + RA review of draft 0 (QA1 review)	EM/ RA	30 May – 1 Jun
	Synthesis team adjustments to address QA1 comments	Team	2–6 June
	OEV QA2 review window	QA2	8–10 June
	TL adjustments to address QA2 comments and submits revised	Team	13–21 June
	QA1+QA2 review followed by QA2 request of additional revisions	QA1+QA2	21–23 June
	Team revision to address QA2 comments	Team	23 June/ 15 July
	EM /QA1 review followed by Team revisions	Team +EM+RA	18–20 July
	QA2 window to review draft	QA2	24–26 July
	Team revisions to respond to QA2 comments	Team	27 July -1 Aug
	Final QA1+QA2 parallel review and final adjustments by the Team before submitting to DDoE	QA1+QA2	2 Aug
	Submission to DDoE / DDoE window of review	DDoE	2–12 Aug
	Team revision to address DDoE comments	Team	15–17 Aug
	QA1+QA2 review and adjustments by the Team before seeking DDoE clearance to share draft with IRG	QA1+QA2	18–29 Aug
DDoE window for final review of the draft before clearance to share it with IRG for comments	DDoE	29 Aug – 1 Sept	
D1	Draft synthesis shared for comments with IRG	DDoE	1–13 Sept
	EM+RA compiles matrix of comments and shares it with the Team ahead of stakeholder meeting	EM	13 Sept
	Stakeholder meeting	IRG+Team	15 Sept
D2	Team submission of revised draft (D3)	Team	27 Sept
	QA1 review of revised draft followed by Team adjustments	EM+RA+Team	28 Sept – 4 Oct
	EM starts preparing the draft Summary Evaluation Report	EM	5 Oct
	DDoE comment window on the revised ER (D3)	DDoE	5–12 Oct
ER D3	TL submits final draft synthesis reflecting DDoE comments	Team	19 Oct
ER D4	EM check to ensure changes made to the synthesis report adequately respond to DDoE comments.	EM / TL	19–21 Oct
	EM consistency check between draft synthesis and draft SER	EM+RA	21–24 Oct
	EM submits zero draft SER comments	EM	24 Oct
	DDoE comment window on the draft SER	DDoE	24–31 Oct
	EM reviews draft SER to address DDoE's comments and seek clearance to send draft SER to Exec Management	EM	1–10 Nov
SER D2	WFP Executive Management/ Oversight and Policy Committee (OPC) comment window on SER	OPC	14-24 Nov
	EM discusses OPC comments received with DDoE and revise and finalize the SER accordingly	EM / DDoE	24–29 Nov
	Final consistency check between SER and Synthesis Report final adjustments made before seeking DDoE approval	EM/ TL	29 Nov – 2 Dec

	Planned activities	By whom	Key dates
	Seek final approval by DDoE on final SER and final synthesis report. Clarify last points as needed	DDoE + EM	2-7 Dec
Final SER+ER	Submission to EBS of final SER and final Synthesis report.	EM / DoE	9 Dec
	EB Secretariat deadline for EB.1/2023		9 Dec
Phase 5. Follow-up and dissemination			
	Submit SER/ recommendations to CPP for management response + Synthesis to EB Secretariat for editing and translation	EM	13 Jan 2022
	Tail end actions, OEV websites posting, EB roundtable etc.	EM	16 Jan 2023
	Presentation of Synthesis to the EB	DDoE & EM	Feb 2023
	Presentation of management response to the EB	CPP	Feb 2023

Note: TL=Team Leader; EM = Evaluation Synthesis Manager; OEV=Office of Evaluation. CPP – Corporate Planning and Performance Division

Annex III: Methodology

A3.1 SCOPE

130. This section provides an overview of the scope of the synthesis. The findings of the screening process are detailed in the scoping note (Annex IV), including the recommended list of 53 evaluations to include in the synthesis.

131. This synthesis is being applied retroactively. As such, the first task for the team was to establish the evaluations that should be included in the synthesis. During the preparation phase of the synthesis, OEV identified a preliminary universe of evaluations that could be included in this synthesis. The team then conducted a scoping exercise during the inception phase to determine the final sample of evaluations to be included in the synthesis.

132. The preliminary list of 68 evaluations identified by OEV was shortlisted based on the following three criteria:

- Evaluation type
 - **CEs¹⁵⁷ – Policy,¹⁵⁸ Strategic,¹⁵⁹ and Country Strategic Plan Evaluations¹⁶⁰**
 - **DEs¹⁶¹ covering activities, pilots,¹⁶² themes, and transfer modalities¹⁶³**
- Time period – evaluations completed between 2018 and 2021
- Quality of the evaluation – evaluations assessed by OEV's post hoc quality assessment (PHQA) system¹⁶⁴ above the 60 per cent threshold (satisfactory)

133. OEV identified an additional seven evaluations that met the same criteria but that were approved after the initial Terms of Reference (TOR) was published, bringing the total number of preliminary evaluation universe to 75. Table 7 below provides a summary of evaluations reviewed during scoping. A full list of evaluations is included in Annex VII, including the evaluation title, evaluation type, commissioner and report approval year.

Table 7: Evaluations reviewed during scoping by type

Centralized evaluations			Decentralized evaluations			Total
Country Strategic Plan	Policy	Strategic	Activity	Thematic	Transfer modality	75 (21 CE/ 54 DE)
12	3	6	48	4	2	

¹⁵⁷ Centralized Evaluations are commissioned and managed by OEV and presented to the Executive Board for consideration.

¹⁵⁸ The Evaluation of the WFP People Strategy (2014–2017) and the Evaluation of the WFP Corporate Partnership Strategy (2014–2017) are not included in this synthesis as considered not relevant by OEV.

¹⁵⁹ The Strategic Evaluation on Funding WFP's work (2020) will not be included in this synthesis as it is considered not relevant by OEV.

¹⁶⁰ Country Portfolio Evaluation are not included in this synthesis because the operations evaluated will have been designed at minimum two years before the Integrated Road Map period, and because WFP no longer operates under this framework.

¹⁶¹ Decentralized Evaluations are commissioned and managed by country offices, regional bureaux or Headquarters-based divisions other than OEV. They are not presented to the Board.

¹⁶² No evaluations of pilots met the other two qualifying criteria for this synthesis, so these are not included in any tables or analysis.

¹⁶³ Operation Evaluations will not be included in the synthesis, as WFP no longer operates under this framework.

¹⁶⁴ Since 2016, OEV has used an outsourced PHQA mechanism, through which independent assessors rate the quality of all completed WFP evaluations against WFP's own evaluation quality standards.

134. During the inception phase, the team developed a scoping note (Annex IV) to determine the sample of reports that would be included in the synthesis, the screening process and a final selection of the evaluation universe. This was validated through consultations with OEV. The approach included the following six steps:

- i. Identifying key information about each report
- ii. Establishing a set of topics related to performance measurement and monitoring
- iii. Reviewing the reports
- iv. Ranking the relevance of the evaluations to the synthesis topic
- v. Conducting a quality assurance review
- vi. Making recommendations for the inclusion of evaluations in the synthesis

135. First, key information about each report – including the commissioner, type of evaluation, language, approval year and PHQA score – was identified.

136. Second, a set of key words or topics related to performance measurement and monitoring were identified to inform the screening process. The TOR identified the following topics: staff capacity; strategy, guidance and tools; monitoring budgeting, planning and execution; monitoring systems; data quality; reporting; and follow-up and utilization of data. During the initial review of the evaluation reports, the team identified the following additional keywords: monitoring and M&E (M&E hereafter); theory of change; indicators, outputs and outcomes; logframe/logical framework; results framework; and data collection.

137. Third, the evaluations were reviewed for references to performance measurement and monitoring topics. Using MAXQDA, a qualitative analysis software, the team reviewed each evaluation's conclusions and recommendations and coded any segments that were relevant to the synthesis. Segments were not coded by individual topic, as that was to be done during the data extraction phase of the synthesis using the analytical framework.

138. The team did not code segments in the evaluation findings during the scoping process. A rapid review found that if performance measurement and monitoring topics were addressed in the conclusions and recommendations, they were also addressed in the findings. In cases where these topics were not addressed or lightly addressed in the conclusions and/or recommendations, the team reviewed the findings to determine whether there was sufficient documentation in the report to justify its inclusion.

139. Fourth, the team ranked the evaluations. Based on the segments identified, the team assessed the reports as having high, medium or low levels of evidence for this synthesis. These rankings were defined as:

- High – One or more topics addressed directly in the conclusions or recommendations. For example, a report in this category may include one or more recommendations about improvements to performance measurement and monitoring systems with specific details and supporting conclusions.
- Medium – Topics addressed in the conclusion or recommendation, but only to a limited extent. For example, some reports may mention performance measurement and monitoring in one or two conclusions, but not have any related recommendations. Or a report could have one recommendation that did not provide any detailed information or did not have any related conclusions.
- Low – Topics not addressed in the report or addressed to a limited extent. In these types of reports, performance measurement and monitoring may appear as a key word but with little contextual information. An example would be an evaluation that cites programme indicators as a data source but does not comment about data quality, availability, relevance, systems, and so on.

140. Fifth, the Team Leader conducted a quality assurance review of all evaluations with a ranking of medium or low to confirm whether they should be included in the synthesis.

141. Sixth, evaluations to be included and excluded were proposed. While guided by the number of segments identified in each report, due to variations between reports ranking was not a count of segments. For example, two evaluations may have the same number of relevant recommendations, but one

evaluation may deal with it in greater detail. In some cases, an evaluation may have a low number of relevant conclusions or recommendations, but the subject matter of the evaluation may provide a unique perspective for the synthesis. Therefore, the team used their experience as evaluators to provide a high/medium/low ranking for each report based on the totality and depth of relevant coverage, and whether they thought enough information was present to be of interest in this synthesis.

142. In total, 53 evaluations were considered to have high coverage of performance measurement and monitoring topics. These were selected because they are most likely to contain relevant evidence to answer the synthesis questions.

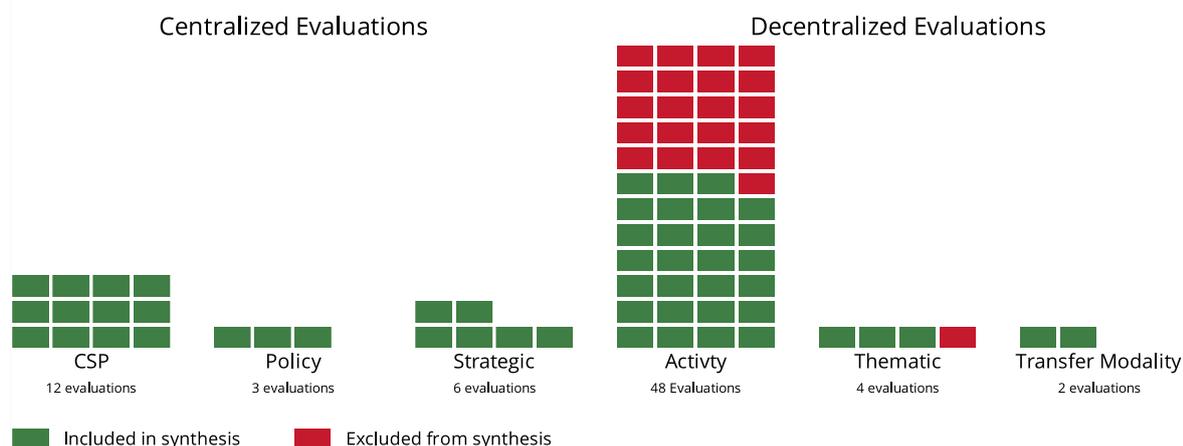
143. The final sample of evaluation types selected for the synthesis is included in Table 8 below. The synthesis will include 21 CEs and 32 DEs.

Table 8: Final sample of evaluations selected for the synthesis by type

Centralized evaluations			Decentralized evaluations			Total
Country Strategic Plans	Policy	Strategic	Activity	Thematic	Transfer modality	53 (21 CE/ 32 DE)
12	3	6	27	3	2	

144. Figure 15 below shows the final sample of evaluations that form part of the synthesis compared to the original universe. All the CEs were included in the synthesis as they all contained high levels of evidence related to performance measurement and monitoring topics. Of the 48 decentralized activity evaluations, 27 had high levels of evidence while 21 included medium or low levels of evidence and coverage. Although most of the evaluations were cut from the decentralized activity category, the team believes that sufficient evaluations remained in the category to be able to adequately answer the synthesis questions. Three of the four thematic evaluations and all the transfer modality evaluations were included in the synthesis.

Figure 15: Evaluations included by type



Source: OEV Management Information System

145. Figure 16 below shows the geographic coverage of the 12 CSP evaluations and 31 DEs that were included in the synthesis across WFP's six regional bureaux. The synthesis includes evaluations from all six RBs which will provide a good geographical balance to the analysis. The highest number of evaluations cover RB Dakar (10) and RB Johannesburg (10), followed by RB Bangkok (9), RB Cairo (6), RB Nairobi (4), and RB Panama (4). The nine centralized policy and strategic evaluations and one DE¹⁶⁵ commissioned by an

¹⁶⁵ Global End-term Evaluation of the Joint Programme on Accelerating Progress towards the Economic Empowerment of Rural Women in Ethiopia, Guatemala, Kyrgyzstan, Liberia, Nepal, Niger and Rwanda from 2014 to 2020.

HQ-based division involved the participation of multiple countries through desk reviews and country missions and are therefore not represented in Figure 16 below.

Figure 16: Geographic distribution of synthesis evaluations



Source: OEV Management Information System

146. Figure 17 below shows the distribution of the synthesis evaluations by the year in which they were completed. Fewer evaluations were completed in the first half of the sample (11 in 2018 and 5 in 2019) than in the second half (19 in 2020 and 18 in 2021).

Figure 17: Distribution of synthesis evaluations by year



Source: OEV Management Information System

147. To provide a preliminary overview of the coverage of activity types in the synthesis evaluations, an analysis drawing from OEV's Management Information System was applied, recognizing the limitations of the approach adopted, described in the paragraph here below. Table 16 in Annex IV provides the indicative coverage of activity categories¹⁶⁶ present in the evaluations covered by the synthesis.

148. For DEs, activity categories were extracted from the OEV Management Information System (MIS). However, for Country Strategic Plan Evaluations, Policy Evaluations and Strategic Evaluations, activity categories were not available in the OEV MIS. Therefore, they were obtained indirectly from the topics reported in the MIS. For instance, if an evaluation addressed the topic of 'nutrition' in general terms, both the activity categories 'nutrition treatment activities' and 'malnutrition prevention activities' were applied. Likewise, if an evaluation contained the topic of 'capacity building', both the categories of 'individual capacity strengthening activities' and 'institutional capacity strengthening activities' were marked. However,

¹⁶⁶ The activity categories presented are those of WFP's Revised Corporate Results Framework (2017–2021).

in several evaluations some topics (such as ‘HIV/AIDS’, ‘purchase for progress’, ‘refugees and IDPs’ and ‘staff and human resources’) could not be converted directly into activity categories. As a result, in these cases the activity category ‘other’ was applied.

149. Table 9 below details the initial analysis of activity coverage across the 53 evaluations included in this synthesis. This analysis is presented in ranges for this report (under synthesis question 5). The activity types were verified during the data collection phase of the synthesis and are presented in the synthesis report. The most widely addressed activity is school meals, which is addressed in over half of the evaluations included in this synthesis. In comparison, fewer than 20 per cent of the evaluations appear to address asset creation and livelihood support, smallholder agricultural market support, unconditional resource transfers to support access to food, and service provision and platforms.

Table 9: Activity coverage in evaluations included in the synthesis

Proportion of evaluations in the synthesis sample that address an activity	Activities
Over 50 per cent	School meals
40–49 per cent	Institutional capacity strengthening
30–39 per cent	Malnutrition prevention
	Nutrition treatment
20–29 per cent	Individual capacity strengthening
	Emergency preparedness
	Climate adaptation and risk management
Less than 20 per cent	Asset creation and livelihood support
	Smallholder agricultural market support
	Unconditional resource transfers to support access to food
	Service provision and platform

150. The team took the uneven distribution of activity type coverage into consideration during the data analysis and reporting by weighting data by activity type where appropriate. For example, synthesis question five examines activity areas in which WFP’s approach and systems have worked better. Data in this analysis were weighted to account for the higher coverage of school feeding activity evaluations, reporting in terms of X per cent of evaluations that covered school feeding activities as opposed to Y number of total evaluations.

A3.2 SYNTHESIS QUESTIONS

151. The synthesis addressed seven main questions, which explored evaluation evidence from the country and corporate levels. During the inception phase, the team worked with OEV to fine tune the synthesis questions from the TOR.

152. The main synthesis questions were broken down into a series of sub-questions as shown below in Table 10. These sub-questions form the basis of the analytical framework, which is presented in Annex V.

Table 10: Synthesis questions and sub-questions

Synthesis question	Synthesis sub-question
1. Do corporate outcome, output and cross-cutting indicators allow for effective measurement of intervention achievements at the country level?	1.1 Does the evaluation address whether corporate outcome, output and/or cross-cutting indicators were used at the country level? If so, what evidence did the evaluation generate about use of corporate indicators at country level?
	1.2 Does the evaluation address whether corporate outcome, output and/or cross-cutting indicators were effective at measuring intervention achievements at the country level? If so, what evidence did the evaluation generate about the effectiveness of corporate indicators at country level?
2. To what extent have WFP's monitoring systems generated credible information? How has the information generated by WFP monitoring systems been used, and by whom?	2.1 Does the evaluation address whether WFP's monitoring systems generated credible information? If so, what evidence did the evaluation generate about the credibility of the monitoring information?
	2.2 Does the evaluation address the use of monitoring information? If so, what findings did the evaluation generate about whether and how monitoring information has been used, by whom?
3. To what extent has WFP's Normative Framework enabled WFP to track programme effectiveness and inform corporate performance reporting?	3.1 Does the evaluation provide evidence on developing the logframe? If so, how does the evidence relate to the Normative Framework to either track programme effectiveness and/or inform corporate performance reporting?
	3.2 Does the evaluation provide evidence on preparing the monitoring, review and evaluation plan? If so, how does the evidence relate to the Normative Framework to either track programme effectiveness and/or inform corporate performance reporting?
	3.3 Does the evaluation provide evidence on developing a monitoring toolkit? If so, how does the evidence relate to the Normative Framework to either track programme effectiveness and/or inform corporate performance reporting?
	3.4 Does the evaluation provide evidence on collecting primary data and collating secondary data? If so, how does the evidence relate to the Normative Framework to either track programme effectiveness and/or inform corporate performance reporting?
	3.5 Does the evaluation provide evidence on capturing, compiling and validating data? If so, how does the evidence relate to the Normative Framework to either track programme effectiveness and/or inform corporate performance reporting?
	3.6 Does the evaluation provide evidence on analysing data and preparing information products? If so, how does the evidence relate to the Normative Framework to either track programme effectiveness and/or inform corporate performance reporting?
	3.7 Does the evaluation provide evidence on making use of monitoring findings? If so, how does the evidence relate to the Normative Framework to either track programme effectiveness and/or inform corporate performance reporting?

Synthesis question	Synthesis sub-question
	3.8 Does the evaluation provide evidence on conducting evaluations or reviews? If so, how does the evidence relate to the Normative Framework to either track programme effectiveness and/or inform corporate performance reporting?
4. To what extent does the evaluative evidence provide learning on the three outcomes of the WFP Corporate Monitoring Strategy?	4.1 Does the evaluation provide evidence on outcome 1 (adequate monitoring expertise)? If so, how? To what extent? 4.2 Does the evaluation provide evidence on outcome 2 (financial commitment)? If so, how? To what extent? 4.3 Does the evaluation provide evidence on outcome 3 (functional capacity)? If so, how? To what extent?
5. What factors contributed to or hindered implementation of performance measurement and monitoring systems? Are there particular activity areas and contexts in which WFP's approach and systems have worked better and why?	5.1 Does the evaluation address any factors that contributed to or hindered implementation of performance measurement and monitoring systems? If so, what were they? To what extent? 5.2 Does the evaluation identify any programme activity in which WFP's performance measurement and monitoring systems worked well? If so, what were they? To what extent? 5.3 Does the evaluation address ways in which the country context contributed to or hindered the implementation of performance measurement and monitoring systems? If so, what were they? To what extent?
6. To what extent is WFP's performance measurement system aligned with national monitoring systems? How has WFP pursued opportunities to strengthen national monitoring systems?	6.1 Does the evaluation address whether WFP's performance measurement systems are aligned with national monitoring systems? If so, what evidence did the evaluation generate about monitoring system alignment? 6.2 Does the evaluation address whether WFP pursued opportunities to strengthen national monitoring systems? If so, how were they pursued?
7. To what extent were cross-cutting issues (gender equality and women's empowerment, protection, accountability to affected populations and the environment) reflected in monitoring practices, guidance and systems?	7.1 Does the evaluation address whether accountability to affected populations was reflected in monitoring practices, guidance and systems? If so, what evidence did the evaluation generate? 7.2 Does the evaluation address whether protection of affected populations was reflected in monitoring practices, guidance and systems? If so, what evidence did the evaluation generate? 7.3 Does the evaluation address whether gender equality and women's empowerment were reflected in monitoring practices, guidance and systems? If so, what evidence did the evaluation generate? 7.4 Does the evaluation address whether environmental sustainability was reflected in monitoring practices, guidance and systems? If so, what evidence did the evaluation generate?

A3.3 METHODOLOGICAL APPROACH AND DATA ANALYSIS METHODS

Data sources

153. The primary data source for the synthesis was the evaluation reports and the evaluability assessments of the CE inception reports. The synthesis draws supplementary data from KIIs and supporting documents for contextualization and verification of findings from the evaluations. A summary of data sources by synthesis question is included in Table 11 below.

Table 11: Data sources by synthesis question

Synthesis question	Data source(s)	Forms of analysis
1. Do corporate outcome, output and cross-cutting indicators allow for effective measurement of intervention achievements at the country level?	<ul style="list-style-type: none"> • Evaluation reports • Evaluability assessments • Interviews with stakeholders that interact with country-level measurement • WFP supporting documents 	<ul style="list-style-type: none"> • Quantitative – examining the frequency of findings across evaluations • Qualitative – combining descriptive findings across evaluations • Identification of best practices and lessons learned • Review and triangulation of specific documentation and interviews
2. To what extent have WFP's monitoring systems generated credible information? How has information generated by WFP monitoring systems been used, and by whom?	<ul style="list-style-type: none"> • Evaluation reports • Evaluability assessments • Interviews with stakeholders that interact with monitoring systems and information at the country, regional and HQ levels, and stakeholders using monitoring information • WFP supporting documents 	
3. To what extent has WFP's Normative Framework enabled WFP to track programme effectiveness and inform corporate performance reporting?	<ul style="list-style-type: none"> • Evaluation reports • Evaluability assessments • Interviews with stakeholders that interact with monitoring systems and information at HQ or have specific responsibilities related to the Normative Framework • WFP supporting documents, specifically the four documents of the Normative Framework 	
4. To what extent does the evaluative evidence provide learning on the three outcomes of the WFP Corporate Monitoring Strategy?	<ul style="list-style-type: none"> • Evaluation reports • Evaluability assessments • Interviews with stakeholders that interact with monitoring systems at HQ or have specific responsibilities related to the Corporate Monitoring Strategy • WFP supporting documents, specifically the Corporate Monitoring Strategy 	
5. What factors contributed to or hindered implementation of performance measurement and monitoring systems? Are there particular activity areas and contexts in which WFP's approach and systems have worked better and why?	<ul style="list-style-type: none"> • Evaluation reports • Evaluability assessments • Interviews with stakeholders that interact with monitoring systems and information at the country, regional and HQ levels • WFP supporting documents 	
6. To what extent is WFP's performance measurement system aligned with national monitoring systems? How has WFP pursued opportunities to strengthen national monitoring systems?	<ul style="list-style-type: none"> • Evaluation reports • Evaluability assessments • Interviews with stakeholders that interact with national monitoring systems • WFP supporting documents 	
7. To what extent were cross-cutting issues (gender equality and women's empowerment, protection, accountability to affected population and environment) reflected in monitoring practices, guidance and systems?	<ul style="list-style-type: none"> • Evaluation reports • Evaluability assessments • Interviews with stakeholders that interact with monitoring systems and information at the country, regional and HQ levels or with specific responsibility for cross-cutting issues • WFP supporting documents on cross-cutting issues 	

Evaluation data extraction

154. The process for reviewing, extracting, and coding evaluation reports and evaluability assessments of inception reports followed a four-part process:

- i. Development of the analytical framework and initial coding system
- ii. Pilot review of one report and related revisions to framework, followed by a second pilot review of one report
- iii. Review of reports using an iterative process to coding and quality assurance
- iv. Extraction of emerging findings and lessons

155. First, the team developed the analytical framework and initial coding system during the inception phase. The analytical framework is structured around the synthesis questions, sub-questions, codes and descriptions of codes. The sub-questions and codes were developed to facilitate consistent extraction of data from the evaluations, based on the initial review of evaluations during the scoping phase and by reviewing supporting documentation. For example, the codes related to synthesis question 3 were developed after review of the components of the Normative Framework.

156. Once the initial analytical framework was approved, the coding system was loaded into the MAXQDA data analysis software. MAXQDA enables a traceable analysis across reports to produce insights and evidence for the synthesis.

157. Second, the team conducted a pilot review to test the validity of the codes and to ensure inter-rater reliability for data extraction. All team members read the same report and extracted and tagged segments in MAXQDA according to the coding framework. The coding was reviewed by the Team Leader, and the team then met to discuss any areas in which team members applied coding differently. The team then undertook a second pilot review, and the Team Leader reviewed the application of the coding framework across the team to ensure consistent application. This was followed by another team meeting to discuss questions related to the understanding and application of the updated coding framework.

158. Third, evaluation reports were assigned to individual team members. The team reviewed and coded the first 26.92 per cent (14 evaluation reports, including the finalization of the second pilot) of the sample. Once this was complete, the Team Leader led QA on 14.29 per cent (two reviews) of the completed reviews. This involved re-reading a report and reviewing the coding to ensure it was being consistently applied. After this the team came together to review the data and discuss the emerging themes and lessons. The team then discussed whether existing codes needed to be refined, sub-themes included, or new codes added. Following this, the next 46.15 per cent (24 evaluation reports) of the sample were reviewed and the process of QA and reviewing the coding framework repeated. The Team Leader undertook QA on 16.67 per cent (four reviews) of the reviews. Then, the remaining 26.92 per cent (14) of the sample was reviewed.

159. This iterative approach to coding and QA allowed the team to systematically build a credible and consistent evidence base against the predefined themes linked to the synthesis questions (deductive), while also enabling emerging themes to be identified (inductive).

160. Fourth, a long list of emerging findings and lessons was produced. This helped to provide the foundation for the KIs and wider document review, which were undertaken alongside the analysis process and subsequent to it. The emerging findings will be validated with OEV.

Secondary data collection

161. After the data extraction and coding process was completed, the team drafted interview guides and protocols based on emerging themes. The team approached stakeholders for interviews in collaboration with OEV. The purpose of the interviews was to corroborate the findings, explore issues in more depth and help explain the patterns identified through the evaluation data extraction process. Depending on the stakeholder, the team also sought their views on common topics emerging from the primary data collection stage of the synthesis, or asked them for details on specific topics included in an evaluation that they had knowledge of. A list of persons interviewed is available in Annex VIII.

162. With the support of OEV, the synthesis team also conducted a detailed review of relevant WFP documentation, including management responses to all CEs in the sample, to enhance the understanding of the broad context for performance measurement and monitoring within WFP. A full list of resources is included as Annex IX.

Data analysis

163. The team then collated the data collected from the evaluation reports, the KIIs and the document review against the synthesis questions and sub-questions. The team analysed the data extracted by examining the findings against each synthesis question. Analysis was both quantitative (examining the frequency of findings across evaluations) and qualitative (combining descriptive findings across evaluations). Some of these findings have been weighted for the quantitative analysis. For example, school feeding activities appeared to be overrepresented in the decentralized activity evaluations.¹⁶⁷ This has been taken into consideration for analysis across report types.

164. The team looked for themes and trends, including similarities, divergences and contradictions in the findings, within and across questions. The team also looked for illustrative examples from the evaluations to highlight as best practices in the synthesis report.

165. After initial analysis, the team then met to validate findings across the entire sample and to draft the conclusions, lessons learned and recommendations.

Limitations

166. The main limitations encountered speak to the following issues:

- Great variability across the sample in the depth of evidence available in relation to the different synthesis questions. This has resulted in much greater reliance on and reference to some evaluation reports compared to others. This is linked to some variability in the evaluation focus and questions asked (for example in DEs) and to the depth and sophistication of the evaluative analysis provided on issues relating to performance measurement and monitoring in each report.
- Perceived bias in the evaluations towards highlighting more prominently examples of shortcomings and gaps relating to monitoring approaches, data, and systems, compared to elaborating on positive examples, and drawing lessons on what works well. This has limited the ability of the synthesis team to showcase more in detail positive examples in that regard.
- Finally, except for country capacity strengthening (CCS) and resilience, limitations emerged in the availability of evidence around successes or shortcomings that were attributed in the evaluation reports specifically to the activity type.

167. In order to mitigate those limitations, the synthesis team supplemented the evaluations with information from supporting documents (corporate and policy/programme-specific) and KIIs for validation purposes where possible.

Reporting

168. In the reporting stage, the team produced a draft synthesis report. The synthesis report has undergone three main rounds of validation and feedback to ensure both its validity and its coherence with the WFP institutional environment. Initial feedback on the draft synthesis report was provided by OEV. A revised draft will be shared with the Internal Reference Group for their comments to provide validation and refinement.

169. OEV will facilitate a virtual workshop (or a hybrid one depending on the evolution of COVID-19, with OEV, CPP and RAM in person and the team virtually) with key stakeholders after the draft synthesis report is circulated for comments. The workshop will be an opportunity to further validate the synthesis' findings, emerging lessons and conclusions, and provide inputs to the proposed recommendations to ensure their focus and targets are appropriate. The team will incorporate these sources of feedback into the final synthesis report.

¹⁶⁷ A description of indicative activity coverage in synthesis evaluations is included in section 2.1.

170. Gender equality and women's empowerment have been addressed in the synthesis as a cross-cutting issue. Findings from the evaluations on how gender is reflected in monitoring practices, guidance and systems were collected during the report data extraction process, and have been included in the synthesis report.

A3.4 ETHICAL CONSIDERATIONS

171. The team recognized that all evaluations (including evaluation syntheses, where relevant) must conform to 2020 United Nations Evaluation Group (UNEG) ethical guidelines. As the firm conducting the synthesis, Itad has been responsible for safeguarding and ensuring that ethical approaches are taken into account at all stages of the evaluation synthesis. In the context of this evaluation synthesis, therefore, the main ethical issues the team foresaw related primarily to ensuring the confidentiality of the interviewees and data protection.

172. The KIs were used to clarify and validate findings from the evaluation reports, such as in relation to failures or successes of internal processes, systems and practices. By nature, some of these issues are considered sensitive. The integrity of the evaluation synthesis requires that interviewees feel comfortable expressing their opinions, without fear of reprisal. As such, the team has protected interviewees' right to confidentiality by:

- Asking stakeholders if they consent to interviews being recorded.
- Paraphrasing findings gathered through the KIs, with no attributable quotes used in reports.
- Using stakeholder group categories where relevant for reporting, never specific titles (for example, 'country office stakeholder', 'stakeholder from the Regional Bureau', 'senior manager'). This will ensure that no findings can be traced back to a specific individual.

173. The team started all interviews stating the interviewees' right to confidentiality and outlining the above points and proceeded in line with Itad's Ethical Principles.

174. Additionally, Itad takes its responsibility for data protection extremely seriously and is certified by IASME¹⁶⁸ for the UK Government's Cyber Essentials Scheme. Itad has put in place a series of policies to establish and maintain the security and confidentiality of information, information systems, applications and networks owned or held by Itad. The team recognized that effective data management is essential to ensure the confidentiality of data, and consistency and quality across the lifetime of the project. For this reason, the team utilized a carefully selected data management system to: (i) store and organize data efficiently; (ii) securely access data at all times and locations; (iii) guarantee the confidentiality and sensitivity of data; (iv) treat the data for analysis; (v) ensure cross-learning among the team; and (vi) consolidate the team's contributions to the stakeholders' workshop. The Project Manager acted as the sole administrator for Itad's data management system with the exclusive ability to grant permissions and oversee the access for team members, which was granted on a 'need to know' basis.

175. These issues were monitored and managed during the implementation of the synthesis. No additional ethical issues arose during the completion of the synthesis.

A3.5 RISKS AND ASSUMPTIONS

176. In Table 12, we outline the potential limitations and risks to the efficient conduct of the synthesis evaluation, which the team has mitigated through the design and careful implementation of the synthesis.

¹⁶⁸ IASME Governance is an Information Assurance standard.

Table 12: Risk matrix (high (H), medium (M), low (L))

Potential challenge	Risk	Proposed solution	Residual risk
The inclusion/exclusion criteria for selecting the evaluation universe significantly limit the sample of eligible evidence (that is, far fewer than the assumed 50 evaluation reports), which limits the strength and breadth of evidence available for the synthesis	M	Our Team Leader, Liz Mason, is an expert at defining synthesis parameters and synthesis design. Liz has overseen this process to ensure any exclusions are reasonable and in line with the methodology and WFP's expectations. The team kept the selection criteria relatively broad to avoid excluding too much evidence. Additional data collection, through the review of key documents and KIIs with WFP stakeholders, enabled the team to address any gaps in evidence where necessary.	L
The team is inconsistent in its application of the coding framework, resulting in the evidence support for common themes not being as sound as it should be, thus limiting the utility of findings and related recommendations	M	Our Team Leader, Liz Mason, brings experience of leading large (15+) teams of researchers in evaluation synthesis and ensuring their consistency in coding. The team also used an iterative approach to coding that built a shared approach and understanding to coding across the team to further mitigate this risk. This was enhanced through regular team check-ins to surface any questions about the application of the coding framework, in addition to a robust QA process that specifically examined the consistency of coding.	L
The team identifies common findings and lessons that are overly general, and therefore do not provide the basis for relevant and actionable recommendations	M	<p>The iterative approach to coding enabled the team to systematically build a credible and consistent evidence base against the predefined themes linked to the synthesis questions (deductive), while also allowing for emerging themes to be identified (inductive).</p> <p>The additional phased data collection enabled further triangulation through document review and KIIs.</p> <p>The team was carefully selected to bring relevant sector- and theme- specific knowledge, with an in-built familiarity with WFP, to help identify nuance. This was further supported by insights from OEV and IRG members.</p> <p>The team scheduled regular check-ins with WFP, and when relevant shared emerging findings to check that the lessons that were being identified were relevant and useful.</p> <p>A stakeholder workshop to validate the findings, conclusions and recommendations helped to ensure that the final report included relevant and actionable recommendations.</p>	L
Disruption to project timelines due to COVID-19 (for example through inability to schedule KIIs with key stakeholders)	H	Itad is proactively managing the risks COVID-19 presents with clients, and brings prior experience of this to this evaluation synthesis. Itad also engaged with WFP staff and stakeholders early in the project to ensure that the value of the synthesis was understood and possible stakeholders were sufficiently engaged.	M

A3.6 QUALITY ASSURANCE

177. WFP has developed a Centralized Evaluation Quality Assurance System (CEQAS) for Evaluation synthesis based on the UNEG norms and standards and good practice of the international evaluation community (Active Learning Network for Accountability and Performance and the Development Assistance Committee of the Organisation for Economic Co-operation and Development). This sets out process maps with in-built steps for QA and templates for evaluation products. It also includes checklists for feedback on quality for each of the evaluation products. CEQAS was systematically applied during this evaluation synthesis and relevant documents were provided to the team.

178. The Project Director was responsible for ensuring technical excellence and the credibility of the evaluation synthesis by operationalizing Itad’s four-stage QA process and CEQAS (where applicable), in close coordination with the team. The Project Director was responsible for assuring the quality of all evaluation synthesis outputs before these were shared with WFP.

179. The synthesis team was also responsible for ensuring the quality of data (validity, consistency, and accuracy) throughout the analytical and reporting phases. The steps in Table 13 below provide an overview of the steps that were taken to ensure the high quality of the evaluation synthesis.

Table 13: Quality assurance steps that have been applied in this synthesis

Evaluation synthesis phases	QA steps and how they were applied
Phase 1: Inception and synthesis design	<ul style="list-style-type: none"> • After ranking the evaluations for the scoping exercise to set the final evaluation universe, our TL reviewed all English language evaluations that were excluded to ensure that the selection criteria were appropriately applied and no relevant evidence was missed. The TL worked with the team’s language specialist to review the excluded Spanish and French evaluations. • A final list of evaluations included and excluded from the sample, and the justifications, were provided to WFP for review and validation (see Annex IV). This provided transparency to WFP on the selection of the evidence for the synthesis. • The Scoping Note and Inception Report were quality assured by the Itad Project Director to ensure the proposed methodology was realistic and credible and would provide a robust response to the EQs.
Phase 2: Review and synthesis	<ul style="list-style-type: none"> • The Team Leader developed the full analytical framework that linked clearly to the evaluation sub-questions, which was reviewed by the Project Director. • The team was trained by the Team Leader on the MAXQDA coding protocol. • All members of the team conducted a pilot review. This involved reviewing a report, highlighting relevant text and coding it using MAXQDA. Reviews were then reviewed by the Team Leader. Following this, the team met to discuss the process, the analytical and coding framework, and whether any adaptation or further guidance was required. • The team then took an iterative approach to reviewing the evaluation universe, with continuous QA processes led by the Team Leader as already mentioned.
Phase 3: Reporting and communications	<ul style="list-style-type: none"> • After the initial round of analysis, the team produced a list of emerging themes. These were discussed on an ongoing basis with WFP before the submission of the first draft of the report and subsequent drafts. • The team validated the recommendations with WFP to help ensure that they were practical and actionable. • The final synthesis report was quality assured by the Itad Project Director to ensure it presents a clear, logical, evidence-based argument, with a strong link between evidence, findings, conclusions, lessons and recommendations.

180. The team has not identified any conflicts of interest relevant to this evaluation synthesis and none were identified during implementation.

Annex IV: Scoping note

A4.1 SCOPING PROCESS

181. This annex details the scoping process and findings. The universe of evaluations that could be included in this synthesis was identified by OEV during the preparation phase of the synthesis. The team then conducted a scoping exercise during the inception phase to determine the final sample of evaluations to be included in the synthesis.

182. During the preparation phase, OEV identified a preliminary list of 68 evaluations based on evaluation types, time period, and the quality of the evaluation (see [Annex I](#) Summary Terms of Reference). OEV identified an additional 7 evaluations that met the same criteria but that were approved after the initial TOR was published, bringing the total number of evaluations to 75.

183. During the inception phase, the evaluation synthesis team developed a scoping approach and validated it through consultations with OEV. The approach included six phases, which are further detailed in the body of this report:

- i. Identifying key information about each report including the commissioner, type of evaluation, language, approval year and PHQA score
- ii. Establishing a set of topics related to performance measurement and monitoring, such as staff capacity; strategy, guidance and tools; monitoring budgeting, planning and execution; monitoring systems; data quality; reporting; follow-up and utilization of data; monitoring and M&E; theory of change; indicators, outputs and outcomes; logframe; results framework; and data collection
- iii. Reviewing the reports and coding the conclusions and recommendations for relevance to performance measurement and monitoring
- iv. Ranking the relevance of the evaluations to the synthesis as high, medium and low, based on the number of relevant segments and the overall strength of evidence
- v. Conducting a quality review of rankings to verify whether the evaluations with a ranking of medium should be included in the synthesis
- vi. Making recommendations for the inclusion of evaluations in the synthesis

A4.2 SCOPING FINDINGS

184. Of the 75 evaluations identified by OEV, the team ranked the coverage of 53 as 'high', 6 as 'medium', and 17 as 'low'. The breakdown of performance measurement and monitoring coverage by evaluation type is shown below in Table 14. Examples of reports with 'high' coverage are detailed in Box 1 below.

Table 14: Performance measurement and monitoring coverage by evaluation type

Evaluation type		Shortlisted	Performance measurement and monitoring coverage		
			High	Medium	Low
Centralized evaluations	Country Strategic Plan	12	12	-	-
	Policy	3	3	-	-
	Strategic	6	6	-	-
	<i>Subtotal</i>	<i>21</i>	<i>21</i>	-	-
Decentralized evaluations	Activity	48	27	6	15
	Thematic	4	3	-	1
	Transfer Modality	2	2	-	
	<i>Subtotal</i>	<i>54</i>	<i>32</i>	<i>6</i>	<i>16</i>
Total		75	53	6	16

Box 1: Examples of reports with 'high' performance measurement and monitoring coverage

Evaluation of The Gambia WFP Country Strategic Plan 2019–2021. Centralized Evaluation Report. October 2021. The evaluation includes **conclusions** on the adequacy of the M&E system and reporting. There are several **recommendations** including developing theories of change, strengthening the M&E system and improving reporting of outputs and outcomes. Relevant **findings** cover M&E capacity, quality of indicators, alignment of indicators with the CRF and availability of outcome data.

General Food Assistance and School Feeding Programmes, Libya 2017–2019. Decentralized Evaluation. March 2021. **Conclusions** include those related to monitoring capacities, adequacy of indicators (including on gender), data collection and reporting. **Recommendations** include the need to incorporate qualitative data collection tools and indicators to capture the impact on beneficiaries including better monitoring of gender results; strengthening the M&E system; mainstreaming conflict sensitivity in the log frame; building partners capacities in reporting; and developing a theory of change. Relevant **findings** cover adequacy of monitoring and reporting, availability of monitoring data and training on monitoring.

Strategic Evaluation of WFP Support for Enhanced Resilience. January 2021. **Conclusions** include the quality of reporting, especially at the outcome level, and the ability to measure and articulate the contribution of programmes to results related to resilience. **Recommendations** include consolidating performance measurement data from resilience-related initiatives for corporate reporting; developing a results framework aligned to the Corporate Results Framework; strengthening the WFP's ability to collect, collate and analyse information on shocks before they happen; and supporting the generation of evidence on the relevance of food security and resilience interventions in conflict and protracted crises. The **findings** include assessments of: the adequacy of indicators; methodologies used for measuring resilience (including the effects on women) and corporate tools used to collect such data; available capacity on monitoring (time, cost and so on); and effectiveness of the monitoring approach.

185. In consultation with OEV, the team recommended that the 53 evaluations ranked as 'high' be included in the synthesis. These evaluations directly addressed performance measurement and monitoring topics. The 6 evaluations ranked as 'medium' by the team had some evidence surrounding these topics but were not detailed enough to be of interest for the synthesis. The 16 evaluations identified as 'low' either did not address these topics or only briefly addressed them. A full list of the component evaluations and rationale for inclusion or exclusion is included in Table 15.

186. All 21 of the CEs directly addressed performance measurement and monitoring. These evaluations should provide evidence at both the corporate and country levels, as well as at the strategic and policy evaluations. For the 54 DEs, 32 (59 per cent) of the evaluations sufficiently addressed performance measurement and monitoring to be of interest for this synthesis. The DEs will most likely provide evidence at the country and/or regional level.

187. Table 15 details the 75 evaluations reviewed by the team at the scoping stage, their rankings for performance measurement and monitoring (PMM) coverage, and notes about the rankings.

188. To provide a preliminary overview of the coverage of activity types in the synthesis evaluations, an analysis drawing from OEV's MIS was applied, recognizing the limitations of the approach adopted, described in the paragraph here below. Table 16 provides the indicative coverage of activity categories present in the evaluations covered by the synthesis.

189. For DEs, activity categories were extracted from the OEV MIS. However, for Country Strategic Plan Evaluations, Policy Evaluations and Strategic Evaluations, activity categories were not available in the OEV MIS. Therefore, they were obtained indirectly from the topics reported in MIS. For instance, if an evaluation addressed the topic of 'nutrition' in general terms, both the activity categories 'nutrition treatment activities' and 'malnutrition prevention activities' were applied. Likewise, if an evaluation contained the topic 'capacity building', both the categories 'individual capacity strengthening activities' and 'institutional capacity strengthening activities' were marked. However, in several evaluations some topics (such as 'gender', 'HIV/AIDS', 'purchase for progress', 'refugees and IDPs' and 'staff and human resources') could not be converted directly in an activity category. As a result, in these cases the activity category 'other' was applied.

Table 15: Evaluation ranking for scoping

Evaluation category	Evaluation type	Title of the report	Conclusions	Recommendations	PMM coverage	Comments
Centralized	CSP	Evaluation of Bangladesh WFP Country Strategic Plan 2016–2019	X	X	High	5 conclusions, 8 recommendations, several relevant references in findings
		Evaluation of Cameroon WFP Country Strategic Plan 2018–2020	X	X	High	6 conclusions, 4 recommendations, several relevant references in findings
		Evaluation of Democratic Republic of the Congo Interim Country Strategic Plan 2018–2020	X	X	High	2 conclusions, 5 recommendations, several relevant references in findings
		Evaluation of Indonesia WFP Country Strategic Plan 2017–2020	X	X	High	5 conclusions, 8 recommendations, a few relevant references in findings
		Evaluation of Timor-Leste WFP Country Strategic Plan 2018–2020	X	X	High	2 conclusions, 1 recommendation, several relevant references in findings
		Evaluation of China WFP Country Strategic Plan (2017–2021)	X	X	High	4 conclusions, 3 recommendations, several relevant references in findings
		Evaluation of El Salvador WFP Country Strategic Plan 2017–2021	X	X	High	1 conclusion 4 recommendations; several relevant references in findings
		Evaluation of Honduras WFP Country Strategic Plan 2018–2021	X	X	High	4 conclusions, 2 recommendations, several relevant references in findings
		Evaluation of Laos WFP Country Strategic Plan Evaluations (2017–2021)	X	X	High	3 conclusions, 2 recommendations, several relevant references in findings
		Evaluation of Lebanon WFP Country Strategic Plan (2018–2021)	X	X	High	3 conclusions, 10 recommendations, several relevant references issues in findings
		Evaluation of the Gambia WFP Country Strategic Plan (2018–2021)	X	X	High	4 conclusions, 4 recommendations, several relevant references in findings
	Evaluation of Zimbabwe WFP Country Strategic Plan 2017–2021	X	X	High	4 conclusions, 8 recommendations, several relevant references in findings	
	Policy	Update of WFP's Safety Nets Policy	X	X	High	4 conclusions, 4 recommendations, several relevant references in findings
Evaluation of the Gender Policy		X	X	High	3 conclusions, 7 recommendations, several relevant references in findings; highly relevant for gender cross-cutting issues	

Evaluation category	Evaluation type	Title of the report	Conclusions	Recommendations	PMM coverage	Comments
		Evaluation of the WFP South-South and Triangular Cooperation Policy	X	X	High	7 conclusions, 5 recommendations, several relevant references in findings
	Strategic	Strategic Evaluation of the Pilot Country Strategic Plans	X	X	High	2 conclusions, 13 recommendations, several relevant references in findings; highly relevant in regard to higher-level performance measurement issues
		Strategic Evaluation of WFP's Support for Enhanced Resilience	X	X	High	4 conclusions, 3 recommendations, several relevant references in findings
		Strategic Evaluation of WFP's Capacity to Respond to Emergencies	X	X	High	4 conclusions, 3 recommendations, several relevant references in findings; highly relevant for monitoring in emergency situations
		Strategic Evaluation of the Contribution of School Feeding Activities to the Achievement of the Sustainable Development Goals	X	X	High	8 conclusions, 8 recommendations, several relevant references in findings; good resource to tie in findings about school feeding beyond the activity level
		Joint Evaluation on the Collaboration among the United Nations Rome-based Agencies	X		High	1 conclusion, 0 recommendations; few direct conclusions or recommendations, but provides unique information in the findings on WFP compared to the other agencies.
		Strategic Evaluation of the WFP's Use of Technology in Constrained Environments	X	X	High	1 conclusion, 1 recommendation, addresses use of technology that may not be included in other reports
Decentralized	Activity	Midterm Evaluation of McGovern-Dole Funded School Feeding Project in Guinea-Bissau (January 2016 – June 2018)	X	X	High	2 conclusions, 1 recommendation, several relevant references in findings
		Global End-term Evaluation of the Joint Programme on Accelerating Progress towards the Economic Empowerment of Rural Women in Ethiopia, Guatemala, Kyrgyzstan, Liberia, Nepal, Niger and Rwanda from 2014 to 2020	X	X	High	2 conclusions, 6 recommendations, several relevant references in findings

Evaluation category	Evaluation type	Title of the report	Conclusions	Recommendations	PMM coverage	Comments
		Decentralized Evaluation of the Results of WFP's Food Assistance to Temporarily Dislocated Persons in Pakistan from 2015–2017_FR	X	X	Medium	3 conclusions, 2 recommendations. Not much information in the findings around PMM.
		End-line Evaluation of United States Department of Agriculture (USDA) McGovern-Dole Grant Food for Education (FFE) Programme for WFP Cambodia 2013–2016	X		Low	2 conclusions, 0 recommendations, limited information in findings. Would not recommend including due to low coverage of topic.
		Evaluation of the DG ECHO funded Emergency Social Safety Net (ESSN) in Turkey	X	X	High	3 conclusions, 2 recommendations, several relevant references in findings
		Evaluation of the National School Feeding Programme in Lesotho , in consultation with the Lesotho Ministry of Education and Training	X	X	High	2 conclusions, 3 recommendations, several relevant references in findings
		Evaluation of the Nutrition Components of the Algeria PRRO 200301	X	X	High	1 conclusion, 2 recommendations, several relevant references in findings
		Final Evaluation of Disaster Preparedness and Response/Climate Change Adaptation Activities under the Office of Foreign Disaster Assistance Fund in the Philippines May 2011 to September 2017_FR	X	X	High	4 conclusions, 4 recommendations, several relevant references in findings
		Final Evaluation of McGovern-Dole-supported School Feeding Programme in Bangladesh (FFE-388-2014/048-00)	X	X	High	1 conclusion, 4 recommendations, several relevant references in findings
		Final Evaluation of the School Meals Programme in Malawi with support from USDA, Government of Brazil and United Kingdom 2014 to 2016	X	X	High	2 conclusions, 2 recommendations. The evaluation does not have a high level of coverage, but may be of interest to the synthesis because it has observations about joint monitoring across two different agency programmes (WFP and FAO).
		Final Evaluation of WFP'S USDA McGovern-Dole International Food for Education and Child Nutrition Programme's Support in Afar and Somali Regions in Ethiopia 2013–2017	X	X	High	4 conclusions, 4 recommendations, several relevant references in findings
		WFP's General Food Assistance to Syrian Refugees in Jordan 2015 to mid-2018		X	Low	0 conclusions, 4 recommendations, limited information in findings. Would not recommend including due to low coverage of topic.

Evaluation category	Evaluation type	Title of the report	Conclusions	Recommendations	PMM coverage	Comments
		Activités du Programme Alimentaire Mondial (PAM) de renforcement des capacités pour l'amélioration du Programme National d'Alimentation Scolaire en Tunisie de 2016 à 2018		X	Low	1 conclusion, 1 recommendation, no monitoring system used in project. No additional insights
		End-line Evaluation of the Target Public Distribution Reforms Project in Bhubaneswar (Odisha) 2014-2019_India	X		Low	1 conclusion, 0 recommendations, almost no information on PMM
		Evaluación final del Proyecto 'Respuesta al fenómeno de El Niño en el Corredor Seco', El Salvador, Guatemala, Honduras y Nicaragua, 2016-2018	X	X	Low	1 conclusion, 1 recommendation, limitation of the indicators, logframe, theory of change mentioned but no further analysis
		Évaluation des activités de renforcement des capacités institutionnelles dans le domaine de l'alimentation scolaire au Togo	X	X	Medium	3 conclusions, 1 recommendation, annex contains suggestions on how to improve the quality of logframe, but information in the report is limited
		Évaluation des programmes intégrés de cantines scolaires financés par l'Ambassade des Pays Bas (provinces Bubanza, Bujumbura rural et Cibitoke) et par l'Union européenne (province Gitega) et mis en oeuvre par le PAM au Burundi 2016 à 2018	X		Low	0 conclusions, 2 recommendations, performance measurement and monitoring are only marginally addressed in the report.
		Evaluation of the Eswatini National School Feeding Programme 2010-2018	X	X	High	3 conclusions, 6 recommendations, several relevant references in findings
		Midterm Evaluation of Integrated Risk Management and Climate Services Programme in Malawi from 2017-2019	X	X	High	2 conclusions, 2 recommendations, several relevant references in findings
		WFP's USDA McGovern-Dole International Food for Education and Child Nutrition Programme's Support in Rwanda 2016-2020 Report: Midterm Evaluation	X	X	High	3 conclusions, 1 recommendation, limited information in the conclusions and, but there is an evaluation question on the monitoring system and related findings
		« Projet lait » au sein du programme d'alimentation scolaire du PAM dans la région du Sahel, Burkina Faso, de 2017 à 2019			Low	0 conclusions, 0 recommendations, performance measurement and monitoring marginally addressed in evaluation
		Contribution des cantines scolaires aux résultats de l'éducation dans le sud de Madagascar (2015 à 2019)	X	X	High	3 conclusions, 3 recommendations, shortcomings in monitoring system are identified throughout the report

Evaluation category	Evaluation type	Title of the report	Conclusions	Recommendations	PMM coverage	Comments
		: Une analyse de la contribution De janvier 2015 à juin 2019				
		End-line Evaluation of United States Department of Agriculture (USDA) McGovern-Dole Grant Food for Education Programme for WFP Cambodia FY 2017–2019	X	X	Medium	1 conclusion, 1 recommendation, limited relevant references in findings
		Evaluation conjointe à mi-parcours du Programme National d’Alimentation Scolaire Intégré (PNASI) Aout 2017 – Mai 2019_Benin	X	X	High	3 conclusions, 4 recommendations, several relevant references in findings
		Evaluation of Namibia National School Feeding Programme		X	High	0 conclusions, 5 recommendations, several relevant references in findings
		Evaluation of the Joint Programme for Girls Education with financial support from the Norwegian Government July 2014 – October 2017_Malawi		X	High	0 conclusions, 3 recommendations, several relevant references in findings including an evaluation question about PMM
		Evaluation of USDA’s Local and Regional Food Aid Procurement Programme (Rwanda 2017–2019)	X		Low	1 conclusion, 0 recommendations, limited relevant references in findings around appropriateness of indicators selected for the activity but not much detail.
		Evaluation Series on Emergency School Feeding in Lebanon 2015–2019	X	X	Medium	1 conclusion, 1 recommendation, limited relevant references in findings with few details provided
		Evaluation Series on Emergency School Feeding in Niger	X	X	High	2 conclusions, 2 recommendations, several relevant references in findings
		Evaluation Series on Emergency School Feeding in Syria	X	X	High	5 conclusions, 10 recommendations, several relevant references in findings
		Evaluation Series on Emergency School Feeding in the Democratic Republic of Congo 2015–2019	X	X	Low	1 conclusion, 1 recommendation, performance measurement and monitoring are only marginally addressed in the report
		FINAL EVALUATION OF THE PROGRAMME ‘Accelerate Progress Towards Millennium Development Goal 1C (MDG1.C Programme)_Mozambique	X	X	High	2 conclusions, 2 recommendations, several relevant references in findings

Evaluation category	Evaluation type	Title of the report	Conclusions	Recommendations	PMM coverage	Comments
		Final evaluation of the USDA-supported Local and Regional Procurement project in Kenya	X	X	Medium	1 conclusion, 2 recommendations, some relevant references in findings but only in passing.
		Final evaluation of WFP Haiti's Food for Education and Child Nutrition Programme (2016–2019)	X		Low	1 conclusion, 0 recommendations, limited relevant references in findings
		Midterm Evaluation of WFP School Feeding USDA Mc Govern Dole Grant for FY 2017–2020 in Bangladesh	X	X	Medium	2 conclusions, 2 recommendations, limited relevant references in findings
		WFP Livelihoods and Resilience Activities in Lebanon 2016 – 2019	X	X	High	4 conclusions, 2 recommendations, several relevant references in findings
		WFP's relief food and cash assistance for conflict-affected people in Kachin and northern Shan January 2016 – December 2019, Myanmar		X	Low	0 conclusions, 1 recommendation, limited relevant references in findings
		Addressing Climate Change Impacts on Marginalized Agricultural Communities Living in the Mahaweli River Basin of Sri Lanka 2013 – 2020	X	X	High	4 conclusions, 1 recommendation, several relevant references in findings
		Contribution du Programme Alimentaire Mondial au Système de Protection Sociale Adaptative en Mauritanie depuis 2018		X	Low	0 conclusions, 3 recommendations, limitations in the logframe and monitoring are mentioned but not analysed in-depth
		End-line Evaluation of USDA Local Regional Procurement project in Nalae District, Luang Namtha Province in Lao PDR [FY 16-19]		X	High	0 conclusions, 4 recommendations, 1 lesson learned, several relevant references in findings
		Evaluación del modelo de descentralización del Programa Nacional de Alimentación Escolar (PNAE) 2016–2019, Honduras	X	X	High	4 conclusions, 2 recommendations, a few relevant references in findings
		Evaluation of the Food Assistance for Assets in the Context of Malawi 2015–2019			Low	0 conclusions, 0 recommendations, limited relevant references in findings. Monitoring data is used in evaluation, a few sentences are included about challenges for monitoring when scaling up activity, but are not dealt with in detail.
		Final Evaluation of Enhanced Nutrition and Value Chains (ENVAC) project 2016–2021, Ghana	X	X	High	2 conclusions, 5 recommendations, several relevant references in findings
		Final Evaluation of McGovern-Dole International Food for Education and Child Nutrition Programme in Guinea-Bissau 2016–2019	X	X	High	2 conclusions, 2 recommendations, several relevant references in findings

Evaluation category	Evaluation type	Title of the report	Conclusions	Recommendations	PMM coverage	Comments
		General Food Assistance and School Feeding Programmes, Libya	X	X	High	3 conclusions, 4 recommendations, several relevant references in findings
		Midterm Evaluation of Nutrition Activities in The Gambia 2016–2019	X	X	High	3 conclusions, 3 recommendations, several relevant references in findings
		Programme Activity Evaluation of Food Assistance for Assets Project in South Sudan March 2016 to December 2019			Low	0 conclusions, 0 recommendations, limited relevant references in findings
		Support for Strengthening Resilience of Vulnerable Groups in Ethiopia : The Fresh Food Voucher Programme Expansion in Amhara Region January 2018 to December 2020			Low	0 conclusions, 0 recommendations, evaluation used activity-monitoring data in the report but did not discuss any relevant topics
	Thematic	Evaluación final de la relevancia del rol y la respuesta del PAM para avanzar en un enfoque de asistencia alimentaria vinculado a los sistemas de protección social en Ecuador	X		Low	0 conclusions, 1 recommendation, little related analysis
		Evaluación de género del Plan Estratégico de País de El Salvador (2017–2021)	X	X	High	1 conclusion, 3 recommendations, several relevant references in findings
		Evaluation of the Intervention for the Treatment of Moderate Acute Malnutrition in Ngozi, Kirundo, Cankuzo and Rutana 2016–2019_ Burundi	X	X	High	2 conclusions, 3 recommendations, some relevant references in findings
		Evaluation thématique sur les questions de genre dans les interventions du PAM au Burkina Faso (2016–2018)	X	X	High	2 conclusions, 3 recommendations, several relevant references in findings
	Transfer modality	An evaluation of the effects and a cost benefit analysis of the General Food Distribution Cash Modality scale-up (Cash-Based Transfers for PRRO 200737) for refugees and host communities in Kenya	X	X	High	2 conclusions, 4 recommendations, several relevant references in findings
		Evaluation décentralisée de la modalité transfert monétaire utilisée dans le programme de cantines scolaires appuyé par le PAM au Sénégal Rapport d'évaluation Juillet 2018. Version Finale	X	X	High	2 conclusions, 4 recommendations, several relevant references in findings, assessment of monitoring limitations and quality of indicators

Table 16: Indicative coverage of activities by report for evaluations included in the final sample

Features of the evaluations				Activity categories ¹⁶⁹ as per Annex I of the revised Corporate Result Framework (2017–2021) ¹⁷⁰												
Evaluation by country	Evaluation type	Year of approval	PHQA	1. URT	2. ACL	3. CAR	4. SMP	5. NTA	6. NPA	7. SMS	8. CSB	9. CSI	10. CPA	11. EPA	12. AAA	13. OTH
Bangladesh	CSP	2020	94		✓	✓	✓	✓	✓		✓	✓		✓		✓
Cameroon	CSP	2020	82		✓		✓	✓	✓		✓	✓		✓		✓
Democratic Republic of the Congo	CSP	2020	84		✓		✓	✓	✓		✓	✓		✓		✓
Indonesia	CSP	2020	89				✓	✓	✓		✓	✓		✓		
Timor-Leste	CSP	2020	91				✓	✓	✓		✓	✓				
China	CSP	2021	83			✓		✓	✓		✓	✓				✓
El Salvador	CSP	2021	93			✓		✓	✓		✓	✓		✓		

¹⁶⁹ Activity categories are listed in [WFP Revised Corporate Results Framework \(2017–2021\)](#). For Decentralized Evaluations, activity categories were extracted from OEV MIS. However, for Country Strategic Plan Evaluations, Policy Evaluations and Strategic Evaluations, activity categories were not available in the OEV tool. Therefore, they were obtained indirectly from the topics reported in MIS. For instance, if an evaluation addressed the topic of ‘nutrition’ in general terms, both the activity categories ‘nutrition treatment activities’ and ‘malnutrition prevention activities’ were ticked in the table. Likewise, if an evaluation contained the topic of ‘capacity building’, both the categories of ‘individual capacity strengthening activities’ and ‘institutional capacity strengthening activities’ were marked. However, in several evaluations a number of topics (such as, ‘gender’, ‘HIV/AIDS’, ‘purchase for progress’, ‘refugees and IDPs’ and ‘staff and human resources’) could not be converted directly in an activity category. As a result, in these cases the activity category ‘other’ was applied.

¹⁷⁰ Activity categories covered by the evaluations are the following: 1. unconditional resource transfers to support access to food (URT); 2. asset creation and livelihood support activities (ACL); 3. climate adaptation and risk management activities (CAR); 4. school meal activities (SMP); 5. nutrition treatment activities (NTA); 6. malnutrition prevention activities (NPA); 7. smallholder agricultural market support activities (SMS); 8. individual capacity strengthening activities (CSB); 9. institutional capacity strengthening activities (CSI); 10. service provision and platforms activities (CPA); 11. emergency preparedness activities (EPA); 12. analysis, assessment and monitoring activities (AAA); and 13. other (OTH). The acronyms of activity categories are reported in the COMET Manual.

Features of the evaluations				Activity categories ¹⁶⁹ as per Annex I of the revised Corporate Result Framework (2017–2021) ¹⁷⁰												
Evaluation by country	Evaluation type	Year of approval	PHQA	1. URT	2. ACL	3. CAR	4. SMP	5. NTA	6. NPA	7. SMS	8. CSB	9. CSI	10. CPA	11. EPA	12. AAA	13. OTH
Gambia	CSP	2021	83			✓	✓				✓	✓				
Honduras	CSP	2021	92				✓	✓	✓		✓	✓				
Laos	CSP	2021	90		✓	✓	✓	✓	✓		✓	✓		✓		
Lebanon	CSP	2021	81				✓	✓	✓		✓	✓		✓		✓
Zimbabwe	CSP	2021	88		✓	✓		✓	✓					✓		✓
Update of WFP's Safety Nets Policy	Policy	2019	74		✓		✓	✓	✓		✓	✓				
Evaluation of the Gender Policy	Policy	2020	96													✓
SST Cooperation Policy	Policy	2021	85				✓	✓	✓		✓	✓				✓
Pilot Country Strategic Plans	Strategic	2018	82													
WFP Support for Enhanced Resilience	Strategic	2018	63		✓	✓										✓
WFP's Capacity to Respond to Emergencies	Strategic	2019	76											✓		
Contribution of SF activities	Strategic	2020	94				✓									
RBA joint evaluation	Strategic	2021	91		✓	✓		✓	✓		✓	✓				✓

Features of the evaluations				Activity categories ¹⁶⁹ as per Annex I of the revised Corporate Result Framework (2017–2021) ¹⁷⁰												
Evaluation by country	Evaluation type	Year of approval	PHQA	1. URT	2. ACL	3. CAR	4. SMP	5. NTA	6. NPA	7. SMS	8. CSB	9. CSI	10. CPA	11. EPA	12. AAA	13. OTH
Technology	Strategic	2021	94											✓		✓
Algeria (DE)	Activity	2018	73					✓								
Bangladesh (DE)	Activity	2018	62				✓									
Lesotho (DE)	Activity	2018	74				✓					✓				
Malawi (DE)	Activity	2018	76				✓									
Philippines (DE)	Activity	2018	69			✓										
Ethiopia (DE)	Activity	2018	70													
Turkey (DE)	Activity	2018	74	✓										✓		
Eswatini (DE)	Activity	2019	66				✓					✓				
Malawi (DE)	Activity	2019	70			✓										
Rwanda (DE)	Activity	2019	72				✓					✓				
Guinea-Bissau (DE)	Activity	2020	71				✓					✓				
Lebanon (DE)	Activity	2020	85		✓											

Features of the evaluations				Activity categories ¹⁶⁹ as per Annex I of the revised Corporate Result Framework (2017–2021) ¹⁷⁰												
Evaluation by country	Evaluation type	Year of approval	PHQA	1. URT	2. ACL	3. CAR	4. SMP	5. NTA	6. NPA	7. SMS	8. CSB	9. CSI	10. CPA	11. EPA	12. AAA	13. OTH
Madagascar (DE)	Activity	2020	86				✓									
Malawi (DE)	Activity	2020	93				✓									
Benin (DE)	Activity	2020	85				✓					✓				
Mozambique (DE)	Activity	2020	73						✓	✓						
Namibia (DE)	Activity	2020	85									✓				
Niger (DE)	Activity	2020	89				✓							✓		
Syria (DE)	Activity	2020	69				✓							✓		
Gambia (DE)	Activity	2021	85					✓	✓			✓				
Ghana (DE)	Activity	2021	80						✓	✓						
Guinea-Bissau (DE)	Activity	2021	75				✓			✓		✓				
Honduras (DE)	Activity	2021	79				✓			✓		✓	✓			
Laos (DE)	Activity	2021	72				✓									
Libya (DE)	Activity	2021	71	✓												

Features of the evaluations				Activity categories ¹⁶⁹ as per Annex I of the revised Corporate Result Framework (2017–2021) ¹⁷⁰												
Evaluation by country	Evaluation type	Year of approval	PHQA	1. URT	2. ACL	3. CAR	4. SMP	5. NTA	6. NPA	7. SMS	8. CSB	9. CSI	10. CPA	11. EPA	12. AAA	13. OTH
Sri Lanka (DE)	Activity	2021	77			✓										
Economic Empowerment of Rural Women (HQ)	Activity	2021	91													
Burkina Faso (DE)	Thematic	2020	82													✓
Burundi (DE)	Thematic	2020	79					✓	✓							
El Salvador (DE)	Thematic	2020	90													✓
Kenya (DE)	Transfer Modality	2018	63	✓												
Senegal (DE)	Transfer Modality	2018	61				✓									

Annex V: Analytical framework

190. The purpose of the analytical framework was to provide a tool against which data was extracted and later analysed. It was shaped around the main questions of the synthesis and contains space to include data from evaluations and any additional information being gathered.

191. Table 7 below details changes from the original synthesis questions. These changes were made in consultation with OEV to clarify the intent of the question and avoid duplication of findings.

Table 17: Synthesis question revisions from TOR to inception report

Original question		Revised question	
1	Do the corporate outcome, output and cross-cutting indicators collected and analysed at country level allow for effective measurement of intervention achievements?	1	Do corporate outcome, output and cross-cutting indicators allow for effective measurement of intervention achievements at the country level?
2	To what extent are the monitoring practices in line with corporate guidance and tools?	-	<i>Question dropped in consultation with OEV as this issue will be addressed through other synthesis questions.</i>
3	To what extent have monitoring systems been able to generate credible and evidence-based information (quantitative and qualitative) to inform intervention design, planning and implementation? How has this information been used, by whom and for what purpose?	2	To what extent have WFP's monitoring systems generated credible information? How has information generated by WFP monitoring systems been used, and by whom?
4	To what extent has WFP's Monitoring Strategy and Normative Framework enabled WFP to track programme effectiveness and inform corporate performance reporting by providing achievement values for the programmatic indicators?	3	To what extent has WFP's Normative Framework enabled WFP to track programme effectiveness and inform corporate performance reporting?
		4	To what extent does the evaluative evidence provide learning on the three outcomes of the WFP Corporate Monitoring Strategy?
5	What factors contributed to or hindered effective implementation of performance measurement and monitoring systems? Are there particular activity areas and contexts in which the approach and systems have worked better and why?	5	What factors contributed to or hindered effective implementation of performance measurement and monitoring systems? Are there particular activity areas and contexts in which WFP's approach and systems have worked better and why?
6	To what extent is WFP performance measurement system aligned with national monitoring systems? How has WFP pursued opportunities to strengthen national monitoring systems through modelling and capacity strengthening interventions?	6	To what extent is WFP's performance measurement system aligned with national monitoring systems? How has WFP pursued opportunities to strengthen national monitoring systems?
	The synthesis will also note the extent to which cross-cutting priorities (accountability to affected populations, protection, gender and environment) in force during the period covered by the synthesis are reflected in monitoring guidance and systems	7	To what extent were cross-cutting priorities reflected in monitoring practices, guidance and systems?

192. Table 18 provides an overview of the analytical framework. It includes the main synthesis questions, sub-questions, codes for use in MAXQDA, and a description of the codes. The codes provide the basis for data extraction and analysis to address the synthesis questions. The codes helped the team gather, collate and analyse evidence consistently and systematically.

Table 18: Analytical framework

Synthesis questions	Codes	Code description
SQ1. Do corporate outcome, output and cross-cutting indicators allow for effective measurement of intervention achievements at the country level?		
1.1 Does the evaluation address whether corporate outcome, output and/or cross-cutting indicators were used at the country level? If so, what evidence did the evaluation generate about use of corporate indicators at the country level?	Corporate indicators – selection	<ul style="list-style-type: none"> Discussion on how the <u>corporate indicators</u> were selected at the country level
	Corporate indicators – use	<ul style="list-style-type: none"> Discussion on how the <u>corporate indicators</u> were used at the country level
	Corporate indicators – data collection methodologies	<ul style="list-style-type: none"> Discussion on data collection methodologies (including frequency of data collection) for <u>corporate indicators</u> at the country level
1.2 Does the evaluation address whether corporate outcome, output and/or cross-cutting indicators were effective at measuring intervention achievements at the country level? If so, what evidence did the evaluation generate about effectiveness of corporate indicators at the country level?	Corporate indicators – relevance	<ul style="list-style-type: none"> Discussion on the relevance of <u>corporate indicators</u> at the country level
	Corporate indicators – credibility	<ul style="list-style-type: none"> Discussion on the credibility of <u>corporate indicators</u> at the country level
	Corporate indicators – quality	<ul style="list-style-type: none"> Discussion on the quality of <u>corporate indicators</u> at the country level
SQ2. To what extent have WFP’s monitoring systems generated credible information¹⁷¹? How has information generated by WFP monitoring systems been used, and by whom?		
2.1 Does the evaluation address whether WFP’s monitoring systems generated credible information ? If so, what evidence did the evaluation generate about the credibility of monitoring information?	Monitoring – credibility	<ul style="list-style-type: none"> Discussion on the credibility of <u>monitoring information</u> collected against the monitoring systems Discussion on whether credible monitoring information that facilitated strategic and operational decision-making was generated
2.2 Does the evaluation address use of <u>monitoring information</u> ? If so, what evidence did the evaluation generate about whether and how monitoring information have been used, by whom ?	Monitoring – use	<ul style="list-style-type: none"> Discussion on whether the <u>monitoring information</u> was used
	Monitoring – how	<ul style="list-style-type: none"> Discussion on how the <u>monitoring information</u> was used
	Monitoring – who	<ul style="list-style-type: none"> Discussion on who used the <u>monitoring information</u>
SQ3. To what extent has WFP’s Normative Framework enabled WFP to track programme effectiveness and inform corporate performance reporting?		

¹⁷¹ Evaluation reports often use data, information, and evidence interchangeably. The synthesis report will outline taxonomies found in the evaluation reports and the taxonomy used in the synthesis report.

Synthesis questions	Codes	Code description
<p>3.1 Does the evaluation provide evidence on developing the logframe? If so, how does the evidence relate to the Normative Framework to either track programme effectiveness and/or inform corporate performance reporting?</p>	<p>Step 1 – logframe</p>	<ul style="list-style-type: none"> - Is there evidence of a logframe being developed? • Discussion on whether and to what extent the Normative Framework has guided tracking of programme effectiveness and/or corporate performance reporting in relation to developing the logframe • Links to the SOPs: <ul style="list-style-type: none"> ○ SOP 1: CSP logframes and monitoring narratives are prepared ○ SOP 2: Completed logframes are entered in COMET • Links to the CRF Business Rules: <ul style="list-style-type: none"> ○ Logframe design and indicator selection ○ Baseline establishment and target setting • Links to CRF – issues related to selecting corporate indicators
<p>3.2 Does the evaluation provide evidence on preparing the monitoring, review and evaluation plan? If so, how does the evidence relate to the Normative Framework to either track programme effectiveness and/or inform corporate performance reporting?</p>	<p>Step 2 – monitoring, review and evaluation plan</p>	<ul style="list-style-type: none"> - Is there evidence of a monitoring, review and evaluating plan being developed? • Discussion on whether and to what extent the Normative Framework has guided tracking of programme effectiveness and/or corporate performance reporting in relation to preparing the monitoring, review and evaluation plan • Links to the SOPs: <ul style="list-style-type: none"> ○ SOP 3: A monitoring, Review and Evaluation Plan is developed ○ SOP 4: A draft M&E budget is prepared reflecting costs of the MRE Plan • Links to MMRs – issues to planning related to: <ul style="list-style-type: none"> ○ monitoring coverage, baselines, targets, data collection exercises, monitoring frequency, applicability, level of disaggregation and sampling requirements
<p>3.3 Does the evaluation provide evidence on developing a monitoring toolkit? If so, how does the evidence relate to the Normative Framework to either track programme effectiveness and/or inform corporate performance reporting?</p>	<p>Step 3 – monitoring toolkit</p>	<ul style="list-style-type: none"> - Is there evidence of a monitoring toolkit being developed? • Discussion on whether and to what extent the Normative Framework has guided tracking of programme effectiveness and/or corporate performance reporting in relation to developing a monitoring toolkit (methodologies, tools, analysis plan, implementation plan) • Links to the SOPs: <ul style="list-style-type: none"> ○ SOP 5: A monitoring toolkit is developed covering the processes listed in the MRE plan ○ SOP 6: Databases to store and process monitoring data are developed ○ SOP 7: A list of sites to be monitored is prepared within a monthly activity plan
<p>3.4 Does the evaluation provide evidence on collecting primary data and collating secondary data? If so, how does the evidence relate to the Normative Framework to</p>	<p>Step 4 – data collection</p>	<ul style="list-style-type: none"> - Is there evidence that primary data were collected and/or secondary data were collated?

Synthesis questions	Codes	Code description
either track programme effectiveness and/or inform corporate performance reporting?		<ul style="list-style-type: none"> • Discussion on whether and to what extent the Normative Framework has guided tracking of programme effectiveness and/or corporate performance reporting in relation to collecting primary data and collating secondary data • Links to the SOPs: <ul style="list-style-type: none"> ○ SOP 8: Primary data is collected, and secondary data synthesized as per the MRE plan ○ SOP 9: Baselines are entered or updated in COMET • Links to the CRF Business Rules: <ul style="list-style-type: none"> ○ Baseline establishment and target setting ○ Monitoring and reporting • Links to MMRs – issues with data collection related to: <ul style="list-style-type: none"> ○ monitoring coverage, baselines, targets, data collection exercises, monitoring frequency, applicability, level of disaggregation and sampling requirements • Links to CRF – issues related to collecting data for corporate indicators
3.5 Does the evaluation provide evidence on capturing, compiling and validating data ? If so, how does the evidence relate to the Normative Framework to either track programme effectiveness and/or inform corporate performance reporting?	Step 5 – data capture	<ul style="list-style-type: none"> - Is there evidence that data were captured, compiled and validated? • Discussion on whether and to what extent the Normative Framework has guided tracking of programme effectiveness and/or corporate performance reporting in relation to capturing, compiling and validating data • Links to the SOPs: <ul style="list-style-type: none"> ○ SOP 10: Monitoring data is entered in databases, compiled and validated • Links to CRF – issues related to capturing, compiling or validating corporate indicators
3.6 Does the evaluation provide evidence on analysing data and preparing information products ? If so, how does the evidence relate to the Normative Framework to either track programme effectiveness and/or inform corporate performance reporting?	Step 6 – data analysis and reporting	<ul style="list-style-type: none"> - Is there evidence that data were analysed, and information products were developed? - Is there evidence of data being analysed and information products being developed? • Discussion on whether and to what extent the Normative Framework has guided tracking of programme effectiveness and/or corporate performance reporting in relation to analysing data and preparing information products <ul style="list-style-type: none"> ○ Links to the SOPs: <ul style="list-style-type: none"> ○ SOP 11: Data is analysed as per the Analysis Plan and information products prepared • Links to CRF – issues related to analysing or reporting on corporate indicators
3.7 Does the evaluation provide evidence on making use of monitoring findings ? If so, how does the evidence relate to the Normative Framework to either track	Step 7 – data use	<ul style="list-style-type: none"> - Is there evidence on making use of monitoring findings? • Discussion on whether and to what extent the Normative Framework has guided tracking of programme effectiveness and/or corporate performance

Synthesis questions	Codes	Code description
programme effectiveness and/or inform corporate performance reporting		reporting in relation to making use of monitoring findings, taking action and documenting lessons <ul style="list-style-type: none"> • Links to the SOPs: <ul style="list-style-type: none"> ○ SOP 12: An escalation system is established to ensure required actions are taken at the appropriate place in a timely manner ○ SOP 13: Regular meetings are conducted between M&E, programme and technical staff to review monitoring findings and implement action • Links to CRF – issues related to utilization of corporate indicator findings
3.8 Does the evaluation provide evidence on conducting evaluations or reviews ? If so, how does the evidence relate to the Normative Framework to either track programme effectiveness and/or inform corporate performance reporting?	Step 8 – evaluation	<ul style="list-style-type: none"> - Is there evidence of conducting evaluations and/or reviews? • Discussion on whether and to what extent the Normative Framework has guided tracking of programme effectiveness and/or corporate performance reporting in relation to conducting evaluations or reviews • Links to the SOPs: <ul style="list-style-type: none"> ○ SOP 14: Monitoring data is consolidated for review and evaluation purposes and made available to evaluators ○ SOP 15: Findings of reviews and evaluations are used to inform the CSP design and monitoring process
SQ4. To what extent does the evaluative evidence provide learning on the three outcomes of the WFP Corporate Monitoring Strategy?		
4.1 Does the evaluation provide evidence on outcome 1 (adequate monitoring expertise)? If so, how? To what extent?	Corporate Monitoring Strategy Outcome 1	<ul style="list-style-type: none"> • Discussion on whether and to what extent there is <u>adequate monitoring expertise</u> (Outcome 1): • Examples/potential sub-codes include: <ul style="list-style-type: none"> ○ Undertake analysis of existing capacity on monitoring against technical and operational needs ○ Provide adequately experienced monitoring staff ○ Retain adequately experienced monitoring staff ○ Support staff skills development on monitoring ○ Sensitize senior management on the importance of providing adequate monitoring expertise
4.2 Does the evaluation provide evidence on outcome 2 (financial commitment)? If so, how? To what extent?	Corporate Monitoring Strategy Outcome 2	<ul style="list-style-type: none"> • Discussion on whether and to what extent there are <u>adequate finances</u> committed to monitoring (Outcome 2) • Examples/potential sub-codes include: <ul style="list-style-type: none"> ○ Provide monitoring budgeting practices to establish resources and gaps ○ Use of a monitoring contingency fund
4.3 Does the evaluation provide evidence on outcome 3 (functional capacity)? If so, how? To what extent?	Corporate Monitoring Strategy Outcome 3	<ul style="list-style-type: none"> • Discussion on whether and to what extent monitoring systems are <u>functional</u> • Examples/potential sub-codes include: <ul style="list-style-type: none"> ○ Standardized monitoring systems and procedures ○ Automated data collection ○ Commitment/support for use of monitoring data

Synthesis questions	Codes	Code description
		<ul style="list-style-type: none"> Evidence-based knowledge management
SQ5. What factors contributed to or hindered implementation of performance measurement and monitoring systems? Are there particular activity areas and contexts in which WFP's approach and systems have worked better and why?		
5.1 Does the evaluation address any factors that contributed to or hindered implementation of performance measurement and monitoring systems? If so, what were they? To what extent?	Factors – positive	<ul style="list-style-type: none"> Discussion on whether any factors (other than the activity type or country context) influenced the implementation of performance measurement and monitoring systems in a <u>positive</u> way and to what extent
	Factors – negative	<ul style="list-style-type: none"> Discussion on whether any <u>factors</u> (other than the activity type or country context) influenced the implementation of performance measurement and monitoring systems in a <u>negative</u> way and to what extent
5.2 Does the evaluation identify any particular activity type in which WFP's performance measurement and monitoring systems worked well or not well? If so, what were they? To what extent?	Activity – positive	<ul style="list-style-type: none"> Discussion on whether performance measurement and monitoring systems worked well or not well for any particular <u>activity types</u> and to what extent Sub-codes to identify programme activity type: <ul style="list-style-type: none"> Unconditional resource transfers Asset creation and livelihood Climate adaptation and risk management School meals Nutrition treatment Malnutrition prevention Smallholder agricultural market support Individual capacity strengthening Institutional capacity strengthening Service provision and platforms Emergency preparedness Analysis, assessment and monitoring Other
5.3 Does the evaluation address ways in which the country context contributed to or hindered implementation of performance measurement and monitoring systems? If so, what were they? To what extent?	Context – positive	<ul style="list-style-type: none"> Discussion on whether any <u>contextual factors</u> influenced the implementation of performance measurement and monitoring systems in a <u>positive</u> or <u>negative</u> way and to what extent Sub-codes to identify context type: <ul style="list-style-type: none"> Sudden crisis/emergency (e.g. climate, conflict, etc.) Protracted crises (e.g. climate, conflict, etc.) Levels of economic development
	Context – negative	
SQ6. To what extent is WFP's performance measurement system aligned with national monitoring systems? How has WFP pursued opportunities to strengthen national monitoring systems?		

Synthesis questions	Codes	Code description
6.1 Does the evaluation address whether WFP's performance measurement systems are aligned with national monitoring systems ? If so, what evidence did the evaluation generate about monitoring system alignment?	National systems – alignment	Discussion on whether and to what extent performance measurement and monitoring systems <u>align</u> with <u>national monitoring systems</u>
6.2 Does the evaluation address whether WFP pursued opportunities to strengthen national monitoring systems ? If so, how were they pursued?	National systems – strengthening	Discussion on whether any activities undertaken to strengthen national monitoring systems (e.g. capacity strengthening interventions), how these were undertaken and to what extent they <u>improved national monitoring systems</u>
SQ7. To what extent were cross-cutting issues (GEWE, protection, accountability to affected population and environment)¹⁷² reflected in monitoring practices, guidance and systems?		
7.1 Does the evaluation address whether accountability to affected populations was reflected in monitoring practices, guidance and systems? If so, what evidence did the evaluation generate?	Accountability to affected populations	<ul style="list-style-type: none"> • Discussion on whether and to what extent <u>accountability to affected populations</u> are reflected in monitoring practices, guidance and systems • Definition: Affected populations are able to hold WFP and partners accountable for meeting their hunger needs in a manner that reflects their views and preferences
7.2 Does the evaluation address whether protection to affected populations was reflected in monitoring practices, guidance and systems? If so, what evidence did the evaluation generate?	Protection to affected populations	<ul style="list-style-type: none"> • Discussion on whether and to what extent <u>protection to affected populations</u> is reflected in monitoring practices, guidance and systems • Definition: Affected populations are able to benefit from WFP programmes in a manner that ensures • And promotes their safety, dignity and integrity
7.3 Does the evaluation address whether GEWE were reflected in monitoring practices, guidance and systems? If so, what evidence did the evaluation generate?	Gender equality and women's economic empowerment	<ul style="list-style-type: none"> • Discussion on whether and to what extent <u>gender equality and women's economic empowerment</u> is reflected in monitoring practices, guidance and systems • Definition: Improved GEWE among WFP-assisted populations
7.4 Does the evaluation address whether environmental sustainability was reflected in monitoring practices, guidance and systems? If so, what evidence did the evaluation generate?	Environmental sustainability	<ul style="list-style-type: none"> • Discussion on whether and to what extent <u>environmental sustainability</u> is reflected in monitoring systems, guidance and practice • Definition: Targeted communities benefit from WFP programmes in a manner that does not harm the environment

¹⁷² The four cross-cutting issues are defined in the Revised Corporate Results Framework (2017–2021).

Annex VI: Findings, conclusions, and recommendations matrix

193. Table 19 outlines the relationship between the findings, conclusions and recommendations in this report.

Table 19: Findings, conclusions, and recommendations matrix

Main findings	Conclusions	Recommendations
SQ1. Do corporate outcome, output and cross-cutting indicators allow for effective measurement of intervention achievements at the country level?		
<p>Finding 1: While corporate indicators allow WFP to aggregate data at the corporate level, they often fall short of enabling country offices to effectively measure and report on a complete set of country level results.</p> <p>Finding 2: Changes to some corporate indicators decreased the effective measurement of achievements at country level over time. However, these changes were made with the intent of improving WFP's ability to measure results more effectively.</p>	<p>Conclusion 2: The Normative Framework continues to provide the structure to support effective performance management and monitoring, but does not encourage WFP to better capture the breadth of its achievements or track them over time.</p>	<p>Recommendations 2, 2.1, 2.2,</p>
SQ2. To what extent have WFP's monitoring systems generated credible information? How has information generated by WFP monitoring systems been used, and by whom?		
<p>Finding 3: Evaluations tended to only address the credibility of monitoring data if they found a shortcoming or problem. Four sets of challenges to credibility identified in the evaluations relate to monitoring frameworks, data gaps, data quality and data disaggregation.</p> <p>Finding 4: Over 90 per cent of all evaluations in the sample recommended improvements be made to monitoring systems or practices. Most of these recommendations focused on improving monitoring frameworks, but some also addressed data gaps, data quality, and disaggregation.</p> <p>Finding 5: The evaluations provided examples of monitoring information used for both accountability and learning. Monitoring information used for reporting within WFP and to donors served an accountability objective, while – to a lesser extent – its use by management to adjust current activities, inform activity design and share lessons with stakeholders served a learning objective.</p> <p>Finding 6: The evaluations highlighted a need to expand qualitative data collection, analysis, and reporting to contextualize WFP's achievements and to support WFP's ability to learn and adapt using its monitoring information.</p>	<p>Conclusion 1: While WFP is collecting credible monitoring data, and the three outcomes of the Corporate Monitoring Strategy remain relevant overall, there are still opportunities for improvement around resourcing the monitoring function, using monitoring data, and supporting monitoring systems and processes. Focusing on factors that influence the approach and conduct of monitoring can also help to guide improvements to monitoring systems and data.</p> <p>Conclusion 3: WFP uses its monitoring data for reporting, but does not consistently leverage it for learning and adaptation of programmes.</p>	<p>Recommendations 1, 1.1, 1.2, 3, 3.1, 3.2, 3.3, 3.4,</p>

Main findings	Conclusions	Recommendations
SQ3. To what extent has WFP's Normative Framework enabled WFP to track programme effectiveness and inform corporate performance reporting?		
<p>Finding 7: While some evaluations refer to the CRF (see questions 1 and 2), the other components of WFP's Normative Framework were not directly discussed. However, some evidence was provided around the steps of the monitoring SOPs that linked use of the SOPs to tracking programme effectiveness and corporate performance reporting.</p>	<p>See conclusion 2 under SQ1.</p>	<p>Recommendations 1, 1.1, 1.2, 2, 2.1, 2.2,</p>
SQ4. To what extent does the evaluative evidence provide learning on the three outcomes of the WFP Corporate Monitoring Strategy?		
<p>Finding 8: Thirty per cent of evaluations included in the synthesis identified shortcoming in staffing levels and capacity, with only six percent assessing as sufficient the staffing levels allocated to the monitoring function. Evaluations also noted concerns about monitoring staff being assigned too many varied responsibilities.</p> <p>Finding 9: Evaluations included in the synthesis rarely addressed the financial requirements for monitoring. When references are made, they pointed to a lack of funding as a hindering factor.</p>	<p>See conclusion 1 under SQ2.</p>	<p>Recommendations 4, 4.1, 4.2</p>
SQ5. What factors contributed to or hindered implementation of performance measurement and monitoring systems? Are there particular activity areas and contexts in which WFP's approach and systems have worked better and why?		
<p>Finding 10: The evaluations note that where engagement with the host government is positive, this contributed to the implementation of performance measurement and monitoring, while challenges with the local government hindered monitoring. Recommendations to strengthen government engagement included starting or improving joint monitoring, creating minimum monitoring requirements to cover multiple donor requests, integrating data collection practices into government systems, and capacity strengthening.</p> <p>Finding 11: The evaluations reveal that a lack of funding and/or staff resources hinder monitoring, and recommend improvements on aspects of resourcing including through: increasing staffing levels and/or improving current staff capacity through training; improving budgeting processes; and making sufficient resources available for staffing and monitoring activities.</p> <p>Finding 12: Effective use of technology – such as mobile data collection, mapping, dashboards, and data visualization – contributed to the implementation of performance measurement and monitoring and was credited with enhancing the ease and timeliness of reporting, improved efficiency, cost reduction, increasing the scale and/or frequency of data collection, and standardizing data collection.</p>	<p>See conclusion 1 under SQ2.</p> <p>Conclusion 4: Measurement of country capacity strengthening and resilience are specific activity areas that require additional development to better capture WFP's performance and achievements.</p>	<p>Recommendations 2, 2.1, 2.2, 3, 3.1, 3.2, 3.3, 3.4, 4, 4.1, 4.2, 4.3, 4.4</p>

Main findings	Conclusions	Recommendations
<p>Challenges that hindered monitoring included the lack of a useful centralized database for monitoring data at country level, fragmentation of data systems, and underutilization of WFP's mapping assets.</p> <p>Evaluation recommendations included merging databases or sharing data between systems; increasing use of digital data collection; and creating dashboards and visualizations to encourage data use.</p> <p>Finding 13: The evaluations found that a lack of knowledge management hindered monitoring. Evaluation recommendations included establishing annual learning events, developing dashboards, disseminating lessons learned internally and externally, creating feedback loops between field monitors and the CSP development process, and improving existing knowledge management systems for use in constrained environments.</p> <p>Finding 14: While recognizing that meeting donor monitoring requirements is necessary, the evaluations found that these requirements placed burdens on country offices, and that the data generated were not used beyond reporting. They also noted a lack of harmonization between donors.</p> <p>Finding 15: Except for the areas of country capacity strengthening and resilience, domains where challenges in measurement and use of indicators are clearly mentioned, the evaluations showed no clear patterns linking activity categories and monitoring challenges and opportunities.</p> <p>Finding 16: Few evaluations provided evidence on country contexts in which WFP's approach and systems for monitoring have worked better and why.</p>		
<p>SQ6. To what extent is WFP's performance measurement system aligned with national monitoring systems? How has WFP pursued opportunities to strengthen national monitoring systems?</p>		
<p>Finding 17: The evaluations documented WFP's overarching alignment with government priorities and plans, but rarely went into detail about the alignment of WFP's performance measurement systems with national monitoring systems.</p> <p>Finding 18: The evaluations noted that where insufficient attention is given to strengthening national monitoring systems, it may undermine efforts for transitioning interventions and sustainability.</p>	<p>See conclusion 1 under SQ2.</p>	<p>Recommendation 4.1</p>

Main findings	Conclusions	Recommendations
SQ7. To what extent were cross-cutting issues (GEWE, protection, accountability to affected population and environment)¹⁷³ reflected in monitoring practices, guidance and systems?		
<p>Finding 19: There is limited evidence across the evaluations of accountability to affected populations in monitoring practices, guidance and systems, beyond a few mentions of data availability and use.</p> <p>Finding 20: There is limited evidence across the evaluations on how protection of affected populations is covered through monitoring practices, guidance and systems beyond the need for improved indicators and additional data collection.</p> <p>Finding 21: The evaluations found limited integration of gender indicators in monitoring frameworks. Also, reporting on GEWE tends to focus on quantitative data, namely reporting beneficiary numbers, with limited analysis and use of qualitative data. The lack of qualitative data collection and analysis, including at the intra- and inter-household levels, limits ability to measure and analyse GEWE achievements, especially at the outcome level, given the need to consider perceptions and other aspects of more sensitive and intangible nature such as issues relating to personal safety and security, discrimination, and intra-household dynamics. In addition, the evaluations found limited use and analysis of gender-related data. Evaluations often reported a need to improve monitoring frameworks and practices, especially the collection of disaggregated data and gender-specific outcomes.</p> <p>Finding 22: There is limited evidence across evaluations on environmental issues in monitoring practices, guidance and systems.</p>	<p>Conclusion 5: Monitoring efforts for gender equality and women’s empowerment largely focus on accountability and reporting, especially quantitative reporting of beneficiary numbers, and they focus less on learning.</p>	<p>Recommendations 2, 2.1, 2.2,</p>

¹⁷³ The four cross-cutting issues are defined in the Revised Corporate Results Framework (2017–2021).

Annex VII: List of component evaluations

194. The evaluations referenced for this synthesis are included below in Table 20. In the body of the synthesis report, evaluations are referred to as follows:

- Country strategic plan evaluations – [Country] CSPE [year] – for example Indonesia CSPE 2020
- Policy evaluations – [Descriptive words] PE [year] – for example Gender PE 2020
- Strategy evaluations – [Descriptive words] SE [year] – for example Resilience SE 2018
- Decentralized evaluations – [Country] DE [year] – for example Lebanon DE 2020

Table 20: List of component evaluations included (53 in total)

Full title of the report	Abbreviation for this report	Evaluation type	Evaluation category	Commissioner	Regional bureau	Year
Strategic Evaluation of the Pilot Country Strategic Plans	Pilot CSP SE 2018	Strategic	Centralized	OEV	OEV	2018
Strategic Evaluation of WFP's Support for Enhanced Resilience	Resilience SE 2018	Strategic	Centralized	OEV	OEV	2018
Strategic Evaluation of WFP's Capacity to Respond to Emergencies	Emergency Response SE 2019	Strategic	Centralized	OEV	OEV	2019
Strategic Evaluation of the Contribution of School Feeding Activities to the Achievement of the Sustainable Development Goals	School Feeding SE 2020	Strategic	Centralized	OEV	OEV	2020
Joint Evaluation on the Collaboration among the United Nations Rome-based Agencies	RBA Collaboration SE 2021	Strategic	Centralized	OEV	OEV	2021
Strategic Evaluation of the WFP's Use of Technology in Constrained Environments	Technology SE 2021	Strategic	Centralized	OEV	OEV	2021
Update of WFP's Safety Nets Policy	Safety Nets PE 2019	Policy	Centralized	OEV	OEV	2019
Evaluation of the Gender Policy	Gender PE 2020	Policy	Centralized	OEV	OEV	2020
Evaluation of the WFP South-South and Triangular Cooperation Policy	SSTC PE 2021	Policy	Centralized	OEV	OEV	2021
Evaluation of Bangladesh WFP Country Strategic Plan 2016–2019	Bangladesh CSPE 2020	CSP	Centralized	OEV	RBB	2020
Evaluation of Cameroon WFP Country Strategic Plan 2018–2020	Cameroon CSPE 2020	CSP	Centralized	OEV	RBD	2020
Evaluation of Democratic Republic of the Congo Interim Country Strategic Plan 2018–2020	DRC CSPE 2020	CSP	Centralized	OEV	RBJ	2020
Evaluation of Indonesia WFP Country Strategic Plan 2017–2020	Indonesia CSPE 2020	CSP	Centralized	OEV	RBB	2020
Evaluation of Timor-Leste WFP Country Strategic Plan 2018–2020	Timor-Leste CSPE 2020	CSP	Centralized	OEV	RBB	2020
Evaluation of China WFP Country Strategic Plan (2017–2021)	China CSPE 2021	CSP	Centralized	OEV	RBB	2021

Full title of the report	Abbreviation for this report	Evaluation type	Evaluation category	Commissioner	Regional bureau	Year
Evaluation of El Salvador WFP Country Strategic Plan 2017–2021	El Salvador CPSE 2021	CSP	Centralized	OEV	RBP	2021
Evaluation of Honduras WFP Country Strategic Plan 2018–2021	Honduras CSPE 2021	CSP	Centralized	OEV	RBP	2021
Evaluation of Laos WFP Country Strategic Plan Evaluations (2017–2021)	Laos CSPE 2021	CSP	Centralized	OEV	RBB	2021
Evaluation of Lebanon WFP Country Strategic Plan (2018–2021)	Lebanon CSPE 2021	CSP	Centralized	OEV	RBC	2021
Evaluation of the Gambia WFP Country Strategic Plan (2018–2021)	Gambia CSPE 2021	CSP	Centralized	OEV	RBD	2021
Evaluation of Zimbabwe WFP Country Strategic Plan 2017–2021	Zimbabwe CSPE 2021	CSP	Centralized	OEV	RBJ	2021
Global End-term Evaluation of the Joint Programme on Accelerating Progress towards the Economic Empowerment of Rural Women in Ethiopia, Guatemala, Kyrgyzstan, Liberia, Nepal, Niger and Rwanda from 2014 to 2020	Economic Empowerment of Rural Women DE 2021	Activity	Decentralized	Gender Unit	OEV	2021
Evaluation of the DG ECHO funded Emergency Social Safety Net (ESSN) in Turkey	Turkey DE 2018	Activity	Decentralized	Turkey	RBC	2018
Evaluation of the National School Feeding Programme in Lesotho, in consultation with the Lesotho Ministry of Education and Training	Lesotho DE 2018	Activity	Decentralized	Lesotho	RBJ	2018
Evaluation of the Nutrition Components of the Algeria PRRO 200301	Algeria DE 2018	Activity	Decentralized	Algeria	RBC	2018
Final Evaluation of Disaster Preparedness and Response/Climate Change Adaptation Activities under the Office of Foreign Disaster Assistance Fund in the Philippines May 2011 to September 2017_FR	Philippines DE 2018	Activity	Decentralized	Philippines	RBB	2018
Final Evaluation of McGovern-Dole-supported School Feeding Programme in Bangladesh (FFE-388–2014/048-00)	Bangladesh DE 2018	Activity	Decentralized	Bangladesh	RBB	2018
Final Evaluation of the School Meals Programme in Malawi with support from United States Department of Agriculture, Government of Brazil and United Kingdom 2014 to 2016	Malawi DE 2018	Activity	Decentralized	Malawi	RBJ	2018
Final Evaluation of WFP'S USDA McGovern-Dole International Food for Education and Child Nutrition Programme's Support in Afar and Somali Regions in Ethiopia 2013–2017	Somalia DE 2018	Activity	Decentralized	Somalia	RBN	2018
Evaluation of the Eswatini National School Feeding Programme 2010–2018	Eswatini DE 2019	Activity	Decentralized	Eswatini	RBJ	2019
Midterm Evaluation of Integrated Risk Management and Climate Services Programme in Malawi from 2017–2019	Malawi DE 2019	Activity	Decentralized	Malawi	RBJ	2019

Full title of the report	Abbreviation for this report	Evaluation type	Evaluation category	Commissioner	Regional bureau	Year
WFP's USDA McGovern-Dole International Food for Education and Child Nutrition Programme's Support in Rwanda 2016–2020 Report: Midterm Evaluation	Rwanda DE 2019	Activity	Decentralized	Rwanda	RBN	2019
Contribution des cantines scolaires aux résultats de l'éducation dans le sud de Madagascar (2015 à 2019) : Une analyse de la contribution De janvier 2015 à juin 2019	Madagascar DE 2020	Activity	Decentralized	Madagascar	RBJ	2020
Evaluation conjointe à mi-parcours du Programme National d'Alimentation Scolaire Intégré (PNASI) Aout 2017 – Mai 2019_Benin	Mali DE 2020	Activity	Decentralized	Benin	RBD	2020
Evaluation of Namibia National School Feeding Programme	Namibia DE 2020	Activity	Decentralized	Namibia	RBJ	2020
Evaluation of the Joint Programme for Girls Education with financial support from the Norwegian Government July 2014 – October 2017_Malawi	Malawi DE 2020	Activity	Decentralized	Malawi	RBJ	2020
Evaluation Series on Emergency School Feeding in Niger	Niger DE 2020	Activity	Decentralized	Niger	RBD	2020
Evaluation Series on Emergency School Feeding in Syria	Syria DE 2020	Activity	Decentralized	Syria	RBC	2020
FINAL EVALUATION OF THE PROGRAMME 'Accelerate Progress Towards Millennium Development Goal 1C (MDG1.C Programme)_Mozambique	Mozambique DE 2020	Activity	Decentralized	Mozambique	RBJ	2020
Midterm Evaluation of McGovern-Dole Funded School Feeding Project in Guinea-Bissau (January 2016 – June 2018)	Guinea-Bissau DE 2020	Activity	Decentralized	Guinea-Bissau	RBD	2020
WFP Livelihoods and Resilience Activities in Lebanon 2016 – 2019	Lebanon DE 2020	Activity	Decentralized	Lebanon	RBC	2020
Addressing Climate Change Impacts on Marginalized Agricultural Communities Living in the Mahaweli River Basin of Sri Lanka 2013 – 2020	Sri Lanka DE 2021	Activity	Decentralized	Sri Lanka	RBB	2021
End-line Evaluation of USDA Local Regional Procurement project in Nalae District, Luang Namtha Province in Lao PDR [FY 16-19]	Laos DE 2021	Activity	Decentralized	Laos	RBB	2021
Evaluación del modelo de descentralización del Programa Nacional de Alimentación Escolar (PNAE) 2016–2019_Honduras	Honduras DE 2021	Activity	Decentralized	Honduras	RBP	2021
Final Evaluation of Enhanced Nutrition and Value Chains (ENVAC) project 2016–2021_Ghana	Ghana DE 2021	Activity	Decentralized	Ghana	RBD	2021
Final Evaluation of McGovern-Dole International Food for Education and Child Nutrition Programme in Guinea-Bissau 2016–2019	Guinea-Bissau DE 2021	Activity	Decentralized	Guinea-Bissau	RBD	2021
General Food Assistance and School Feeding Programmes, Libya	Libya DE 2021	Activity	Decentralized	Libya	RBC	2021

Full title of the report	Abbreviation for this report	Evaluation type	Evaluation category	Commissioner	Regional bureau	Year
Midterm Evaluation of Nutrition Activities in The Gambia 2016–2019	Gambia DE 2021	Activity	Decentralized	The Gambia	RBD	2021
Evaluación de género del Plan Estratégico de País de El Salvador (2017–2021)	El Salvador DE 2020	Thematic	Decentralized	El Salvador	RBP	2020
Evaluation of the Intervention for the Treatment of Moderate Acute Malnutrition in Ngozi, Kirundo, Cankuzo and Rutana 2016–2019_Burundi	Burundi DE 2020	Thematic	Decentralized	Burundi	RBN	2020
Evaluation thématique sur les questions de genre dans les interventions du PAM au Burkina Faso (2016–2018)	Burkina Faso DE 2020	Thematic	Decentralized	Burkina Faso	RBD	2020
An evaluation of the effects and a cost benefit analysis of the General Food Distribution Cash Modality scale-up (Cash-Based Transfers for PRRO 200737) for refugees and host communities in Kenya	Kenya DE 2018	Transfer Modality	Decentralized	Kenya	RBN	2018
Evaluation décentralisée de la modalité transfert monétaire utilisée dans le programme de cantines scolaires appuyé par le PAM au Sénégal Rapport d'évaluation Juillet 2018. Version Finale	Senegal DE 2018	Transfer Modality	Decentralized	Senegal	RBD	2018

Source: OEV Management Information System

Annex VIII: List of persons interviewed

195. The synthesis team conducted a series of key informant interviews (see Table 21) with a range of stakeholders involved in monitoring at corporate, regional and country levels during May 2022. The main purpose of these interviews was to triangulate and validate the emerging findings, and ensure that the synthesis team was aware of recent and relevant organizational context that could be factored into the final report.

Table 21: List of persons interviewed

Jennifer Nyberg	Deputy Director CPP	Corporate Planning and Performance Division
Rebecca Lamade	Chief, CPPM	Corporate Planning and Performance Division
David Ryckembusch	Chief, CPPS	Corporate Planning and Performance Division
Genevieve Wills	Chief, CPPP	Corporate Planning and Performance Division
Ronald Tran Ba Huy	Deputy Director RAM	Research, Assessment and Monitoring Division
Simon Renk	Programme Officer	Research, Assessment and Monitoring Division
Andres Chamba	Policy Programme Officer	Social Protection
Vera Mayer	Programme Officer	Climate and Disaster Risk Reduction Programme Unit
Yumiko Kanemitsu	Gender Consultant	Gender Division
Mariemesarra Ahmed	Gender Programme Policy Officer	Gender Division
Stien Gijssel	Chief Knowledge Management and Digital Innovation	Nutrition Division
Siti Halati	M&E Focal Point	Nutrition Division
Luna Kim	Regional Monitoring Advisor	Regional Bureau for Asian and the Pacific
Flaminia Mussio	Regional Monitoring Officer	Regional Bureau for Asian and the Pacific
Marta Fontan	Programme Cycle Advisor	Regional Bureau for the Middle East and Northern Africa and Eastern Europe
Caterina Kireeva	Regional Monitoring Advisor	Regional Bureau for Southern Africa
Federico Doehnert	Regional Monitoring Advisor	Regional Bureau for Western Africa
Zarrina Kurbanova	Regional Monitoring Advisor	Regional Bureau for Eastern Africa
Rossella Bottone	Regional Monitoring Advisor	Regional Bureau for Latin America
Katri Kangas	Programme Adviser	Technical Assistance and Country Capacity Strengthening Service (PROT)
Daniel Dyssel	Programme Policy Officer	Technical Assistance and Country Capacity Strengthening Service (PROT)
Tomilyn Rupert	CCS M&E Team	Technical Assistance and Country Capacity Strengthening Service (PROT)

Annex IX: Bibliography/evaluation library

196. Below is a list of documents that were consulted for the synthesis, excluding the evaluation reports.

BIBLIOGRAPHIC ENTRIES	Year
EVALUATIONS INCLUDED IN THE SYNTHESIS SCOPE	
Evaluation reports, plus annexes	2018–2021
Evaluability assessments of centralized evaluations inception reports	2018–2021
Management responses to centralized and decentralized evaluations	2021
PERFORMANCE MEASUREMENT AND MONITORING – POLICY AND STRATEGIES	
WFP's Performance Management Policy	2014
WFP Corporate Monitoring Strategy 2014–2017	2014
WFP Corporate Monitoring Strategy 2017–2021	2017
PERFORMANCE MEASUREMENT AND MONITORING – GUIDANCE	
Corporate Monitoring Guidance	2017
Minimum Monitoring Requirements	2020
Standard Operating Procedures for CSP Monitoring	2017
CRF Business Rules	2017
Annual Performance Planning Guide for Country Offices	2019
Monitoring Toolkit	
CRF Management Performance structure under the CRF	2017
COMET Directive	2020
COMET Manual	2017
CRF Guidance Note to Select Indicators for CSP Logframes 1.0	2022
GENDER	
Gender policy	2009, 2015 & 2022
Gender Policy 2009 Evaluation, Annexes and Management Response	
WFP EVALUATION SYNTHESSES AND MANAGEMENT RESPONSES	
ANNUAL AND REGIONAL OPERATION EVALUATIONS SERIES	
Annual Synthesis of OpeEval 2013–2014	2014

BIBLIOGRAPHIC ENTRIES	Year
Annual Synthesis of OpeEval 2014–2015	2015
Annual Synthesis of OpeEval 2015–2016	2016
Annual Synthesis of OpeEval 2016–2017	2017
RBB OpeEval Series Regional Synthesis	2017
RBC OpeEval Series Regional Synthesis	2017
RBD OpeEval Series Regional Synthesis	2017
RBJ OpeEval Series Regional Synthesis	2017
RBN OpeEval Series Regional Synthesis	2017
RBP OpeEval Series Regional Synthesis	2017
Synthesis report on four evaluations of the impact of WFP programmes and Management response	2017
Synthesis report of WFP's country portfolio evaluations in Africa (2016–2018) and Management response	2019
Synthesis report of evidence and lessons from WFP's Policy Evaluations (2011–2019)	2020
Synthesis Report on Evidence and Lessons on Country Capacity Strengthening from Decentralized Evaluations (2016–2019)	2021
Management Responses for all Centralized Evaluations	2018–2021
Review of the implementation of recommendations from thematic evaluations of a strategic/global nature. Synthesis Report	2022
ANNUAL EVALUATION REPORTS	
Annual evaluation reports	2017–2020
AUDIT REPORTS	
Internal Audit of Monitoring in WFP	2018
Consolidated Insights on the State of Country Office Monitoring within WFP	2021
Report of the External Auditor on the management of information on beneficiaries	2021
WFP STRATEGIC PLANS AND RELATED DOCS	
WFP Strategic Plan (2014–2017)	2013
WFP Strategic Results Framework 2014–2017	2013
WFP Management Results Framework 2014–2017	2013
Indicator Compendium 2014–2017	2013
WFP Strategic Plan (2017–2021)	2017
Financial Framework Review 2017–2021	2016
Corporate Results Framework 2017–2021	2016
Revised Corporate Results Framework 2017–2021	2018

BIBLIOGRAPHIC ENTRIES	Year
Midterm review of the Corporate Results Framework	2020
Midterm Review of WFP's Strategic Plan 2017-2021	2020
2017 - 2021 Corporate Results Framework Outcome and Output Indicator Compendium	2018
2017 - 2021 Programme Indicator Compendium - Revised Corporate Results Framework	2020
WFP Strategic Plan (2022-2025)	2022
Corporate Results Framework 2022-2026	2022
Country Capacity Strengthening Policy update	2022
ANNUAL PERFORMANCE REPORTS	
Annual Performance Reports	2017-2020
MOPAN	
MOPAN WFP Report	2019
MOPAN WFP Brief	2019
CORPORATE PLANNING MATERIALS	
ANNUAL PERFORMANCE PLAN	
WFP Annual Performance Plan	2018
Annual performance report for 2017	2018
Annual performance report for 2018	2019
Annual performance report for 2019	2020
Annual performance report for 2020	2021
Annex III-C: Supplemental Reporting on Top Ten Efficiency Gains in 2020	2021
ANNUAL PLANNING MATERIAL	
Country Strategic Planning Manual - Integrated Road Map	n.d.
Annual Performance Plan for Country Office	n.d.
Annual Performance Planning Process Flow	n.d.
APP and Risk Register mid-year review	n.d.
Country Operations Management Plan	n.d.
Full Manual _ Annual Country Report Guidance	n.d.
Headquarter divisions planning tools for 2022	n.d.
Internal controls and planning	n.d.

BIBLIOGRAPHIC ENTRIES	Year
Introduction	n.d.
Organizing your Country Office's APP	n.d.
Regional bureau planning tools for 2022	n.d.
PLANNING A Revised Approach	n.d.
The CRF Management Key Performance Indicator Compendium	n.d.
Workplans	n.d.
ANNUAL PERFORMANCE PLAN GUIDE	
Annual Performance Plan Guide – Compulsory indicators lists January 2018	2018
PERFORMANCE MANAGEMENT	
CRF Management Performance – detailed structure October 2017	2017
Behind the performance management cycle	2018
Performance Management Data Map	2018
Retail Performance Management Guidance Supply Chain Cash-based Transfers and Markets 2019	2019
RESEARCH ASSESSMENT AND MONITORING	
COMET	
Country Office Tool for Managing (programme operations) Effectively (COMET): Purpose, Scope, Roles and Responsibilities	2020
COMET and Integrated Road Map	n.d.
COMET – Design Module: Logframe Design & Results	n.d.
COMET Map and integration with other systems	n.d.
CORPORATE MONITORING GUIDANCE	
Country Office Monitoring Strategy	n.d.
Data Collection and Analysis	n.d.
Data Quality Guidance	n.d.
EMEP: Emergency M&E Package	n.d.
Guidance Note on Estimating and Counting Beneficiaries	n.d.
Logframe Guidance	n.d.
Country Office Guidance for Monitoring and Review of WFP Country Strategic Plans Developed by	n.d.
Outcome Monitoring	n.d.
Output Monitoring	n.d.

BIBLIOGRAPHIC ENTRIES	Year
Process Monitoring Issues Ticketing	n.d.
Process Monitoring	n.d.
Qualitative Research Guidance for WFP Monitoring Performance Management and Monitoring Division	n.d.
Monitoring Recommendations for Covid-19 Response a Practical Guide for M&E Officers in Regional Bureaux and Country Offices	n.d.
Remote monitoring	n.d.
Risk-based Monitoring Framework	n.d.
Guidance on Roles, Responsibilities and Oversight	n.d.
Guidance for Sharing WFP Monitoring Information with Donors or Third-Party Monitors	n.d.
Theories of Change	n.d.
Third-Party Monitoring Guidance	n.d.
MONITORING TOOLKIT	
Monitoring Toolkit	n.d.
CORPORATE MONITORING STRATEGY	
Corporate Monitoring Strategy 2015 – 2017	2014
WFP Corporate Monitoring Strategy 2018 – 2021	2018
OTHER	
WFP Nutrition Monitoring and Evaluation	2017
School Meals Monitoring Framework and Guidance	2017
Standard Operating Procedures for CSP Monitoring	2017
Collection on Quarterly Monitoring Reports – Guidance for Building a Quarterly Monitoring Report in Publisher	2018
Quarterly Monitoring Report Template	2018
Chapter 7 – Monitoring, Evaluation, and Review Food Assistance for Assets Manual	2020
Minimum Monitoring Requirements	2020
REGIONAL MONITORING DOCUMENTS	
RBB Evidence Position Paper	
RBB Programme Monitoring Results in 2020 Focusing on Relief, Nutrition, and Country Capacity Strengthening activities from the Annual Country Reports (2020), with additional RAM analysis – Internal Only	2020
RBB Key monitoring results and takeaways from School Feeding operations during COVID-19 Take Home Rations – Internal Only	2020
RBB Regional Monitoring Strategy for the Asia and the Pacific Region	2020
RBD Interim Guidance on Tier 2 and Tier 3 Beneficiaries In the Context of the COVID-19 Pandemic	2020

BIBLIOGRAPHIC ENTRIES	Year
RBD Monitoring Recommendations for Covid-19 Response a Practical Guide for M&E Officers in Regional Bureaux and Country Offices	2020
RBD Decision Tree for mobile surveys for Outcome Monitoring	
RBD Resilience Monitoring Framework Road map	2021
RBD Resilience Monitoring Framework Road map – Annexes	2021
RBD Regional Monitoring Road map 2018–2019	
RBD Qualitative Data Analysis and Reporting – Practical Guidelines	2020
RBJ WFP Southern Africa Regional Monitoring Strategy (2019–2021, Interim)	2021
RBJ Monitoring Outcome Data Status Report	2021
RBJ Remote Monitoring Survey Results	2021
RBJ Regional Analysis of Monitoring Findings (January–May 2021)	2021
RBN Oversight Principles	2021
RBN Regional RAM Strategy	2020
RBN Standard Operating Procedures Issue Tracking and Prioritization	
RBN Terms of Reference: Oversight Mission	
RBP Oversight 2021 – Final Results	2021
RBP M&E – Consulta con COs	2020
RBP Oversight Activities Conducted in 2019 and 2021	2019–2021
RBP Joint Oversight Missions	

Annex X: Acronyms and abbreviations

AAA	Analysis, assessment and monitoring activities
ACL	Asset creation and livelihood support activities
CAR	Climate adaptation and risk management activities
CCS	Country Capacity Strengthening
CE	Centralized evaluation
CEQAS	Centralized Evaluation Quality Assurance System
CO	Country Office
COMET	Country Office Tool for Managing Effectively
CPA	Service provision and platforms activities
CPP	Corporate planning and performance
CRF	Corporate Results Framework
CSB	Individual capacity strengthening activities
CSI	Institutional capacity strengthening activities
CSP	Country Strategic Plan
CSPE	Country Strategic Plan evaluation
DDoE	Deputy Director of Evaluation
DE	Decentralized evaluation
DED	Deputy Executive Director
DRC	Democratic Republic of Congo
EB	Executive Board
EM	Evaluation manager
FY	Financial year
GEWE	Gender equality and women's empowerment
HoP	Head of Programmes
IDP	Internally displaced person
IR	Inception report
IRG	Internal Reference Group
M&E	Monitoring and evaluation
MIS	Management information system
MMR	Minimum monitoring requirements
MRE	Monitoring, review and evaluation plan
NPA	Malnutrition prevention activities
NTA	Nutrition treatment activities
OEV	Office of Evaluation

OPC	Oversight and Policy Committee
OTH	Other
PAM	Programme Alimentaire Mondial
PE	Policy evaluation
PHQA	Post hoc Quality Assessment
PMM	Performance measurement and monitoring
QA	Quality Assurance
RAM	Research, Assessment and Monitoring Division
RB	Regional Bureaux
RBA	Rome-based agencies
RBB	Regional Bureau Bangkok
RBC	Regional Bureau Cairo
RBD	Regional Bureau Dakar
RBJ	Regional Bureau Johannesburg
RBN	Regional Bureau Nairobi
RBP	Regional Bureau Panama
RMA	Regional monitoring advisor
SE	Strategy evaluation
SER	Summary Evaluation Report
SMP	School meal activities
SMS	Smallholder agricultural market support activities
SOP	Standard operating procedures
SQ	Synthesis question
SSTC	South–South triangular cooperation
TL	Team leader
TOR	Terms of Reference
UNEG	United Nations Evaluation Group
URT	Unconditional resource transfers to support access to food
USDA	United States Department of Agriculture
VAM	Vulnerability Analysis and Mapping
WFP	World Food Programme

Office of Evaluation

World Food Programme

Via Cesare Giulio Viola 68/70

00148 Rome, Italy

T +39 06 65131 wfp.org/independent-evaluation