

## Technical Note

### Pilot Evaluation<sup>1</sup>

Version April 2017

#### 1. Introduction

1. The purpose of this TN is to clarify specificities of a pilot evaluation noting that the DEQAS Process Guide is the main reference to describe key characteristics common to all evaluations. This note only includes information specific to PILOT evaluations.

Key characteristics of a pilot evaluation	
<b>Focus</b>	Pilot to be understood as means of determining the best way to deliver assistance to improve humanitarian and development outcomes. This often includes introducing new instruments, institutional platforms, partnerships, processes and programme designs to assess what benefits will arise, for whom, and in what form.
<b>Scope</b>	All activities of the selected pilot, covering the period of implementation in a defined geographical area.
<b>Utility</b>	To demonstrate results and make decisions about scaling up, replicating or closing the activities.

#### 2. What is a pilot evaluation?

2. Whilst there is no agreed definition of a 'pilot' in WFP, the organization uses pilots and innovations and prototypes as a means of determining the best way to deliver assistance to improve humanitarian and development outcomes. This often includes introducing new instruments, institutional platforms, partnerships, processes and programme designs to assess what benefits will arise, for whom, and in what form.
3. Pilots aim to test the design, implementation, and contextual implications of a new or innovative programme in a limited area, or for a limited time before extending the pilot and/or scaling up for a larger population with a more extensive resource investment. Pilots may include prototypes which are new technical or operational models.

<sup>1</sup> For this Technical Note, a pilot is an initial small-scale implementation that is used to prove the viability of an intervention and/or approach. This could involve either the exploration of a new approach or the application of an approach that has already been used by other country offices but it is new to the country office, or the areas it is being implemented.

**Table 1: Examples of pilots in WFP**

Examples of pilots
<ul style="list-style-type: none"><li>• Testing different transfer modalities e.g. fresh food, electronic vouchers, paper vouchers, cash account, card cash;</li><li>• Testing different transfer-based activities e.g. cash for Assets etc;</li><li>• Trying the use of cash for social insurance;</li><li>• Using new products and crops (e.g. fortified food, locally sourced food);</li><li>• Technical assistance (e.g. enhancement of agribusiness models);</li><li>• Using new incentives (e.g. take home ration for girls/boys.</li></ul>

4. Pilot evaluations are especially important for WFP as it moves along the trajectory from Food Aid to Food Assistance, expanding the range of programming modalities and instruments at its disposal. The importance of a robust evidence base for these new modalities and instruments, and learning about their effects on target populations and scalability, is therefore key.

### 3. How is a Decentralized Pilot Evaluation used?

5. Findings from pilot evaluations are often used **to provide credible evidence** about the pilot in achieving specific outcomes of interest, in order to provide internal and external stakeholder groups information for decision making about replication, upscaling or closing of the activities piloted. As WFP increasingly aligns with, and provides advice to, national social protection systems, robust evidence from pilots help to understand what works, where and under what circumstances, to better inform national choices and investments.

### 4. When should a Pilot evaluation be conducted?

6. Pilot evaluations are usually conducted **towards the end** of the pilot process, to ensure that findings can feed into the design of future policy and/or any scaled-up of WFP intervention. It is also possible to conduct an evaluation following the pilot's completion, but this may lead to a loss of momentum for the design process of the scaled-up intervention, unless the key decisions are to be made in future.
7. Accordingly, since decisions on whether to proceed with designing and seeking funding for a fully scaled-up intervention, based on the pilot, are often taken whilst the pilot is still operational, it is very important that evaluations are **built into the planning and programming cycle** of the pilot intervention from the outset. This means that the data requirement to enable the evaluation are established at the onset, costed and rigorously implemented. Other resources (human and financial) can be planned for and protected, and ensure that the continuity and momentum of the intervention is not lost. It will also allow learning from the evaluation to be maximized.

### 5. What criteria and questions should be applied for Pilot evaluations?

8. The DEQAS [Technical Note on Evaluation Criteria and Questions](#), provides the definitions of these evaluation criteria, and explains how they should be applied in evaluations. Since a pilot evaluation is often **strongly focused on learning**, evaluation questions may include many 'explanatory', or 'how and why' questions, as well as the 'what worked where' type of questions.
9. For the evaluation of a pilot, evaluation questions generally lead up to **Lessons/Good Practices**, which provide generalizable information to inform other similar interventions in similar settings, and towards **Recommendations**, which confirm whether or not, and in what degree/where/for whom/how the pilot should be scaled up (see Table 2 for examples of lessons and recommendations these evaluation should lead to).

**Table 2: Example of lessons and recommendations to pay attention to in pilot evaluation on the basis of evidence generate while addressing the evaluation questions**

Lessons and Recommendations	
<b>What are the key Lessons from the pilot?</b>	<ul style="list-style-type: none"> <li>• How was the logic model/theory of change designed for the pilot realized in practice? Were there any differences, and how?</li> <li>• Which of the underlying assumptions that the pilot was based on held true? Which ones did not?</li> <li>• What has been learned - <i>concerning efficiency, effectiveness, relevance, sustainability and impact</i> - that might apply to other, similar, interventions, in different contexts?</li> </ul>
<b>What would be the Recommendations based on the results of the pilot and the lessons?</b>	<ul style="list-style-type: none"> <li>• Should the pilot be scaled up or replicated? Recommend whether the pilot or some of its sub-elements should be: <ul style="list-style-type: none"> <li>✓ Terminated – any pilot</li> <li>✓ Extended - any pilot</li> <li>✓ Scaled up – any pilot</li> <li>✓ Replicated – any pilot</li> <li>✓ Handed over – any pilot whose objective was to test an initiative/technology before handing over to government or communities.</li> </ul> </li> <li>• If the pilot should be extended/scaled up/replicated or handed over, then indicate: <ul style="list-style-type: none"> <li>✓ Which aspects</li> <li>✓ Where</li> <li>✓ For whom (target group)</li> <li>✓ How (applying which transfer modalities/intervention modalities)</li> <li>✓ With whom (partnerships)</li> <li>✓ At what scale</li> <li>✓ With what budget</li> <li>✓ Suggestions for design</li> </ul> </li> </ul>

## 6. Key issues to consider when undertaking a pilot evaluation

10. Key issues to consider when undertaking a pilot evaluation include evidence base for its replicability or scale up plan, timing of the evaluation, data requirements, partnerships/stakeholder interests, provisions for impartiality and independence, risks and dissemination and use plans. These are briefly explained in Table 3.

**Table 3: Key considerations for pilot evaluations**

<b>Replicability/ Scaling up Plan</b>	WFP's Coverage Norms for decentralized evaluation under its 2016-2021 Evaluation Policy point to the importance of evaluating pilots before scaling up or replicating. It is important therefore that the evaluation of any such pilot is planned to take place before the design process of any successor intervention, in order that the new design includes scalable elements. Equally important is that the evaluation assesses the evidence underpinning proposed scale-up, and provides conclusive judgements.
<b>Timing</b>	<p>Pilot evaluations can assess the intermediary or final results of a pilot before it is replicated or scaled up. The timing of the evaluation should therefore be considered given the intervention cycle and the coverage of recent/planned evaluations. Pilots selected to be evaluated should:</p> <ul style="list-style-type: none"><li>• be conducted at a time when they can inform future decisions on extension, scaling up, replication or hand over, depending on the type and objective of the pilot</li><li>• have produced intermediary or final results that can be assessed</li></ul>
<b>Data requirements</b>	Pilot evaluations must consider the new or innovative aspects of an intervention itself as well as the contextual factors that may influence results. For that reason, baseline data and contextual information is critical to determining what and how a pilot has achieved results.

## 7. Data and approaches for Pilot evaluation

11. Considering the decisions that have to be made at the outset of the pilot it is important the specific monitoring data are collected to inform the evaluation and potential lessons and recommendations for scale up, replication or closure. Application of a rigorous logic framework and theory of change is strongly advised.
12. To ensure that evaluations meet quality standards, WFP recommends that **a combination of mixed methods be applied**. This includes qualitative and quantitative methods for data collection and analysis. See the [Technical Note on Methodology](#) for more guidance on how particular methods can be combined.
13. The **balance of qualitative and quantitative** methods will vary depending on the type of pilot. A transfer modality or social insurance pilot project for example would likely combine quantitative data on delivery and results with qualitative information on how the transfer was used by beneficiaries. A logistics prototype may well emphasize more quantitative aspects.

For more information on Decentralised Evaluations visit our webpage

<http://newqo.wfp.org/how-do-i/do-an-evaluation>

Or contact the DE team at: [wfp.decentralizedevaluation@wfp.org](mailto:wfp.decentralizedevaluation@wfp.org)