

WFP Efficiency Analysis Guidance Materials – Definitions, Concepts and Methodology

Robrecht Renard¹

Stage 1 & 2 Final Report – 30 March 2013

Contents

1.	Introduction	3
2.	Definitions of effectiveness and efficiency	4
3.	The problem of attribution	10
4.	A taxonomy of WFP interventions	11
5.	Allocative efficiency analysis (Cost-Benefit Analysis)	15
6.	Technical efficiency analysis (Cost-Effectiveness Analysis)	17
7.	Other considerations in evaluating WFP interventions	20
8.	The BMZ efficiency toolbox: what is less useful for WFP	21
9.	The BMZ efficiency toolbox: what is useful for WFP	25
10.	Summary and recommendations	42
	Bibliography	44
	Acronyms	47
	Glossary	48
	Annex A: A selective review of rigorous WFP efficiency studies	49
	Annex B: An analysis of 13 selected OE evaluations	52
	Annex C: Fortified Blended Food Factory Project Timor Vita – Comments from an efficiency perspective	63

¹ Karel Verbeke at the Institute of Development Policy and Management of the University of Antwerp (IOB-UA) contributed the analysis for annex B. My thanks to Stephen Lister of Mokoro for many insightful suggestions, and the staff of OE, in particular Helen Wedgwood and Marian Read for their reactions to previous drafts.

Tables:

Table 1	Effectiveness and Efficiency in the Logical Framework for WFP	12
Table 2	An overview of BMZ efficiency tools.....	22
Table 3	Characteristics of selected BMZ tools	26
Table 4	Economic CBA of a food fortification factory (ex ante) (thousands of 2009 \$)	28
Table 5	Profitability (ex ante).....	29
Table 6	Economic CBA of a food fortification factory (ex post)	30
Table 7	Profitability (ex post)	30
Table 8	Alternative outcome rankings of three food modalities.....	34
Table 9	Cost-effectiveness of selected nutritional interventions	35
Table 10	Decision tree for selecting modalities	37
Table 11	Benchmarking of school-meal-only costs (standardised annual costs (\$) per recipient).....	39

Figures:

Figure 1	Efficiency and effectiveness in the logical framework	5
Figure 2	Value for money and the logframe at DFID	7
Figure 3	The “three Es” as interpreted by DFID	7
Figure 4	Efficiency notions throughout the logical framework at WFP.....	8

1. Introduction

1. The purpose of this consultancy is to produce user-friendly guidelines and templates for the country programme evaluations (CPEs) and Operations Evaluations (OpEvs) commissioned by the Office of Evaluation (OE) of WFP. Evaluation Managers and evaluation teams are expected to draw on these guidelines and templates in the future to select appropriate methods for efficiency analysis and presentation of findings in CPEs and OpEvs. The main focus is on efficiency as defined by OECD-DAC and on the tools described in a document by the Germany's Federal Ministry for Economic Cooperation (Palenberg 2011), some of which are expected to be relevant to evaluation in WFP. The closely related concept of value for money serves as a background, but is not the focus of this consultancy.

2. The exercise is carried out in tandem with the CPE for Timor-Leste undertaken by Mokoro, taking place over the same time period. Mokoro suggested the consultant and provides quality assurance. Timor-Leste was selected for testing the efficiency analysis methods.

3. The consultant visited WFP HQ in Rome on 19 and 20 November 2012 for a briefing with OE and other departments. He studied the reference document from Germany's Federal Ministry for Economic Cooperation proposed in the ToR, and undertook a broader literature review on the basis of documents from WFP, the DAC, other donors, and scientific journals and books. He participated in the Timor-Leste field visit from 27 November until 4 December 2012. The present document constitutes the draft stage 1 output² and stage 2 output³. The consultancy report will further consist of a stage 3 output⁴ and a stage 4 output⁵.

² Described in the ToR as follows: *short report outlining for efficiency analysis*: A. Definitions B. Concepts and C. Methods that reflect where the international evaluation community/organizations have reached on efficiency analysis methods; and WFP's and OE's status on the same.

³ Described in the ToR as follows: *progress report, testing strategy and guideline templates*. Approach to the development of efficiency analysis methods guidelines for OE evaluation managers and evaluation teams. This will include the methods to be tested, including a procedure for testing and developing the different methods throughout the CPE TL evaluation process. The set of templates for the guidance modules will be developed at this stage and included in the report, including a sample of at least one of the methods.

⁴ Described in the ToR as follows: a set of stand-alone guidance modules on efficiency analysis methods – already tested and assessed as appropriate to WFP OE CPEs and OpEvs. For each method: main concepts/definitions, tools, methods of analysis, ways of interpreting and presenting results and standards, where appropriate, etc. are to be presented in a user-friendly, electronic-publishing friendly format

⁵ Described in the ToR as follows: a short report on the results of the consultancy (achievements against plans) including a set of recommendations for where the Evaluation Quality Assurance System (EQAS) (CPE and other EQAS as appropriate) needs to be updated in light of the new guidance available on efficiency analysis methods.

2. Definitions of effectiveness and efficiency

4. Like many donor agencies, WFP prides itself on aligning its definitions to those of the OECD DAC in Paris. The OECD, and in particular the DAC, is indeed an important standard setter in the field of development evaluation. Historically, the organisation has sponsored important research into the application of efficiency analysis to developing countries⁶. When the aid industry began to devote more resources to evaluation, the DAC provided a forum for agreeing on standards and on terminology. Although inevitably some variation in the definitions has crept in to suit the needs, capacities and sensibilities of individual agencies, by and large aid agencies speak the same evaluation language. A major problem is that this common donor language is not necessarily consistent with the generic concepts underpinning the scientific literature, as we will show below. We start from the broadest DAC definition of efficiency, which specifies that it is a measure that relates inputs to results. In the logical framework terminology also standardised by DAC, “results” stand for either output, outcome, or impact. We suggest one small modification to include the level just below output, i.e. activities⁷. Indeed there is no reason not to include it, and the practice of aid agencies, including WFP, is to make this inclusion implicitly. In the resulting broad definition, efficiency is then applied to four hierarchical levels. There is however another, more narrow definition that also appears in some DAC documents, and is used by donors, in which efficiency is limited to the output level only (Michiels 2012: 6). We follow the BMZ study (Palenberg 2011: 7-8) and adopt the broad definition here, for reasons that we will explain in more detail below. Incidentally, support for this broad rather than the narrow definition was widely shared by the representatives in the OECD DAC network on Development Evaluation interviewed for the BMZ study. It is important to note that only the broader definition of efficiency is fully consistent with the large scientific literature on this topics from economics⁸.

5. We also adopt the general DAC definition of effectiveness. It refers to the extent to which a development intervention’s objectives are expected to be achieved (appraisal and interim evaluation), or have been achieved (end of project and ex post evaluation). In this definition “objectives” can again be best understood as pertaining to multiple levels of the logical framework: activities, output, outcome, and impact. This makes sense: an intervention may be effective in providing its output (the right food at the right time for the right beneficiaries) notwithstanding some “pipeline breaks” at activities level, yet may not achieve its outcome (improved nutrition). Another intervention may contribute both to output and outcome, but without being successful in its intended impact (reduced child mortality). It is obviously important to distinguish these different levels of effectiveness.

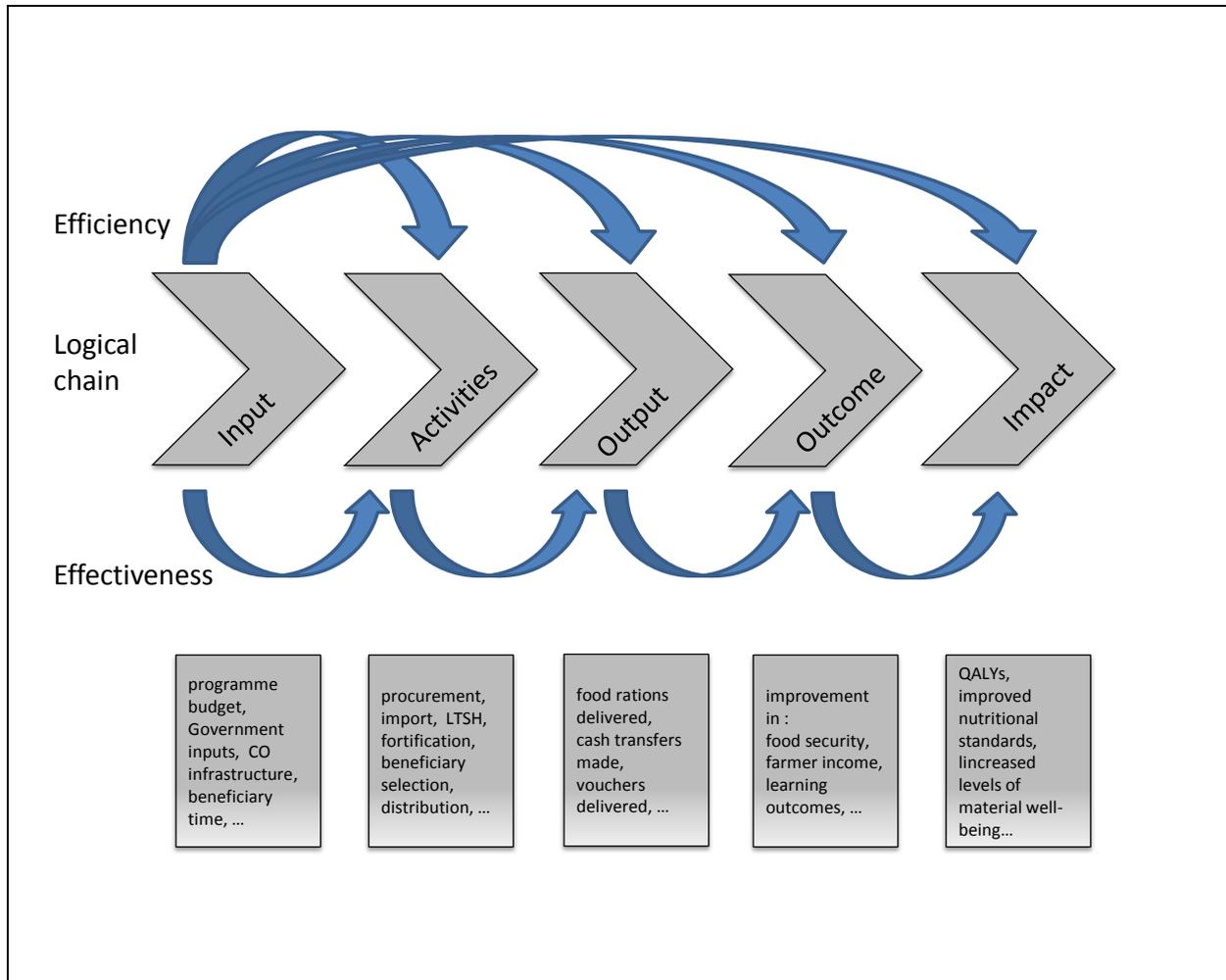
⁶ The prequels of two classic textbooks in the field - Rossi et al. (2004) and Little and Mirrlees (1974) – were commissioned and published by the OECD at the end of the 1960s.

⁷ Also sometimes called throughput. Whereas activities refers to the process, throughput refers to the intermediate results that are generated at that level. As efficiency is a relation between inputs and results, throughput is a useful addition to the terminology. Examples are warehouses, or locally produced fortified food. However, we will restrain from using “throughput” in the rest of the text, as this term is not in use at WFP and not essential to the argument.

⁸ As can be verified by consulting popular academic textbooks on the topic such as Boardman et al. (2011), Drummond et al. (1997), and Levin and McEwan (2001).

6. Figure 1 summarizes these general concepts of efficiency and effectiveness and the links between them.

Figure 1 Efficiency and effectiveness in the logical framework



Source: author

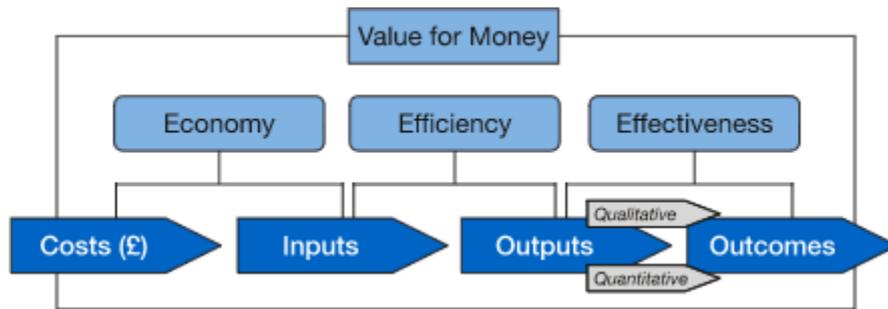
7. Four levels of effectiveness analysis and four corresponding levels of efficiency analysis are thus identified. They will be referred to as *activities effectiveness*, *output effectiveness*, *outcome effectiveness*, and *impact effectiveness*. The corresponding concepts for efficiency are *activities efficiency*, *output efficiency*, *outcome efficiency*, and *impact efficiency*. The top arrows express that in efficiency analysis the comparison is between inputs and a result level, whereas the bottom arrows express that in effectiveness analysis the relationship is one of causality with the previous level.

8. We will refer to this as the “Preferred Standard Terminology” (PS terminology) and contrast this with the “Economy, Effectiveness, and Efficiency Terminology” (“Three Es terminology”) currently popular with donor agencies. There are some concepts used by WFP and other donors that are easily integrated in the former. Take the distinction between “doing the right thing” and “doing things right”. “Doing the right things” refers to strategic choices, for instance between alternative modalities of food assistance (food-in-kind, vouchers or cash), or between alternative ways of boosting learning outcomes (reducing class

sizes, increasing teacher salaries, or providing school meals). The efficiency tools to do this, using PS terminology, are outcome and impact efficiency analysis. But aid agencies must also “do things right”. This is about implementation, and the tools here are effectiveness and efficiency at the level of activities and output. Another concept that is easily reconciled with PS terminology is “Value for Money (VfM)”. We interpret it as synonymous with impact efficiency. Impact efficiency requires that aid agencies both “do the right things” and “do things right”. This is also the approach taken in DFID where value for money is described as maximizing “the impact of each pound spent to improve poor people’s lives” (DFID 2011: 3). And it is also how it is mostly understood at WFP, although there is some unease about the concept (Michiels 2012 : 11-12).

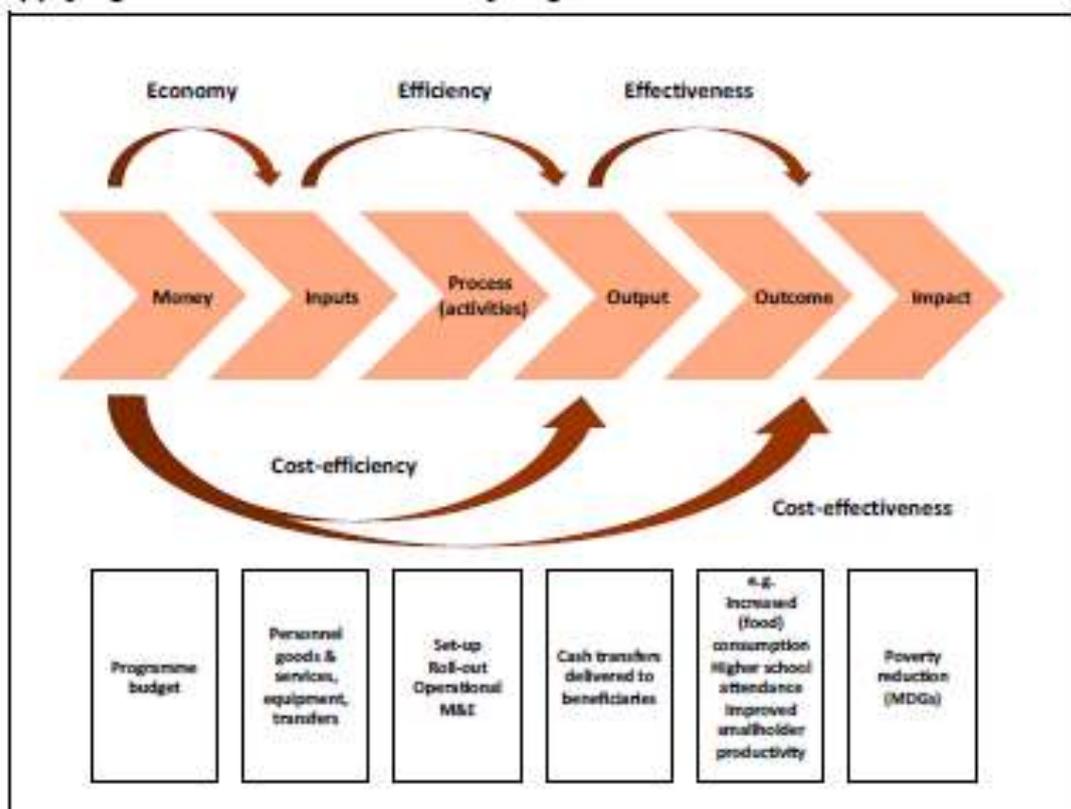
9. As already indicated, many aid agencies, including WFP, use a “three Es” terminology that is claimed to be consistent with DAC terminology but that is in fact difficult to square with PS terminology and the economic literature. Figure 2 and Figure 3 show some recent examples from DFID, probably the bilateral aid agency that has invested most in getting its evaluation concepts clear:

Figure 2 Value for money and the logframe at DFID



Source: Barnett et al. (2010: 6)

Figure 3 The “three Es” as interpreted by DFID



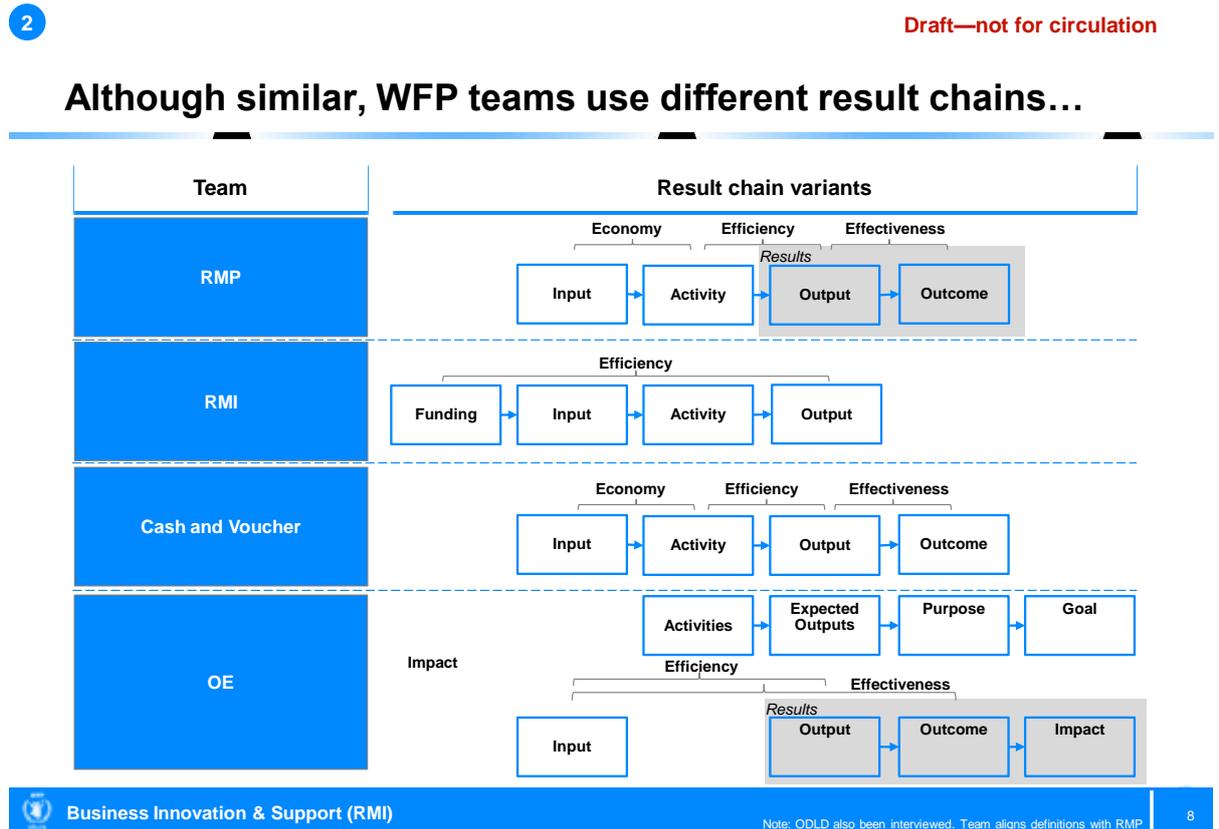
Source: Hodgson et al. (2011: 5)

10. The “three Es” of economy, efficiency and effectiveness appear in both examples. At first sight they seem in line with the DAC definitions, just adding another layer of understanding. On closer inspection, there are awkward tensions with the broad definitions of effectiveness and efficiency. There are two sets of problems. First, the “three Es” fail to acknowledge the simple but important fact that the notion of efficiency applies to all stages of the logical framework beyond input, rather than just to output, and that there is some inherent hierarchy between these stages. Secondly, they obfuscate a fundamental distinction in evaluation theory between two questions: what is the effect of the intervention at a certain

stage in the logical framework (effectiveness), and could we have achieved the results at lower cost by choosing another alternative (efficiency).

11. At WFP a uniform set of definitions has not yet been adopted. Here is an overview of how efficiency and effectiveness are at present defined in different units at WFP. Figure 4 is copied from an internal WFP presentation.

Figure 4 Efficiency notions throughout the logical framework at WFP



Source: WFP(2012c: slide 8)

12. Again the notion of efficiency is narrowed to the output level. The recent change in WFP strategy illustrates why this is unhelpful. In its move from food aid to food assistance, WFP has broadened its modalities from food in kind to vouchers and cash. In order to make appropriate choices and justify to its own donors that aid is well used, the organisation must think through how in a given country setting each of the three modalities affect human wellbeing. Let us concentrate on how this can be verified in ex post evaluation and start with the effectiveness part. For each modality that has been implemented the effectiveness at impact can be analysed. An appropriate indicator may be the number of deaths averted. Likewise, the effectiveness at outcome level can be analysed. An interesting indicator may be the increase in nutritional intake by beneficiaries. Effectiveness at output level can also be measured, for instance through the nutritional value of the food made available to targeted beneficiaries. Next we can perform efficiency analysis. We start with impact efficiency: if the three modalities were implemented and it is found that all three modalities were in their different ways effective at averting deaths, we can calculate the dollar cost of every death

averted. The modality for which this metric is the lowest is the most efficient. If instead of impact efficiency we measure outcome efficiency, we can calculate the dollar cost of a given increase in nutritional value intake, and see which modality reaches this result at the lowest cost. Note that if we use this outcome efficiency analysis as a basis for future interventions or for accountability purposes, the implicit assumption is that for the three modalities a unit of outcome is being translated into units of impact in a way that does not change the ranking of the modalities in terms of cost-effectiveness. If this assumption is not satisfied, outcome efficiency is a misleading guide to impact efficiency. Finally, output efficiency can be estimated, for instance by calculating the dollar cost per unit of nutritional value provided. But for output efficiency to be a reliable guide to value for money, there is again the assumption that the nutritional value offered to beneficiaries is translated into outcome and then into impact in ways that do not alter the ranking of the modalities. These assumptions about the transformation of outputs into outcomes, and outcomes into impact, are however problematic. There are many reasons why a ranking might change in the process of moving from output to outcome to impact⁹. For this reason it is imperative to rely on the insights from the theoretical and empirical literature to make appropriate assumptions in cases full impact efficiency analysis cannot be performed. What this description should make clear is that different levels of efficiency analysis are related to each other, and that there is a natural hierarchy among them¹⁰. By defining efficiency as relating to the output level only, this crucial insight risks being lost and the impression may be conveyed to staff and to evaluators that efficiency type reflections need not be pushed to the higher levels. This is most unhelpful.

13. The other feature is that the terminology used at WFP tends to confuse effectiveness and efficiency. This leads to a sloppy use of language, and worse, to situations where efficiency issues are being addressed with effectiveness issues or vice-versa¹¹.

⁹ If for instance mothers have more control over how food-in-kind is allocated within the household, and fathers more control over cash transfers, then an identical nutritional output may lead to different outcomes for undernourished children under these two modalities. Or an identical nutritional outcome may lead to a better impact in maternal and child health and nutrition (MCHN) than in school feeding because the former is conditioned on attending a medical health clinic.

¹⁰ In the economics literature it is taken for granted that efficiency analysis aims for the highest level in the logical framework, outcome as a minimum, and impact where possible. Efficiency analysis that is restricted to the output level is frowned upon (Drummond et al. 1997: 102-103; Levin and McEwan 2001: 111-112). From this perspective it is particularly odd that so many donors are so focused on that lower level, to the point of banning the use of the term efficiency at higher levels altogether!

¹¹ Consider the following WFP description of cost-effectiveness cited in Michiels (2012: 9): “Cost effectiveness looks at whether it makes sense to carry out the activity/project or not given its capacity to achieve development objectives and the prevalent conditions (transportation constraints, etc.)”. This is confusing to anyone familiar with the economic literature where the concept of cost-effectiveness has a very different meaning. As we will see below, a major limitation of cost-effectiveness analysis compared to cost-benefit analysis is exactly that it does *not* allow to say that an activity is worth undertaking or not. Or take the distinction between efficiency and effectiveness. In an Excel file called “Background Conclusion sheet ex-post” apparently used for internal monitoring of programmes at WFP, one worksheet is on efficiency, another on effectiveness. Given the narrow definition used by WFP, one would expect that most of the 19 monitoring questions raised in the worksheet “efficiency”

14. We strongly recommend that OE keeps the broad DAC definitions in mind, and invites its evaluators to address the whole logical sequence, from inputs to impact, in making efficiency assessments.

3. The problem of attribution

15. The “three Es” terminology presents efficiency and effectiveness as belonging to different stages in the logical framework. In the figures 1 to 4 above this corresponds to a vertical decomposition of the figure. In contrast, the broad DAC definitions are based on a horizontal decomposition. It is this horizontal perspective that we need to adopt to understand the crucial issue of attribution, and how it relates to efficiency analysis. There is a considerable literature on the rigorous measurement of effectiveness, in particular at impact level, drawing on analysis from such disparate scientific fields as medicine, psychology and education, and from the method of scientific investigation in applied (laboratory) sciences. The problem is as follows. In principle we want to compare the results of an intervention with what would have happened in the absence of the intervention. But this counterfactual cannot be directly observed. It can at best be estimated indirectly. A random assignment of potential beneficiaries to an exposed and a control group, taking into account relevant characteristics such as age, gender, education, wealth, and income, is the best way to ensure that both groups are as similar to one another as possible. Such an experimental design is the gold standard in science, and has in the recent past been vigorously advocated in the field of development aid (see e.g. Savedoff et al. 2006, and Banerjee and Duflo 2011). In its absence, a host of less robust quasi-experimental designs are available, but they are less precise, not less complex analytically, and require data that is not routinely collected by either donor agencies or national governments. The intrinsic difficulty of effectiveness analysis tends to increase as we move up in the logical chain from activities to output, outcome and impact. The previous section has provided illustrations from the line of work of WFP. We can establish relatively easily whether WFP has been effective at activities level (e.g. whether food is available in a central warehouse in the country). It is much more difficult to establish that the right food has been distributed at the right time to the right beneficiaries, and thus that output effectiveness has been achieved. And there is concern about whether WFP achieves outcome effectiveness: are beneficiaries really better nourished than they would have been in the absence of the food aid? And whether impact indicators of malnourishment such as stunting have gone down as a consequence of the WFP intervention, has not been established in any convincing manner. Indeed, WFP has come in for considerable criticism for what critics consider an inadequate M&E system and an inability to adequately address effectiveness concerns (Harvey et al. 2010: 50-51). Yet some excellent analysis is performed. Annex A contains a summary of some of the methodologically robust studies that have been produced in recent years at WFP or in collaboration with WFP, and the list is incomplete.

16. This is important because efficiency analysis builds on effectiveness analysis at every step in the logical framework, as Figure 1 has illustrated. Was the right food available for distribution at central warehouses, and was it protected against pests? Was it delivered in time to the point of final delivery? Were beneficiaries well targeted? How many beneficiaries received how much food aid? Was there sharing with non-targeted family members? Was

would address what in our terminology we label output efficiency, but this is not the case. They instead overwhelmingly address issues of activities and output effectiveness.

some of the food sold, and if so, was the female household head in control of the proceeds? How well educated was the mother? How knowledgeable was she about preparing the food so as to preserve its nutritional effect? Was there enough safe water and fuel to cook the food? Did illness not prevent children from gaining strength from the additional food consumed? Only after knowing the answers to such effectiveness queries can we start calculating efficiency measure such as the cost per food ration delivered, the cost per unit of nutritional value obtained, or, most challengingly, the cost per unit of decrease in stunting or the cost per death averted.

17. The author of the BMZ study acknowledges the problem: “Some of the experts interviewed for this report described the analysis of attribution as the single most challenging hurdle for reliable effectiveness and efficiency analysis” (Palenberg 2011: 19). But the author argues that this is beyond the scope of his study. This is fair enough, but when devising ways for WFP to increase the efficiency analysis of its interventions, this prior need to improve the volume and quality of effectiveness analysis, both through better monitoring of results and through more rigorous analysis of the results, should be kept in mind, and arguably constitutes the most important domain for improvement. To give an example, cost-effectiveness ratios (CERs) are a useful instrument for assessing efficiency, as we will later explain in more detail. CERs can be applied at the level of output (cost per nutritional unit delivered), outcome (cost per nutritional unit consumed), or impact (cost per death averted). In all three calculations the denominator is basically the same. The fact that at WFP unit costs are only occasionally produced at output level, seldom at outcome level and almost never at impact level is not due to the difficulty of calculating the cost side, or a failure of evaluators to grasp the economics underscoring efficiency analysis, but is due to a failure to properly monitor and analyse the benefits of WFP interventions, in other words an issue of effectiveness analysis.

4. A taxonomy of WFP interventions

18. The vast majority of WFP interventions have nutritional improvements as an objective. In Table 1 below we propose a generic logical framework for such interventions. We mention, in *italic*, some other dimensions, such as educational achievements (school feeding), better health through attendance of medical clinics (mother and child feeding) or farm income (P4P), to illustrate that many WFP interventions are multi-purpose, a fact that complicates efficiency analysis. In such cases the logical framework branches out, but even in such cases nutrition is a key link in the causal argument. The table contains a list of often encountered indicators of effectiveness and efficiency at different levels of the logical framework, both for nutrition and the other dimensions.

Table 1 Effectiveness and Efficiency in the Logical Framework for WFP

	Logframe	PS effectiveness indicators	PS efficiency indicators
Impact	Nutritional status improved Sustainable human development broadly defined	<ul style="list-style-type: none"> Δ weight for height (short term measure) Δ height for age (long term measure) Δ BMI # deaths averted QALYs Δ household income 	<ul style="list-style-type: none"> Cost per quality-adjusted life year or QALY (Cost utility analysis) Net benefits using value of statistical life (VSL) estimates Cost per 1% improvement in nutritional indicator Beneficiary satisfaction scores Equity scores
Outcome	<ul style="list-style-type: none"> Food security enhanced <i>Improved education (learning outcomes)</i> <i>Mother-and-child health improvement</i> <i>Farmer income increased (P4P)</i> 	<ul style="list-style-type: none"> Δ food diversity index Δ food consumption score Δ food consumption group Δ nutritional value score (NVS) <i>Δ farmer income per capita</i> <i>Δ school enrolment and dropout (gender differentiated)</i> <i>Δ end of schooling cognitive scores</i> 	<ul style="list-style-type: none"> Cost per 1 % improvement in summary food security indicator Cost per kcal consumed Cost per nutritional value unit consumed
Output	<ul style="list-style-type: none"> Aid (food, cash, vouchers) delivered <i>Assets created</i> <i>Disaster risk mitigation (DRM) measures in place</i> 	<ul style="list-style-type: none"> # beneficiaries reached # MT food distributed # Kcal distributed # standardised daily nutrition rations distributed # cash/vouchers distributed # units of FBF as % total food distributed <i># community assets created</i> <i># of schools assisted, also as % of total</i> <i># pupils trained</i> Beneficiary scores of project and staff Description of DRM measures realised 	<ul style="list-style-type: none"> Cost per recipient Cost per standardised ration delivered Food market value of transfer of food commodity/cost of the transfer of food commodity (alpha ratio) Cost of transfer of \$10f food or food purchasing power (1/alpha) Delivery cost per 1\$ of food or food purchasing power (1/alpha -1) Cost per kcal provided Cost per nutritional value unit provided Cost per nutritional value unit provided modality 1/cost per nutritional value unit provided modality 2 (omega ratio)

Efficiency Analysis for WFP – Definitions, Concepts and Methodology

	Logframe	PS effectiveness indicators	PS efficiency indicators
Activities	Needs assessment Preparation of cash, vouchers modalities Securing food (donor, triangular, local) Transport, LTSH Food fortification (FBF) Capacity building Monitoring, ...	MT of food procured/warehoused # months of delays in construction Food “pipeline break” statistics Staff turnover statistics Staff recruitment statistics Timeliness of (donor and government) contributions in relation to the budget cycle	Cost per MT, broken down for commodities, transport, LTSH, DSC, ODOC, ISC, costs borne by government or beneficiaries CBA of production of FBF, financial analysis of warehouse construction, cost analysis of truck transport, ... Administrative costs of raising funds
Input	Donor financial support through WFP Donor food in kind donated through WFP Partner government resources Beneficiary contributions	Donor funding (actuals) Government contributions (actuals) Beneficiary contributions (actual)	Not applicable ¹²

Source: author’s elaboration

¹² This is not to argue that inputs can be taken for granted in WFP interventions, like manna from heaven. On the contrary, securing sufficient and timely contributions by donors and governments is an important challenge, involving costs of its own, and failings in this respect may affect the efficiency of operations in a major way. For reasons of the syntax of the logical framework these considerations are here listed in the row above, of activities, the level at which they start to have an effect.

4.1 Effectiveness analysis of WFP interventions

19. In Table 1 the indicators listed at **activities** level were taken from different WFP documents. Not mentioned for reasons of space is technical assistance to the partner government. This will become more important as partner countries reach the status of Middle Income Countries while poor citizens remain food insecure (WFP 2012a). In these circumstances WFP's role is increasingly to help the government improve its policies rather than implementing policies on its behalf.

20. The overwhelming majority of WFP projects have as an **output** “measures that directly support the consumption of food”, a useful description of food assistance borrowed from Harvey et al. (2010: 28). The output level indicators in Table 1 were all found in WFP documents. As elsewhere in the table, we do not aim at an exhaustive list.

21. The **outcome** of WFP interventions is mostly about nutritional intake. Outcome indicators thus measure the improvement in nutrition that can be attributed to the intervention. We came across the following useful indicators in our reading of WFP documents: food diversity index, food consumption score, food consumption group, and nutritional value score (NVS). Maybe there are others. In school feeding learning outcomes and attendance of children from poor families are also important. Other outcomes include women empowerment (e.g. enticing families to send their girls to school), health improvements through attending health clinics (mother and child supplementary feeding), and income of farmers (P4P), but we did not pursue these points in more detail.

22. If the outcome is about nutritional intake, corresponding indicators of **impact** are improvements in weight for height (short term measure), height for age (long term measure), or changes in BMI (for adults). And of course lives saved, and quality adjusted life years (QALYs), a concept explained further in the text. There are other indicators that may be used in some circumstances.

4.2 Efficiency analysis of WFP interventions

23. Many non-food interventions undertaken by governments, NGOs and other donors *indirectly* support the consumption of food. For instance, if an irrigation project increases family income of the participating farming households, a considerable part of that extra income will normally be devoted to food, because of the high income elasticity of demand for food among the poor. Similarly, microcredit targeted at women may have a significant impact on child nutrition. The same is the case for many other types of projects and programmes undertaken by countless multilateral, bilateral, and non-governmental organisations. Indeed it is difficult to imagine successful interventions aimed at poverty reduction that do not have some positive effect on food consumption. The same is true for growth in per capita income in general (see e.g. Headay 2012). And there is no a priori reason why the direct approach favoured by WFP would be more advisable than the indirect approach of most other agencies. In this regard efficiency analysis will have to demonstrate that, in comparison with such indirect alternatives, the WFP food assistance approach provides “Value for Money”. There is in principle an important exception to this need to allow for comparison with indirect alternatives: when people suffer acutely from undernourishment and timeliness is of the greatest importance, indirect approaches are less effective, and thus less efficient. In such emergency circumstances no efficiency comparison with indirect methods must be undertaken, and the efficiency analysis can concentrate on alternative modalities of direct food assistance. Note however that a sharp divide between humanitarian and development

interventions does not exist. Reality is best described as a broad spectrum with pure humanitarian and pure development interventions at the extremes and most WFP interventions being situated somewhere in between these extremes, exhibiting features of both. Hence the need to carefully consider what the alternatives are in any efficiency analysis of WFP operations.

24. Efficiency analysis can be performed at different levels, from **activities** upward. Indicators are mentioned in the last column of Table 1. WFP is commendably attentive to the efficiency of its logistics chain at activities level, and has elaborate procedures to minimise the cost of procurement, transport, local transport and warehousing while maintaining minimum standards of timeliness and quality. To monitor this, WFP among other things uses a breakdown in commodities, transport, LTSH, DSC, ODOC, and ISC. But it does not routinely include costs borne by government and beneficiaries.

25. At **output** level, WFP collects systematic information that allows to calculate cost per MT and recipient. The number of recipients are apparently not always comparable because no distinction is made according to length of time or frequency of assistance provided. This is a key failure in the M&E system at WFP. From a methodological perspective, the following are more interesting than cost per recipient: cost per standardised ration delivered, cost per kcal provided (although less relevant if the main focus is on nutrition), and cost per nutritional value unit provided. An important output efficiency indicator is the so-called alpha ratio, but it has gone out of fashion (as mentioned in the ToR). We return to this later.

26. The **outcome** indicators presented in Table 1 all come from WFP studies: cost per 1% improvement in summary food security indicator, cost per kcal consumed, and cost per nutritional value unit consumed of WFP interventions.

27. The most important efficiency indicators at **impact** level are CBA using value of statistical life (VSL) estimates for which the World Bank has proposed a simple way of calculating it across countries, and the cost per QALY. We will discuss these measures in more detail in the following two sections.

5. Allocative efficiency analysis (Cost-Benefit Analysis)

28. In sections 5 and 6 we draw on the scientific literature on efficiency. The distinction between the two sections follows a suggestion in the BMZ study and is also consistent with the treatment in most economic textbooks on the topic.

29. Efficiency analysis consists of a systematic comparison of the disadvantages (costs) and advantages (benefits) of an intervention. For the sake of both scientific rigour and precision of the decision tool, costs and benefits are expressed in commensurate quantitative terms. Costs and many benefits (e.g. increased farmer income in P4P) are naturally expressed in money terms, and the use of monetary units throughout is the gold standard in economics. Many efforts have been made to extend the principles of cost-benefit analysis (CBA) to the specific situation of developing countries, such as poorly functioning capital, labour and commodity markets and the need to take equity issues into account¹³.

¹³ Palenberg (2011: 63-77) provides a non-technical survey and a useful bibliography for further reading.

30. CBA, properly conducted, answers the fundamental “value for money” question of whether WFP interventions produce net developmental benefits. For this reason, it is considered a more powerful tool of efficiency analysis than some related but simpler techniques that will be discussed in the next section. The BMZ study gives it the label of “optimisation efficiency”, contrasted with less powerful “transformation efficiency” (Palenberg 2011: 16). The distinction is important, but the terminology proposed is not standard and not wholly convincing because all efficiency analysis is about optimisation, and all efficiency analysis is equally about transformation of inputs into results. We prefer to stick to the standard terminology used in economic textbooks such as Boardman et al. (2011: 484) where the class of tools to which CBA belongs is labelled “allocative efficiency analysis”, and the less powerful variety “technical efficiency analysis”.

31. Crucial WFP benefits are difficult to express in money terms. Most WFP interventions aim to save human lives, but there is no market where life years or quality of life are traded across individuals or across time and where prices can be observed. Therefore any monetary valuation of life is indirect. There are basically three techniques that can be used. The *human capital approach* calculates the income a person would have earned in the additional life years, and postulates this as a minimum estimate of the value of this life. The *stated preference approach* uses interview techniques to let people indicate the monetary values they attach to a set of hypothetical situations involving life, illness and death. The *revealed preference approach* studies the actual behaviour of individuals in risk-related markets¹⁴. This third method can be applied to public sector interventions. Governments and international donors spend (or force individuals to spend) money to reduce the risk of accidents and death. Safety standards in new cars and buildings, technical controls of used cars, stiff penalties for traffic rule violations, spending on hospital ambulance services, efforts to eradicate communicable disease vectors are all examples. Implicitly, by spending so much, and not more, on any such measure, public authorities assign a monetary value to human life thus saved that can in principle be calculated. Unfortunately, the three methods have theoretical and empirical limitations. In addition there are moral reservations about the whole idea of putting a monetary value on life.

32. Many estimates of the value of a statistical life exist for developed countries, but hardly any are available for developing countries. Cropper and Sahin (2009), in a background study for the joint World Bank – UN work on disaster risk reduction, suggest a quick and rough method for deriving the value of a statistical life for developing countries from that in the US on the basis of income per capita data. We advise against using these estimates for comparisons across countries, as this would imply that WFP attaches a higher value to lives saved in countries that are more developed. But for within-country analysis the approach can sometimes prove useful.

33. Several alternative **decision rules** are available in CBA. The most simple of these is just to subtract the costs from the benefits. A project passes the efficiency test if the resulting net benefit is positive. Because projects often have a time horizon of several years, discounting is applied to express all costs and benefits in their equivalent value at the start of

¹⁴ If a given job increases the daily risk of death by 4/10000, and an additional daily salary of \$30 is sufficient to convince workers to take on such a job, the implicit value of a statistical life can be calculated (in this particular case it is \$75,000).

the project. This leads to the net present value (NPV) criterion: if the NPV is positive, the project passes the efficiency test. Alternatively, the (discounted) benefits can be divided through by the (discounted) costs to obtain the (discounted) benefit-cost ratio (BCR). If the BCR exceeds unity, the project passes the efficiency test¹⁵. Finally it is possible to do a sensitivity analysis by calculating the switching value for one of the economic parameters used in the calculation. The best known such criterion is the internal rate of return (IRR). If the IRR exceeds the chosen discount rate, which expresses the opportunity cost of capital invested in the project¹⁶, the project passes the efficiency test¹⁷. All these are simple algebraic transformations of the NPV rule, and all will normally give the same yes-or-no answer to the question whether a project is worth doing¹⁸.

34. Sometimes a simple yes-or-no answer is insufficient. The decision on several projects may be linked, either because they are mutually exclusive (as when a WFP CO hesitates between food, cash or voucher transfer modalities to address a given food insecurity situation: by choosing one modality, or a combination, other modalities are ipso facto excluded) or subject to a constraint (as when a WFP CO has identified a series of not mutually exclusive projects in different parts of the country, all with positive NPV, but exceeding the total budget that donors are willing to put up). Among mutually exclusive projects, the rule is to choose the one with the highest NPV. In case of a single input constraint, the rule is to calculate the ratio of the NPV to the number of the constrained inputs required, rank projects in decreasing order, and select projects until the input is exhausted^{19,20}. Here is a crucial link with cost-effectiveness ratios (CER) discussed below and arguably the major tool suited for WFP: it can be shown that under certain not unrealistic circumstances a CER gives the same ranking as the NPV per unit of the constrained input. We return to this point below.

6. Technical efficiency analysis (Cost-Effectiveness Analysis)

35. When the benefits of a project cannot be expressed in monetary terms, but it is possible to express them in some other metric (number of daily food rations delivered, nutritional value units delivered, number of cases of child stunting avoided, number of lives

¹⁵ The reverse of the BCR, the cost-benefit ratio (CBR) can also be calculated. The decision rule here is that a project passes the efficiency test if the CBR is less than unity.

¹⁶ In fact, the discount rate is often called upon to express other things as well: inflation, risk aversion, and intertemporal equity. Obviously one and the same parameter cannot satisfactorily express all these. How this overload is addressed in CBA is beyond this document.

¹⁷ Another example, not discussed by Palenberg, is the domestic resource cost (DRC) often used in the evaluation of agricultural projects in developing countries. A project passes the efficiency test if the DRC is lower than the (shadow) exchange rate.

¹⁸ Palenberg (2011: 68-73) draws attention to the possible pitfalls of using the IRR, and explains why academic researchers argue that the only fully reliable decision rules are NPV-based (NPV and BCR).

¹⁹ Palenberg (2011: 16) erroneously advocates selecting the projects with the highest NPV in such a case.

²⁰ If there is more than one input constraint (say both a budget constraint and a lack of critical WFP staff to implement all desirable projects) a more complex mathematical algorithm is needed to solve the allocation problem.

saved, number of quality adjusted live years saved, ...), then a cost-effectiveness ratio (CER) can be calculated, expressing the cost per unit of result. Obviously, projects with lower CERs or unit costs are to be preferred to projects with higher CERs. CERs of different projects can be meaningfully compared as long as the metric used in the numerator (e.g. discounted 2012 US dollars) and in the denominator (e.g. % change in the dietary diversity index) are the same everywhere. We have seen above that ratios are also used in allocative efficiency analysis in the form of benefit-cost ratios (BCR) or their reverse (CBR)²¹. CERs are closely related to CBRs. The difference between the two is that in the CBR denominator benefits are expressed in money terms, whereas in the CER denominator they are expressed in a non-monetary metric. So what is lost when we express benefits in a non-monetary metric instead? Have we stumbled on an alternative efficiency approach that has the advantage of being more simple to calculate and of being less controversial, as many evaluators believe? Yes, but there is a caveat.

36. The CER is closely related to the CBR but has three drawbacks:

- a) It misses the CBR's ability to establish the desirability of an intervention. To say that a project is worth doing because the cost expressed in dollars (CBR numerator) is lower than the benefits expressed in the same dollars (CBR denominator) makes perfect sense. But it does no longer make sense to say that it is worth doing because the cost expressed in dollars (CER numerator) is lower than the benefits expressed in food rations (CER denominator). So we lose the capacity to perform allocative efficiency analysis. This does not matter if we are so convinced that the intervention we are studying is worth doing that no calculation is needed to justify it, in which case the only question is how best to implement it. This is probably the case for most EMOPs, but questionable for most PRROs and development interventions of WFP.
- b) The CER, like its close relative the CBR, is an appropriate ranking device only in case of a budget constraint. When there is no constraint, the ranking it provides is only correct if all projects are scalable so as to provide the same level of results, e.g. nutritional intake per person. This need not be the case, witnessed by WFP studies comparing cash, vouchers and food in kind. It must therefore be checked in each case whether this is a reasonable assumption to make.
- c) It must be possible to express all significant benefits either in monetary terms (in which case they can be included as negative costs in the numerator) or in the metric of the denominator²². If the three conditions are fulfilled, the use of CERs or unit costs is a full substitute for CBA. For this reason evaluation based on CERs is considered a valid but lower form of efficiency analysis in the scientific literature. Boardman et al. (2011: 484) call this "technical efficiency", which, as indicated above, we prefer to the expression "transformation efficiency" proposed in the BMZ study.

²¹ The BMZ study erroneously associates allocative efficiency with decision rules involving subtractions and technical efficiency with decision rules involving ratios (Palenberg 2012: 16). There is no basis for this claim.

²² P4P procurement is a case in point. The benefits consist of both increased income for local farmers and increased food security of food aid beneficiaries. By treating increased income to farmers as a negative cost, a CER calculation can be performed.

37. Depending on which non-monetary metric is selected, a CER can express activities efficiency (e.g. cost per tonne of food brought to a warehouse near the point of food aid distribution), output efficiency (e.g. cost per food ration delivered), outcome efficiency (e.g. cost per nutritional improvement realised), or impact efficiency (e.g. cost per life saved). As already argued above, there is a natural priority ordering of these different CER calculations. The higher the indicator is situated in the logical framework, the better. In case there is a contradiction in two rankings, the CER that is higher in the logical framework trumps the lower one. For instance, selecting interventions with the lowest cost of procurement, transport and warehousing per MT (an application of activities efficiency analysis) may increase the time before food reaches the point of delivery. In an emergency, this may cost lives. If instead the cost per life saved is minimized (an application of impact efficiency analysis), this bias is overcome because the effect of timeliness will in principle be fully reflected in the measure of lives saved. And it is obvious what the best alternative is: the project design that achieves the best impact efficiency will be preferred (“value for money”), even if its activities efficiency is lower than that of another design. To give an example from WFP itself, in a case study on Malawi, poorly functioning local markets meant that food in kind had a better output efficiency than cash transfers, but when the analysis was repeated at outcome level, cash transfers had a better nutritional efficiency (Oadsley et al. 2009). Obviously, in this case the cash transfer programme is to be preferred. At least, if we have reasons to believe that the better ranking of the cash programme would be maintained if we had been able to do an impact efficiency analysis.

38. It does not follow from the preceding argument that the use of lower level CERs should be shunned. There is a trade-off, in that higher level CERs tend to be more difficult to compute. The difficulty does not reside in the efficiency analysis per se but in the underlying effectiveness analysis. In fact, the numerator expressing the costs will often be the same in both cases. It is the denominator, expressing the benefits in non-monetary terms that usually causes the trouble. As explained in more detail above, attribution is the major culprit. It becomes more complex as we proceed to the higher levels in the logical framework of a project. So much so that efficiency analysis may no longer be possible or worth the cost. On the other hand the assumption that the transformation of outputs into outcome and impact is so comparable that no ranking reversal take place at higher CERs should not be made lightly.

39. Arguably the most sophisticated non-monetary indicator used in CER is “quality adjusted life years saved” (QALY). It was developed in the health sector as a non-monetary metric at impact level. The indicator takes into account both the length of life saved and its quality. If two three-year old children are saved from starvation, but only the first has been reached in time and no lasting harm to its long term health occurred, then the benefits of the same food rations are different. The first child will have a longer life expectancy than the second (quantity), but it will also have a healthier and more productive life (quality). The QALY captures both dimensions. A powerful advantage of the QALY indicator is that it allows ranking health projects whose output and intermediary outcome indicators are different but whose final outcome indicator is the same. An example is a project that vaccinates children against measles and another project that offers curative treatment for HIV AIDS. In both cases the final benefit is length and quality of live years saved, but the indicators at lower levels are different. QALYs can be compared not only across interventions and sectors, but also across countries. The latter is not the case when a monetary value is attached to life.

40. As mentioned in the BMZ study, two special cases of efficiency analysis are **cost minimization** and **yield maximisation** (Palenberg 2011: 21-22). They are simpler to apply than CERs as in the latter both costs (numerator) and benefits (denominator) have to be estimated. Cost minimisation is an appropriate decision rule when comparing alternatives that have identical results, whereas yield (benefit) maximisation is an appropriate decision rule when comparing alternatives that have identical costs. Obviously this will not often be the case, but if it happens, efficiency analysis is much simplified. There is some tendency among evaluators who are not economically trained to use such measures in cases where they are not fully justified, so it bears repeating that these are only appropriate decision rules when the assumptions mentioned are satisfied.

41. Using CERs or unit costs always involves comparisons. Typically these comparisons are within countries and restricted to a particular type of intervention, for instance emergency food aid. But at WFP and elsewhere broader comparisons across types of interventions and across countries are also being made. A good example is provided in the World Development Report 1993. It proposes a series of **league tables** containing the cost per DALY (Disability Adjusted Life Year)²³. The data are most often averages for all developing countries, but sometimes a distinction is made between low-income and middle-income countries (World Bank 1993, in particular the tables and figures on pages 62, 82, 106, and 117).

42. Cost per QALY or cost per death averted are relevant for WFP as well, as nutrition is the main underlying mechanism, and as nutritional impacts can be expressed in such terms. Although QALYs are not easy to estimate, an impressive literature from the health sector provides many examples of how it can be done. The results of hundreds of painstaking empirical studies have led to a fairly good capacity to predict QALYs for a broad range of interventions. Once QALYs have been estimated, CERs can be calculated fairly easily. Note that from here it is only a small further step to full “allocative” efficiency analysis using CBA. See our comment on the use of full CBA at WFP at the end of the previous section.

7. Other considerations in evaluating WFP interventions

43. EQUITY is an important consideration in evaluating donor projects in developing countries. One possibility is to integrate equity directly in the efficiency calculation. Benefits going to poorer people can be assigned a higher weight than those accruing to the rich, as is done in social cost-benefit analysis. But such social CBA, fully developed in the 1970s, and methodologically very elegant, is seldom applied in practice because it is too demanding in data and also because it requires explicit value judgements. Another possibility is to compute partial efficiency measures only focusing on the poor. The BMZ study describes both methods of equity analysis in the section on allocative efficiency analysis (Palenberg 2011: 72). We add that a similar procedure can also be applied to technical efficiency in the form of CER calculations focusing exclusively on the poor. A third approach is to leave efficiency analysis as it is and score projects separately on their redistributive qualities, and then apply multi-attribute decision models. See sections 8 and 9 for more information.

²³ DALYs are a specific approach of calculating QALYs, developed for the 1993 World Development Report, and since that time regularly published by WHO.

44. At first sight equity considerations may seem to trouble evaluators at WFP somewhat less than for instance their colleagues at FAO or IFAD. Consider two irrigation projects their colleagues may be involved in, one targeting poor illiterate farmers and the other literate commercial farmers. The first may score better on equity, the second on efficiency. It may appear that WFP is less confronted with such stark choices. The reason is that WFP projects overwhelmingly target and benefit the poor. To that the extent, the equity-efficiency trade-off in WFP interventions may be less sharp. But on closer inspection the equity problem is often very present. For instance, food-in-kind and vouchers are said to be less prone of being captured by the not so poor than cash transfers, because only the needy will take up a form of assistance with strong social stigma attached. And in school feeding the assumption usually is that if you feed a school you have to feed everyone in it. Equity issues then become crucial through decisions of which schools or which school districts to include or exclude from the programme. Similarly, with many programmes seriously underfunded and with stark targeting choices having to be made, equity aspects may become key aspects of efficiency analysis.

45. ACCOUNTABILITY TOWARDS BENEFICIARIES can be included alongside efficiency analysis by applying multi-attribute decision models, or, more simply, by a separate analysis that is joined to the efficiency evaluation. See next sections on the tools that the BMZ study proposes in this respect.

8. The BMZ efficiency toolbox: what is less useful for WFP

46. The ToR explicitly refers to a recent BMZ study (Palenberg 2011). This is indeed an exhaustive overview of possible evaluation tools to assess efficiency. Table 2 lists the 15 tools discussed in the BMZ study. They are rated in three categories by Palenberg, according to their “analytical power”: level 2 (allows for an overall estimate of the efficiency of an intervention compared to other interventions), level 1 (focuses on some partial aspect of efficiency, typically within a particular intervention), and level 0 (provides a qualitative but inevitably subjective expert assessment on certain dimensions pertaining to efficiency).

Table 2 An overview of BMZ efficiency tools

	Level 2	Level 1	Level 0
Well-known methods	CBA	Benchmarking of unit costs	Expert judgement
Somewhat less well-known methods	CEA	Follow the money	
		Financial analysis	
		Stakeholder-driven approaches	
		Benchmarking of other indicators than unit costs	
Methods unknown to a substantial fraction of evaluation experts	Intuitive MADM	Stakeholder effectiveness and cost rankings	Specific efficiency questions
	Scientific MADM		
	Effects method	Stakeholder efficiency ratings	
	CUA		

Source: Palenberg (2011: 92)

47. But not all the 15 tools surveyed in the BMZ study are equally important or relevant for our purpose. We propose not to discuss : expert judgement; financial analysis; stakeholder-driven approaches; stakeholder effectiveness and cost rankings; stakeholder efficiency ratings; multi-attribute decision model (MADM) using scientific scores; and the effects method. This is not a criticism of their inclusion in the BMZ study, but reflects our assessment of their usefulness for WFP. In Table 2 the names of these 7 tools have been struck through. The remaining 8 tools can be considered for use in isolation or in combination in evaluating WFP interventions. We comment on why we suggest to drop the 7 tools in this section, and then discuss the remaining 8 in the next section.

48. What is useful for WFP depends, among other things, on what can be realistically expected in the context of OE, its resources and mandate. To help us get an insight in past evaluation work and what realistic improvements may be expected in the future, OE selected 13 CPE and OpEv completed evaluation reports that had addressed the efficiency questions rather better than average. Annex B contains a description of the relevant sections in these evaluation reports.

49. EXPERT JUDGEMENT (level 0): too descriptive and subjective to be considered a scientific evaluation method of efficiency for which guidelines can be developed. The author of the BMZ study himself expresses grave doubts (Palenberg 2011: 99). This is not to say that there is no room for intuitive judgement. But such judgement is best provided as part of a more robust evaluation. In fact any rigorous evaluation reports will contain some qualitative assessment. For instance, cost-benefit analysis (CBA) needs a narrative description and interpretation of the findings and considerations on how elements which inevitably have not been included in the analysis might influence the overall assessment.

50. FINANCIAL ANALYSIS (level 1): basically a CBA from the perspective of a particular actor (the donor, an implementing agency, the beneficiaries, farmers, ...). Whenever full economic CBA is performed from the perspective of recipient society as a whole, financial CBA is de facto the first step. But if financial CBA can be performed, then full CBA is often within reach. And from a Value for Money perspective (that is: efficiency at impact level) it is of course full economic CBA that is relevant. The use of financial CBA, which is expensive and time consuming, is best considered as a first step of full CBA rather than as a stand-alone tool.

51. STAKEHOLDER-DRIVEN APPROACHES (level 1)²⁴: concerns highly participative methods where stakeholders “provide not only most of the data, but also determine what data should be considered important and what standards efficiency should be compared to” (Palenberg 2011: 55). The author himself acknowledges that the analytical power of this tool is possibly undermined by the fact that (1) stakeholders may find it difficult to see beyond their own interests and may for this reason fail to assess the impact on society as a whole, that (2) they may not have enough information, and that (3) their conceptual understanding may be inadequate. We agree. All three weaknesses in our view apply to WFP projects, and in particular the third element may cause serious methodological problems. Implicitly, food is treated in WFP interventions as what economists call a “merit good”, where consumer sovereignty is discarded in favour of paternalistic standards set by nutritional experts. The nutritional content of food is something that beneficiaries may not be aware of or may not care about, but that WFP judges to be very important. This is not to say that there is no room for listening to stakeholders in evaluation. On the contrary, accountability to beneficiaries is an important topic. But in our view this particular tool, as defined by Palenberg, will not help WFP enhance efficiency in the sense intended in the ToR.

52. STAKEHOLDER EFFECTIVENESS AND COST RATINGS (level 1): This is the milder version of two tools discussed by Palenberg (2011). See the next section for the stronger version. In both cases beneficiaries are asked to rank the results of an intervention. In the milder version discussed here, the questions relate to components of the intervention such as timeliness of delivery, quality of the food, or the cost borne by the beneficiaries or the community. It is left to the evaluator to bring the different elements together in appropriate measures of efficiency, also using other sources of information and making appropriate calculations. After discussion with OE, this tool is not retained. This is not to argue that the opinions of beneficiaries are not important, and should not be gauged. On the contrary, one of the good reasons for doing fieldwork during evaluations is to get a feeling of how beneficiaries feel about an intervention. But it is considered a way of collecting qualitative information rather than an efficiency tool.

53. STAKEHOLDER EFFICIENCY RATINGS (level 1): closely related to “stakeholder effectiveness and cost rankings”, and discussed in the previous section. The same argument for not retaining the milder version applies. In addition, asking beneficiaries to assess efficiency directly is fairly complex to explain and subject to more errors of understanding on their part. Note that the BMZ study has not much to offer in terms of a description of these

²⁴ Not to be confounded with the less ambitious “stakeholders effectiveness and cost rankings” proposed in the BMZ report and discussed in the next section, in which beneficiaries are asked to assess some specific aspects of an intervention that targeted them.

two tools. No references to extant studies are provided and the author admits that they may not be fully established methods.

54. **MULTI-ATTRIBUTE DECISION MODEL (MADM) USING SCIENTIFIC SCORES** (level 2): suppose irrigation project A has a higher NPV than alternatives B and C, but project B has the advantage of favouring poor farmers, and project C is more ecologically sustainable. Scoring the three irrigation projects on each dimension and attaching a weight to each dimension allows an overall score to be calculated. Some projects score better on one of these objectives, some on others. So how should these projects be ranked overall? All the dimensions are relevant from a development perspective, but they cannot be readily aggregated. A solution is for the evaluation report to provide the information separately, and leave it to decision makers to make up their mind. Alternatively, the evaluator can apply the MADM method, proceed to translate the impacts on the different dimensions in scores (say, on a scale from 1 to 10), attach weights to the different scores, and calculate weighted overall scores. This can be done through weights that are selected directly, by the evaluator, or preferably by someone with authority in the organisation. This is labelled the “intuitive” scoring system (Palenberg 2011: 61). We suggest to retain this method for further consideration, and discuss it in the next section. By contrast, a “scientific” scoring system may be developed, whereby for instance members of the WFP executive board or senior management reveal their implicit (and supposedly rational) preference functions by participating in a series of stylised computer simulation exercises where they are exposed to hypothetical decision situations and asked to pick the most desirable decision. On this basis an analyst then derives an overall objective function with certain desirable features such as completeness and transitivity, from which quantitative weights can then be derived. We suggest to drop this tool of scientific scoring as both methodologically problematic and unrealistic. The principle of MADM is hereby not rejected, but only the variety with intuitive weighting is retained.

55. **EFFECTS METHOD** (level 2): a French method developed in the same period and as a competitor to economic cost-benefit analysis which had a more Anglo-Saxon (or more correctly an international) flavour. Requires more detailed data than all but the most sophisticated version of CBA. Relies less on market price information and more on input-output tables. But these tables are seldom available in sufficient detail. Some of the underlying logic can be used in CBA, as for instance in estimating “conversion factors” (Little and Mirrlees 1974), but CBA is more flexible as it mainly relies on market-based information that is easier to obtain and often very rich in insight. And there is the question whether the two methods - the effects method and CBA - are fully equivalent, as is sometimes claimed. There is no convincing positive answer to this. To the extent that they are not, CBA is methodologically superior because it is grounded in optimisation theory²⁵.

²⁵ The effects method can be understood as a way of maximizing GDP. However GDP is an indicator of economic activity, and not necessarily of economic welfare.

9. The BMZ efficiency toolbox: what is useful for WFP

56. In what follows we discuss the efficiency tools that we feel may be useful for WFP in the present context. An overview of the different methods is provided in Table 3. For each tool, we discuss in turn key features, data and skill requirements, applicability to OE evaluations, and a worked-out numerical example.

57. **COST-BENEFIT ANALYSIS (CBA) features.** For the reasons explained in section 5 CBA is the gold standard in economic analysis. The BMZ study rates it at 2, the highest level. CBA has been primarily conceived for application at the level of impact. It can however also be applied for intermediary results, such as warehouses, locally produced fortified food, or emergency road improvements, that can be regarded as sub-projects. CBA sits more uncomfortably at output level, because the benefits of interventions cannot yet be identified and measured at that level²⁶. CBA can be applied throughout the life cycle of an intervention, from the ex ante (decision making), to the mid-term (managerial improvements, strategic revisions) and finally to the end of project stage (policy lessons and accountability). It can also be used for advocacy, as in the “business case” developed by WFP to illustrate the potential worth of school feeding (Regnault de la Mothe and Molinas 2012).

58. **COST-BENEFIT ANALYSIS (CBA) requirements.** In terms of data needs, CBA is the most demanding of all the tools discussed here. As all efficiency analysis, it can only be applied if the relevant results from effectiveness analysis are available, and this may be the major part of the challenge. We need to know how much time beneficiaries have put in to secure the benefits, what the impact of emergency feeding is on beneficiary mortality and morbidity throughout their further life, or what the impact of school feeding is on learning outcomes, and from there on increased productivity and ultimately income earning power. Once these effects have been estimated, CBA can be used to translate them in monetary terms and aggregate them into a measure of project worth, such as NPV, BCR, or IRR. This will require information on national parameters such as a discount rate or a shadow wage rate. When CBA is applied to sub-projects, benefits typically consist of foreign exchange or fuel savings that can be more readily measured in money terms, and data requirements are moderate. But when CBA is applied at outcome and impact level, data requirements become extremely heavy. Advanced economic skills are required to perform CBA.

²⁶ See also paragraph 61 below on this.

Table 3 Characteristics of selected BMZ tools

		CBA	CEA	CUA	MADM	unit cost benchmarking	other benchmarks	follow the money	specific efficiency questions
features	analytical rigour (BMZ)	2	2	2	2	1	1	1	0
	location in logical framework	activities; impact	all levels	impact	impact	inputs, activities	inputs, activities	activities	all levels
	PS terminology	efficiency	efficiency	efficiency	efficiency	efficiency	effectiveness	efficiency	both
	intervention cycle	throughout	throughout	throughout	throughout	mid-term and ex post	mid-term and ex post	mid-term and ex post	mid-term and ex post
requirements	data needs	moderate to very heavy	heavy to very heavy	very heavy	moderate	light	moderate	light	very light
	economic skills required	advanced	advanced	advanced	none	limited	none	limited	none
applicability	WFP interventions	all	single objective	nutrition and/or health	all	all	all	all	all
	individual interventions	✓	✓	✓	✓	✓	✓	✓	✓
	country portfolios				✓	✓	✓	✓	✓
	role OE evaluators	verify	verify	verify	verify	interpret	interpret	do	do

59. **COST-BENEFIT ANALYSIS (CBA) applicability.** CBA can in principle be used fairly widely. Almost all WFP interventions have an important nutrition component that at the level of impact is expressed by indicators such as lives saved, or QALYs. CBA involves putting a monetary value on the length and the quality of human life thus measured. As we indicated in section 5, this is controversial but not impossible. By construction the value of a life differs between countries, with higher values in economically more developed countries. This makes CBA unsuited for comparisons across countries²⁷. Cost-utility analysis using QALYs may be the most sophisticated alternative on offer, at least if interventions have health as their principal objective. When health is not the only major objective, CUA is not a valid substitute, and neither is CEA. CBA then remains the only advanced efficiency tool. CBA cannot be meaningfully applied to a country portfolio. This is not a problem, however, as the results of the CBA of individual interventions (EMOPs, PRROs, DEV,...) can be easily aggregated, provided the same methodology has been used in all the components. As for the role of evaluation teams contracted by OE to assess CPEs and OpEvs, it is not realistic to expect that they will be able to perform such analysis. What they may do, is verify whether appropriate efforts have been made by the CO to make use of the tool in selecting interventions and modalities, and where CBA has been applied, to verify whether the recommendations were followed and whether its major assumptions are still valid.

60. **COST-BENEFIT ANALYSIS (CBA) applicability (continued).** In the past alpha values were often calculated for food-in-kind operations at WFP. Alpha values are the local market value of a food-in-kind transfer, expressed in dollars, divided by the total dollar cost to WFP. In the 1980s WFP promoted the calculation of alpha values at the appraisal stage. It helped the organisation “doing the right things”, in the sense of either deciding to buy locally rather than import food, or by staying out altogether, for instance when the cost of bringing in food aid from the outside is very high, and another donor agency can provide cash support to beneficiaries. One possible interpretation of the alpha ratio is to consider it a BCR. The numerator can indeed be regarded as a measure of the benefit to beneficiaries. For it expresses what, in the absence of the food aid package, they would have paid for the same food, or what they would obtain if they sold the food²⁸. And the denominator expresses the costs of the food package. As both numerator and denominator are expressed in the same dollar units, the ratio can indeed be interpreted as a simple BCR at output level. It is however problematic to interpret alpha values in this fashion, because current market prices may be inappropriate for estimating benefits. First this reasoning assumes current food prices are not affected by the intervention. But this may be an untenable assumption. If we are for instance comparing a food-in-kind with a cash transfers for instance, we must take into considerations that in a poorly integrated and tiny market, a cash transfer programme may

²⁷ It would make identical WFP interventions more efficient if implemented in recipient countries with a higher living standard. If equity considerations are brought in, the opposite should be the case. If equity is not considered, the results should be neutral.

²⁸ This assumes that there is no effect on food prices. But this may be an untenable assumption. If we are for instance comparing a food-in-kind with a cash transfers for instance, we must take into considerations that in a poorly integrated and tiny market, a cash transfer programme may cause a considerable jump in the price of food.

cause a considerable jump in the price of food. Second, even the changed market price is only a good approximation of the value of a commodity for people who in the absence of the intervention have purchasing power and access to the market. In WFP emergency interventions, these conditions are typically not satisfied, and the benefit of providing one unit of food aid is higher than its market value. In the case of WFP development interventions, market prices are a poor guide for another reason: the underlying merit good argument. We will come back to alpha values later on, because they are useful in benchmarking the costs of food assistance transfers, that is in helping the organisation keeping prices under control, i.e. “doing things right”.

61. **COST-BENEFIT ANALYSIS (CBA) example.** In Table 4 a cost-benefit analysis is presented of a food fortification project. It is loosely based on the Timor Vita case study in Timor Leste, but missing values have been invented, and some parameters have been changed to make the case study more didactic. It is thus about an imaginary project in an imaginary country, with subsistence agriculture as the dominant form and poorly developed markets. We suppose that the information in the table was used to justify the support of WFP and the government for the local production of the fortified food, and that this information is made available to the evaluators.

**Table 4 Economic CBA of a food fortification factory (ex ante)
(thousands of 2009 \$)**

YEAR	0	1	2	2 to 12	13
INVESTMENT					
Land	24.4				
Buildings	450.0	300.0			
Machinery	200.0	54.2			
residual value					-34.0
TOTAL	674.4	354.2	0.0	0.0	-34.0
RECURRENT COSTS					
skilled labour	56.0	56.0	112.0	112.0	
unskilled labour	12.0	24.0	36.0	36.0	
commodities (local)			1248.0	1248.0	
commodities (import)			412.8	412.8	
Packaging			3.6	3.6	
Electricity			3.1	3.1	
Water			6.0	6.0	
Maintenance			3.4	3.4	
laboratory tests			24.0	24.0	
TOTAL	68.0	80.0	1849.0	1849.0	0.0
BENEFITS					
import substitution			2016.0	2016.0	
NET BENEFITS	-742.4	-434.2	167.0	167.0	34.0

62. The NPV and IRR of the project are provided in Table 5 for discount rates (dr) ranging from 0% to 10% . The calculation served to give the project the go-ahead. It is based on the hypothesis that soya, maize and sugar can be bought from the local market, even if some nutrients are imported.

Table 5 Profitability (ex ante)

Dr ^a	NPV
0%	527.6
2%	323.1
4%	157.6
6%	23.2
8%	-86.3
10%	-175.8
6%	IRR

Note a: discount rate

63. At a discount rate of 8%, used by the donor community for investment projects in this country, the NPV is negative. The CO was aware of this but argued that this was because the project had to go through an infancy stage, and that in the future (for instance when the project would be renewed after the initial phase of 13 years) it would show a healthy profitability. However, when the evaluation team visits the country, in year 5 of the project, it appears that all agricultural ingredients are imported. There is no prospect that these ingredients will be bought locally any time soon, and this for a number of reasons that were spelled out in a feasibility study that was commissioned at appraisal but whose recommendations were then ignored. The CO acknowledges the problem but remains convinced that the project is worth doing. It sees an additional economic advantage in the local value added that is created by the project, some of which is passed on to workers in the form of wages, some to the government in the form of taxes, and some in to the local entrepreneur in the form of profits. There are also political reasons: the government is very keen to support a local industry that is very slow to emerge. The evaluators rerun the calculations on the basis of the actual situation. The results are presented in Table 6 and Table 7.

Table 6 Economic CBA of a food fortification factory (ex post)

YEAR	0	1	2	3 to 12	13
INVESTMENT					
Land	24.4				
buildings	450.0	300.0			
machinery	200.0	54.2			
residual value					-34.0
TOTAL	674.4	354.2	0.0	0.0	-34.0
RECURRENT COSTS					
skilled labour	56.0	56.0	112.0	112.0	
unskilled labour	12.0	24.0	36.0	36.0	
commodities (local)			0.0	0.0	
commodities (import)			2210.4	2210.4	
packaging			3.6	3.6	
electricity			3.1	3.1	
Water			6.0	6.0	
maintenance			3.4	3.4	
laboratory tests			24.0	24.0	
TOTAL	68.0	80.0	2398.6	2398.6	0.0
BENEFITS					
import substitution			2016.0	2016.0	
NET BENEFITS	-742.4	-434.2	-382.6	-382.6	34.0

Table 7 Profitability (ex post)

dr	NPV
0%	-4,968.4
2%	-4,422.0
4%	-3,963.8
6%	-3,576.9
8%	-3,248.1
10%	-2,966.7
	IRR

64. Table 7 suggests that the project does not create any economic value added, compared with not doing the project and instead using the same resources (including aid) in their best alternative use. On the contrary, the project generates a considerable negative value added for the economy. What this means is that any benefits the owner, workers or the government derive from the project are compensated by much higher losses elsewhere in the economy. Doing sensitivity analysis, they also reveal that the project can just break even at a 8% discount rate (NPV=0) if the price at which the finished product can be bought from outside the country would be 28% higher. But a brief check with people knowledgeable in the sector leads to the conclusion that this is highly unlikely²⁹. The conclusion is that there is no

²⁹ We could run many other sensitivity tests on the basis of this example, but then we would need to explain the project and the model assumptions in much more detail. We could also perform CBA from

justification for supporting the project until the moment in time local farmers are able to supply raw materials to it. A project to support local farmers in this regard may therefore be worth studying.

65. **COST-EFFECTIVENESS ANALYSIS (CEA) features.** This tool is rated level 2 by the BMZ study. It is closely related to CBA from which it can be derived as a special case. Compared to CBA it has the following drawbacks. (1) It does not tell us whether an intervention is worth doing, only how it ranks in comparison with other interventions. (2) When interventions are compared that have different benefit scales, CEA may not be strictly valid and may have to be complemented by appropriate CBA considerations³⁰. (3) It must be possible to express all benefits in the one selected non-monetary indicator, which is a limiting factor when doing efficiency analysis of projects with multiple objectives or when comparing projects with different objectives. CEA is often treated as the poor cousin of CBA in the theoretical literature. But CBA is in practice very difficult to apply and has never fully caught on in the field of aid. In contrast CEA is highly appreciated by practitioners³¹. This is especially the case in non-productive sectors, such as health, education, or social safety nets, where benefits are difficult to monetize. Whereas the theory of CBA has not seen any significant additions in the last 40 years, several recent efforts have been made to refine CEA (Ravallion 1999; Dhaliwal et al. 2012). CEA can be applied at all levels of the logical framework: cost per MT of food delivered to a central warehouse (activities), cost per nutritional value unit delivered to beneficiaries (output), cost per nutritional improvement achieved (outcome), and cost per life saved (impact). Note that technically the calculation of unit costs can be considered an application of CEA. It also concerns a ratio of costs and some relevant unit of activity or output. The author of the BMZ study nevertheless treats them separately. We agree with his approach because the ratio in that case is just a tool to standardize costs and compare them with benchmarks drawn from a larger number of projects that are not part of the same decision situation. See further in the text. Just as CBA, CEA can be applied throughout the life cycle of an intervention, from ex ante (to assist in decision making), to mid-term (to help management) and finally end of project (policy lessons and accountability). It is less convincing for advocacy purposes, because of the first drawback listed before in this paragraph.

66. **COST-EFFECTIVENESS ANALYSIS (CEA) requirements.** CEA is less demanding than CBA in that benefits expressed in the units used in the underlying effectiveness analysis can be directly used and do not have to be translated in money terms. Apart from this, the data needs are the same. CEA also appears less demanding from a methodological perspective than CBA, but its intuitive appeal is somewhat deceptive: mistakes are easily made, for instance in interpreting the rankings obtained. Economic skills are required that in some cases are only slightly less demanding than in CBA.

the perspective of the private firm that is producing the product to highlight the difference between financial and economic CBA.

³⁰ This is a bit technical. We will later provide a numerical example relevant for WFP to illustrate what the problem is and how it may be overcome.

³¹ After comparing the two methods, the DFID guide on VfM in cash transfer programmes comes to the following conclusion: "For these reasons, a well-conducted cost-effectiveness analysis is always to be preferred to a poorly substantiated CBA" (Hodgson et al. 2011: 40).

67. **COST-EFFECTIVENESS ANALYSIS (CEA) applicability.** An attractive feature of CEA is that the analysis can already be applied to the activity and output level, without having to go all the way up to outcome or impact measurement. And the relevant statistics such as the cost per school meal ration or per MT of food delivered are generally available for evaluators. However, for reasons given before, output-level CEA can be highly misleading as an indication of higher level efficiency, and it can therefore be questioned whether it constitutes a meaningful measure of efficiency. Such output-level ratios are best used in another more restricted tool, unit cost benchmarking. A more convincing feature of CEA is that there is no need to put money values on difficult-to-monetize benefits. WFP has developed the “omega value” to select the appropriate modality of food assistance at outcome level: food-in-kind, vouchers or cash. See the example below. Note that this works fine as long as the emphasis is on nutrition. If interventions have multiple objectives, such as FFW, FFA or school feeding, then we will probably wish to look at the efficiency from these other perspectives as well. We could try to rank such projects by their cost per nutritional improvement, and then also by their cost per educational improvement, and by their cost per asset created. Sure enough, each of these ratios gives a relevant part of the answer to the efficiency question. But in many cases these partial ratios will provide different rankings of the projects, and we need a more comprehensive measure to overall efficiency. CBA would then be the first option³². Another possibility is to use MADM, a tool discussed below. We conclude that CEA is suitable for analysing WFP interventions from a single perspective such as nutrition. It is thus probably more appropriate for EMOP than for multiple-objective WFP interventions such as school feeding. For the same reason the role of CEA in evaluating country portfolios is problematic. A typical portfolio is composed of several interventions (i.e. the EMOP, PRRO, DEV, ...) that cannot be reduced to one single objective, and thus cannot be aggregated in one single overarching non-monetary metric. QALYs are more powerful in this regard, although less powerful than full monetisation or the use of MADM. And what is the role of OE evaluators? CEA analysis will only rarely be performed from scratch during OE evaluation missions. What evaluators can do is to verify CEA calculations that were performed at the appraisal stage. But this requires that appropriate data on implementation are generated in the WFP M&E system.

³² An hybrid between CEA and CBA can be imagined. If some limited non-nutrition benefits can be expressed in monetary terms, they can be subtracted from the costs in the numerator of the CER, with nutritional benefits captured in the denominator in a non-monetary metric. Such a procedure is valid, but must to be applied consistently in all the interventions being compared.

68. **COST-EFFECTIVENESS ANALYSIS (CEA) example.** Table 8 summarizes information on Egypt from a WFP study on cash, vouchers and food-in-kind modalities (Ryckembusch 2012a). The information draws on household surveys, and is claimed to measure nutritional outcome³³. Column (1) provides information in the form of a CER: the monthly cost (in local currency) of inducing the consumption of an additional daily ration of 1300 kcal/person under each of the three modalities. Column (2) gives the ranking of the three modalities in terms of CEA. Cash is the clear winner. The research also allowed to calculate not only the energetic but also the nutritional value of the rations that beneficiaries consume under the three modalities. Column (3) provide provides the nutritional value score (NVS) in each case. In column (4) the corresponding CER is presented: the cost in local currency per nutritional unit consumed. The ranking, presented in column (5), is now different: cash is the least attractive alternative, vouchers the most attractive, with food-in-kind in the middle position. Between the two possible approaches to measure nutritional outcomes, energy (kcal) and nutrients (NVS), WFP seems to favour the second. Vouchers are thus the preferred option. Note that the results are identical to those in Ryckembusch (2012B). However the calculations are here presented using cost-effectiveness ratios rather than the omega values he proposes³⁴.

³³ The logical framework only distinguishes between five levels: inputs, activities, output, outcome, and impact. Sometimes it would be useful to have more. In this particular instance, we would argue that consumption of kcal or nutrients might be best understood as “intermediate outcome indicators”, and that for a full outcome indicator we would have to measure a reduction in some measure or malnutrition, as we do in a subsequent example in table 10 below. For examples of possible “intermediary outcome indicators” for food-based interventions, see Grosch et al. (2009: 240-241).

³⁴ The omega value is the outcome of a division of two CERs. If for instance we wish to compare cash and in-kind transfer modalities in our example, the two CERs in column (4) can be divided as $214/197=1.08$. An identical result is obtained in Ryckembusch (2012b), but using a different route. The underlying economic logic suggests the following interpretation of the omega value: if the ratio exceeds unity the modality in the denominator (in this case food-in-kind) is the better option, if less than unity it is the modality in the numerator (cash). Using omega values thus leads to the same conclusion as a direct comparison of CERs, but lack the intuitive feel of CER rankings, and can only be compared two at the time. We therefore suggest to use CERs as illustrated in Table 8.

Table 8 Alternative outcome rankings of three food modalities

treatment	CER _{outcome, kcal} : cost in local currency of providing 1300 kcal/person/day	CER _{outcome, kcal} ranking	nutritional value score (NVS)	CER _{outcome, NVS} : cost in local currency of providing one unit of NVS	CER _{outcome, NVS} ranking
	(1)	(2)	(3)	(4)	(5)
cash	1317	1	6.15	214	3
voucher	1450	2	7.70	188	1
in-kind	1717	3	8.70	197	2

69. COs are from now on expected to consider the choice of modalities explicitly and do an analysis along the lines of Table 8 before deciding. This requires prior information that can come from running an experiment, from household surveys, or from extrapolating from other countries. The task of evaluators will increasingly be to check whether such a calculation has been performed, and whether the results are convincing, also in light of international evidence.

70. **COST-UTILITY ANALYSIS (CUA) features.** Rated at level 2 by the BMZ study, the same as CBA and CEA. In the academic literature there are diverging views on the pecking order. Some authors see CUA as the highest form of efficiency analysis. Others see it as just an application of CEA, which itself can be derived as a special case of CBA, and therefore place it lower in the pecking order. Even worse, some see it as poor efficiency analysis. The method for calculating DALYs, a specific version of QALYs, which was developed in support of the 1993 World Development Report (World Bank 1993), has for instance been strongly criticised on methodological grounds (Anand and Hanson 1997, Williams 1999). Note that QALYs or DALYs are situated at the impact level in the logical framework. For this reason, CUA can only be applied at this level. It can be used throughout the intervention cycle. At the ex ante stage, results from theoretical studies and empirical research can be used to generate predictions about what impact a certain intervention is expected to have if well implemented. During mid-term reviews, evaluators can correct for assumptions that have proved too optimistic or pessimistic or for implementation failures. During ex post evaluations actuals can be substituted for planned results.

71. **COST-UTILITY ANALYSIS (CUA) requirements.** Data needs are very heavy, but this is mostly for obtaining the underlying effectiveness indicators, and this is a job for medical experts and bio-statisticians. WHO now publishes annual data on the burden of disease which can be used as DALYs. As argued before, once a reliable value for the denominator has been obtained, the CUA calculation itself is not so demanding: on the cost

side the data are the same as for CEA at output or outcome levels. The economic skills required are nevertheless advanced, identical to those for CEA.

72. **COST-UTILITY ANALYSIS (CUA) applicability.** CUA is applicable to all WFP operations where impact can be expressed in QALYs. This means in practice all operations where the objective is nutrition, health, or a combination of both. This makes the method suitable to some interventions with multiple objective, such as MCHC. Note that QALYs look at the effect on the whole lifespan of a beneficiary. From this perspective, it will be more efficient to save the life of younger people than of elderly people. If the moral principle is that all human beings, independent of their life expectancy, have the same right to be helped, then it is better to use as a metric lives saved, which is an application of CEA. The role of CUA in evaluating country portfolios is limited. The same limitations invoked for CEA apply. Even if the unit of measurement is now broader than with CEA, nevertheless in most cases some of the interventions that compose a country portfolio will aim at objectives distinct from nutrition and health, rendering CUA at this aggregate level inappropriate. As to what can be expected of evaluators working for OE, the points made in connection with CEA also apply here. Evaluators can play a role in verifying existing calculations made at the appraisal stage. But when they find that no such study was performed, they have neither the qualifications, time or the resources to do CUA themselves.

73. **COST-UTILITY ANALYSIS (CUA) example.** We do not think it makes much sense to explain through a worked-out example how to operate a tool that it is beyond the reach of OE evaluators. But if consultants evaluating CPEs and OpEvs need not have a hands-on expertise on estimating QALYs or DALYs, it is useful that they know that such information is available in the literature in a way that enables broad comparisons across policy instruments and countries. An example from the World Development Report 1993 is reproduced in Table 9

Table 9 Cost-effectiveness of selected nutritional interventions

Intervention	Target group	Approximate cost (dollars)	
		Per death averted	Per DALY saved
Iron supplementation	Pregnant women	800	12
Iron fortification	Entire population	2000	4
Iodine supplementation	Women of reproductive age	1250	19
Iodine supplementation	Entire population	4650	37
Iodization of salt or water	Entire population	1008	8
Vitamin A supplementation	Children under age 5	50	1
Vitamin A fortification	Entire population	154	4
Food supplementation	Children under age 5	1942	63
Food supplementation ^a	Pregnant women	733	24

Notes

a : Deaths averted and DALYs saved are for fetal deaths.

Source: World Bank (1993: 82)

74. In the absence of robust information on the impact efficiency of a particular intervention, such comparative information can help establish how much credibility claims made on the efficiency of an intervention deserve.

75. **INTUITIVE MULTI-ATTRIBUTE DECISION MODEL (MADM) features.** In the academic literature MADM is presented as an enrichment and also a correction of CBA, CUA and CEA. It is a way to bring dimensions to bear that are important to decision making but have not or not sufficiently been taken into account in the more quantitative economic analysis. Suppose that a series of WFP projects are being evaluated, and apart from efficiency impact analysis, scores have been calculated for the way projects have an effect on women empowerment, accountability to beneficiaries, political fragility, and environmental sustainability. If these contributions to different objectives are rated on the same scale, then a weighted summary score can be computed to rank the projects. The BMZ study rates MADM as level 2 analysis, even though the method is fairly straightforward and easy to apply, once the underlying information is available. Care must be taken to select the objectives retained for MADM. In particular, to avoid double counting, they must be distinct. Efficiency measures at output, outcome and impact level are a good case of the contrary. The following formulae show how the CER at a lower level is an integral part of the calculation of a CER at higher levels.

$$CER_{output} = \frac{cost}{output}$$

$$CER_{outcome} = \frac{cost}{outcome} = \frac{cost}{output} \frac{output}{outcome}$$

$$CER_{impact} = \frac{cost}{impact} = \frac{cost}{outcome} \frac{outcome}{impact}$$

76. It would thus be double counting to include these different CERs alongside each other in MADM, as is sometimes suggested in WFP documents on modality selection. Instead, only the one that is highest in the logical framework should be used in decision making.

77. This tool has been developed mainly as a decision tool, and therefore its first and most important level of application is during ex ante evaluation. But it can be applied at mid-term and ex post to verify whether the right choice was effectively made and whether it is still valid.

78. **INTUITIVE MULTI-ATTRIBUTE DECISION MODEL (MADM) requirements.** Once the underlying information from the analysis of economic efficiency and from the assessment of the other dimensions are available, the additional data requirements to do MADM are light. Some analysis of data may be required to translate the scores to the same scale. But once this is done MADM is essentially a weighting system involving the judgement of the decision makers without relying on external data. No specific economic skills are involved.

79. **INTUITIVE MULTI-ATTRIBUTE DECISION MODEL (MADM) applicability.** This tool can be used to reconcile the information from different efficiency and other metrics that are not expressed in the same units. In this sense it is suitable to WFP interventions with multiple objectives where partial indicators are available, for instance on nutrition, education and asset creation. In this fashion, this tool can also be used to assign an overall score to country portfolios. The question here is not about technical feasibility, but rather of desirability. Does OE wish to get overall scores for all interventions and country portfolios? If

so, it should decide on the dimensions to be scored by evaluators, and also impose the scale and the weights.

80. **INTUITIVE MULTI-ATTRIBUTE DECISION MODEL (MADM) example.** We give an example in Table 10 of how MADM can be used for assessing the choice of modalities. It is inspired by Ryckembusch (2012b, slide 52) and draws on his example for selecting objectives and even for assigning scores, but corrects for double counting. The denominator of the CER used here measures the change, attributed to the intervention, in the percentage of beneficiaries whose food consumption score is rated as acceptable. The exercise suggests that in this particular case vouchers is the best choice, then food-in-kind, and cash is the least attractive modality. This obviously depends on the scores assigned, as well as on the objectives selected and the weight assigned to each objective. Suppose that the list of objectives and their weights is decided in Rome, whereas scores are attributed by the CO during appraisal. Evaluators could be asked to re-score the table. If they come to a different ex post ranking than the ex ante ranking by the CO, it may be because there have been unforeseen changes in external circumstances beyond the control of the CO, or it may be that the CO has not done its job very well. In both cases, it would be a significant finding.

Table 10 Decision tree for selecting modalities

	Weight	In kind	Vouchers	Cash
	Score 1- 10	Score 1-5		
Timeliness	3	4	3	5
CER _{outcome, %FCS}	8	4	4	2
Beneficiary preferences	1	2	4	2
Gender and protection	3	2	4	2
Impact on local market	2	2	4	3
Overall score		56	65	45

81. **BENCHMARKING UNIT COSTS features.** Rated at level 1 in the BMZ study. Unit costs are calculated as a ratio of costs to some relevant unit (food, staff, beneficiary,...) and algebraically resemble CERs. But the emphasis is not on comparing alternative strategies but on keeping the costs down of the alternative that has been chosen. This is not always understood, and the author of the BMZ study warns that such cost analysis does not inform us about the efficiency of development interventions. Some of the experts he interviewed even did not see a place for it in the efficiency toolbox³⁵. However, if it is understood that this

³⁵ “These experts saw too great a risk of explicit or implicit conclusions based on insufficient facts. Indeed, in our own review of evaluation reports that contained unit costs and other partial efficiency benchmarks, observations regarding overall efficiency of interventions were sometimes based on a fragment of the necessary information only.” (Palenberg 2011: 44). To illustrate this point in the health sector: the cost per vaccine stocked in health clinics refrigerators (activities) or cost per patient

is not a tool for “doing the right things” but for “doing things right”, this criticism is unwarranted. In the logical framework, this tool is situated at the level of inputs (food imported or secured locally, staff,...) and activities (transport, warehousing, food fortification, delivery...), up to the level of output, which at WFP is mostly food assistance delivered. The calculating of unit costs and their comparison with some external benchmark can be usefully done during mid-term evaluation. Care must be taken to establish appropriate benchmarks.

82. **BENCHMARKING UNIT COSTS requirements.** Data requirements are fairly limited provided a good system of accounting and monitoring is in place. If that is not the case, getting good data may be difficult and even impossible. The economic skills required are limited and do not require specific training.

83. **BENCHMARKING UNIT COSTS applicability.** Given the complex logistics involved in providing food assistance in difficult circumstances and inaccessible places, managing costs is a major concern at WFP. Unit benchmarking therefore has an important role to play. It can be calculated for all types of individual interventions as well as country portfolios as a whole (for instance with regard to staffing or infrastructure). Evaluators should be able to interpret unit costs against benchmarks provided by HQ. At WFP international corporate comparisons are made. Two issues must be addressed in this regard. First, these benchmarks should be sufficiently disaggregated. For instance, data on the cost per MT of food delivered, irrespective of modality, is not very easy to interpret. A cost study has been recently completed at WFP HQ that addresses this issue in the field of school feeding WFP (2012b)³⁶. This type of analysis should be done in other fields as well. Second, only part of the global corporate variation in unit costs can be attributed to the difference in efficiency of individual interventions. Other drivers of cost differences are project age, project size, country political situation, and geography, and should be statistically removed³⁷. A useful metric is an algebraic variation on the alpha value: the cost-transfer ratio (CTR), the ratio of administrative costs to the value of the transfer. This is applicable to food-in-kind, but also to cash and vouchers.

84. **BENCHMARKING UNIT COSTS example.** We draw on the cost analysis for school feeding (WFP 2012b) just mentioned. Suppose that this analysis has been pushed to a higher level by correcting for the exogenous factors already identified. This leads to the result for a given (again hypothetical) country in Table 11. We first look at the top part of the table. Columns (1) and (2) give the mean average standardised cost benchmarks for all WFP school-meal-only interventions. Columns (3) and (4) are a corrected set of benchmarks which filter out the influence of a number of exogenous variables, such as the geography of

seen at health clinics (output) are useful tools for keeping costs under control, but are a very poor substitute for the cost per QALY when it comes to establishing the efficiency of a medical treatment.

³⁶ A distinction is made between "meals only", "meals + take-home rations", "take-home rations only", and "biscuits only".

³⁷ See Collier et al. (1997) for an illustration of how statistical regression analysis can be used to eliminate such unwanted influences in the case of establishing good national performance in the use of budget support. The authors of the cost study on school meals are aware of the problem (WFP 2012b slide 44), but we have not seen the final results of their efforts.

the country, its distance from major food supplying countries, the fact that this is a pilot project or not, etcetera³⁸. This leads to a considerable increase in the overall cost benchmark for the country in question, from \$40 to \$63. The total cost in the country is however \$70, much higher still. The breakdown along the usual lines of WFP cost accounting allows to identify the major cost driver: ODOC. Equipped with this information, the evaluation team can start discussing with the CO to understand what is the reason for this higher than average cost, and whether it is justified or not. The lower part of the table calculates some other unit costs: The alpha value gives the value transferred to beneficiaries per \$ of total cost. As indicated above, we propose to use an algebraic transformation of the alpha value, the Cost-Transfer Ratio (CTR), which expresses the total cost of transferring 1\$ in value to the beneficiary, which can be calculated as $((1/\alpha) - 1)$. As can be seen in Table 11, the CTR at 0.67 is problematically high in this country. The evaluators will have to study this result in more depth.

Table 11 Benchmarking of school-meal-only costs (standardised annual costs (\$) per recipient)

	WFP benchmark		WFP benchmark		Actuals	
	Uncorrected		corrected ^b		country	
	(1)	(2)	(3)	(4)	(5)	(6)
total cost	40.0	100%	63.0	100%	70.0	100%
commodity	19.2	48%	22.5	36%	23.8	34%
sea transport	1.6	4%	4.7	7%	3.7	5%
LTSH	8.0	20%	17.3	27%	19.4	28%
ODOC	2.4	6%	6.1	10%	13.7	20%
Support	8.8	22%	12.4	20%	9.4	13%
market value ^a	33.0		49.6		42.0	
alpha	0.83		0.79		0.60	
1/alpha	1.21		1.27		1.67	
CTR	0.21		0.27		0.67	

notes

a: food plus imputed costs of home cooking

b: exogenous cost drivers removed

85. **BENCHMARKING PARTIAL INDICATORS OTHER THAN UNIT COSTS features.** Is similar to benchmarking of unit costs, and explained under the same heading in the BMZ study, even if it appears as a separate tool later on, as reproduced in our Table 2 above. Only a few examples are given, among them time until a loan has become effective. This tool is rated level 1 in the BMZ study. It is located in the lower levels of the logical framework (inputs and activities). It does gauge effectiveness rather than efficiency (PS terminology) ,and in this sense it is somewhat strange that it features in the BMZ toolbox. But this is not an issue here. As explained above, effectiveness must be verified before efficiency can be analysed, and therefore some effectiveness analysis is almost always required before

³⁸ In reality such calculations have not yet been made available by WFP Rome. We assume that this will be the case in some point in the near future.

evaluators can pronounce on efficiency. Can be usefully performed during mid-term and ex post evaluations.

86. **BENCHMARKING PARTIAL INDICATORS OTHER THAN UNIT COSTS requirements.** Data needs are limited, again provided that some relevant information can be generated from the M&E system. If this condition is not fulfilled, it will often be impossible to construct these indicators as an outside evaluator with limited time and resources. No economic skills are required to use this tool.

87. **BENCHMARKING PARTIAL INDICATORS OTHER THAN UNIT COSTS applicability.** Is applicable at WFP for all types of interventions and for evaluating country portfolios. Some useful partial indicators are already generated as part of the annual CO performance plans, such as share of funding needs secured, share of funding unspent at the end of the intervention, share of projects starting within targeted time, share of operations with full ration size distributed during the reporting period. Evaluators should be able to interpret these indicators against corporate benchmarks. As we are dealing with partial indicators, trade-offs with other results should be considered. For instance, timeliness may be achieved but at a very high delivery cost, and for that reason also at the cost of a reduced distribution cycle. It is a question of striking the right balance.

88. **FOLLOW THE MONEY features.** a pragmatic approach, not rigorously described in the literature. Rated level 1 in the BMZ study. Involves to the detailed recording of all the steps in the implementation of an intervention, and the search at each step for cost minimization measures or yield maximization measures. Applicable at the activities level of the logical framework. In the PS terminology it belongs to efficiency analysis. Can be a useful tool during mid-term evaluations.

89. **FOLLOW THE MONEY requirements.** When this tool is used, it will probably involve questioning of CO staff, government and implementing NGOs. The need for the evaluator to collect independent corroborating data is presumably limited. No advanced economic skills are required. An understanding of the complex logistics of food assistance is probably important.

90. **FOLLOW THE MONEY applicability.** Applicable at the level of individual interventions and country portfolios.

91. **FOLLOW THE MONEY example.** A natural application for this tool is the logistics part of WFP interventions. During evaluation, this tool can be used to verify whether WFP procedures have been followed. It can be deployed to understand why certain unit cost or other indicators deviate from corporate benchmarks. In the previous example, developed in ¶ 84 above, the suggestion at the end was that evaluators look into the high ODOC of the school feeding programme. Another interesting application is cases of national execution. One way to follow the money is to start at the end: how much of the food assistance that is mobilised arrives at its destination, and where are the breaks in the pipeline?

92. **SPECIFIC EFFICIENCY QUESTIONS features.** This is considered a “level o” tool of efficiency by the author of the BMZ study. It is best understood in the context of theory-based evaluation. Following this methodological approach, the underlying logical theory of an intervention is made explicit, so that the evaluation can zoom in on particularly sensitive sections of the logical chain. If for instance a particular operation has not been implemented as foreseen, the chain of cause-to-effect of the underlying theory may have been violated at one particular point, and the hoped for results will not be achieved. Depending on the particular question, it can be either focusing on effectiveness issues or efficiency issues, or both. A typical tool for mid-term evaluations.

93. **SPECIFIC EFFICIENCY QUESTIONS requirements.** Data requirements will vary but will in general be very light. No economic skills are required.

94. **SPECIFIC EFFICIENCY QUESTIONS applicability.** In the course of field visits evaluation teams often come across unexpected issues that suggest a break in the underlying theory. In the present practice at OE it is left to evaluators to decide what they want to look for. That freedom must be preserved. But it may be useful that in the ToR of an evaluation some critical sequences in the logical chain are identified where either insight from similar interventions or the history of the intervention itself suggest that there may have been an occurrence that has either a positive or negative effect on efficiency. The evaluation team can then be asked to look into this occurrence and give its expert opinion of how it may affect efficiency. The tool is applicable to both individual interventions and country portfolios.

95. **SPECIFIC EFFICIENCY QUESTIONS example.** It is less relevant to propose a case study here. Mostly, there is no calculation involved or standard procedure to be followed.

10. Summary and recommendations

96. The BMZ toolbox (Palenberg 2011) does not fully address what in the “three Es” terminology is called efficiency. First, in contrast to efficiency in the “three Es” terminology, efficiency in the terminology that underscores the BMZ study - and that we have called Preferred Standard terminology - is not about effectiveness (even if it builds further on effectiveness results). Second, most of the BMZ tools aim higher than the output level in the logical framework. This has inevitably led to some degree of mismatch between the content of the BMZ toolbox and the expectations of OE. We refrain from recommending that OE and WFP drop the “three Es” terminology altogether, although we feel that there are very sound intellectual reasons for doing so, because we understand that this is not realistic. We have therefore tried to find a way of applying the insights from an economic efficiency literature that is based on PS terminology to WFP procedures and terminology without forcing a radical vocabulary change. In order not to leave terms undefined, we often had to juxtapose the two types of terminology. It will be important for OE to acknowledge this difference in terminology in the future and to provide more clarity. At present there is a considerable amount of terminological confusion in WFP documents.

97. We have used the theoretical literature to frame the efficiency tools that are available. We then reviewed the 15 BMZ tools, and assessed their usefulness for WFP. In the end, 8 tools were judged to be less immediately useful for use in ex post evaluations by OE. The remaining 8 tools were reviewed in more detail. For each tool we discussed in turn their key characteristics, in particular where they are situated in the logical framework, what they would be called in the PS terminology, and how they can be applied in the intervention cycle. Next we looked at the requirements in terms of data and economic skills to operate them. Then we addressed their applicability to ex post evaluations commissioned by OE. We did this by contrasting operations with single and those with multiple objectives, by distinguishing between OpEv and CPE, and also by how active and with how much precision evaluators can be expected to handle these tools. And where the tools allowed for this, we provided numerical (but hypothetical) WFP examples.

98. OE has requested that we to propose the three tools that are most suitable for WFP at this stage. After an initial proposal from our side and further discussion with OE and Stephen Lister of Mokoro, we suggest that these are, in order of priority: Unit Cost Benchmarking, Specific Efficiency Questions, and Cost-Effectiveness Analysis (CEA):

- (1) Unit Cost Benchmarking is at present the most widely used (and de facto often the only) efficiency tool in OE evaluations, as can be seen in our brief review in Annex B. But it is widely misused, and the improvement of this tool and its interpretation should be high on the priority list. It is both desirable and feasible. We suggest to reintroduce a variant of the alpha value that was popular in the 1980s, in the form of the Cost-Transfer Ratio (CTR). We also suggest that there is some work to be done by WFP Rome to refine the benchmarks per modality. The corporate average cost per metric ton of food, probably the benchmark at present most readily available to evaluators, is really not very meaningful. We have argued in which way benchmarking could best be refined and given an example of how evaluators could then use this improved tool.
- (2) When OE staff draw up ToR and subsequently negotiate evaluation matrices with evaluation teams, they de facto propose (or approve) specific evaluation questions on

efficiency. It would be good to have a set of semi-standard questions that avoid conceptual confusion, but also get at issues of concern in ways that are likely to yield credible answers, and ideally answers that are somewhat comparable across evaluations.

- (3) Ex ante CEA is now advocated by WFP to select among food assistance modalities under the new food assistance strategy. It is quite natural for ex post evaluations to look into these same decisions with the hindsight of time. There will be some interventions that have to be evaluated that were set up before the new strategy was launched, but as time goes by, it should be possible to look into this CEA for most operations. It is feasible in the sense that evaluators do not have to start from scratch. They are rather expected to verify the calculations that were performed ex ante. One disadvantage of this tool is that it can only look at one objective at a time, and in the way it is promoted by WFP Rome for modality selection, this objective is nutritional improvement. There are however other objectives that may be relevant, such as learning outcomes, access to health, or asset creation. For this purpose, appropriate CEA measures may have to be developed in the future.

Bibliography

- Anand, S., Hanson, K. (1997), “Disability-adjusted Life Years: a Critical Review”, *Journal of Health Economics*, Vol. 6, pp. 685-702.
- Audsley, B., Halme R., Balzer, N. (2009), “Comparing Cash and Food Transfers: a Cost-Benefit Analysis from Rural Malawi”, in : Omamo et al. , pp. 89-102.
- Bailey, S., Hedlund, K. (2012), *The Impact of Cash Transfers on Nutrition in Emergency and Transitional Contexts - A Review of Evidence*, London, Humanitarian Policy Group, Overseas Development Institute.
- Banerjee, A.V., Duflo, E. (2011), *Poor Economics. A Radical Rethinking of the Way to Fight Global Poverty*, Public Affairs Books.
- Barnett, C., Barr, J., Christie, A., Duff, B., Hext S. (2010), *Measuring the Impact and Value for Money of Governance & Conflict Programmes*, ITAD, London.
- Barrett, C., Maxwell. 2005, D. (2005), *Food aid after fifty years*, Routledge, Taylor and Francis Group.
- Boardman, A.E., Greenberg, D.H., Vining, A.R., Weimer, D.L. (2011), *Cost-Benefit Analysis – Concepts and Practice*, fourth revised edition, Pearson.
- Cropper, M.L., Sahin, S. (2009), Valuing Mortality and Morbidity in the Context of Disaster Risks, *World Bank Policy Research Working Paper 4832*, Washington D.C., World Bank.
- Collier, P., Guillaumont, P., Guillaumont, S., Gunning, J.W. (1997), “Redesigning Conditionality”, *World Development* , Vol. 25, pp. 1399-1407.
- Dhaliwal, I., Duflo, E., Glennerster, R., Tulloch, C. (2012), *Comparative Cost-effectiveness Analysis to Inform Policy in Developing Countries*, Abdul Latif Jameel Poverty Action Lab (J-PAL), MIT.
- DFID (2011), *DFID’s Approach to Value for Money (VfM)*, July 2011.
- Drummond, M.F., O’ Brien, B., Stoddart, G.L., Torrance, G.W. (1997), *Methods for the Economic Evaluation of Health Care Projects*, second revised edition, Oxford University Press.
- Grosch, M., del Ninno, C., Tesliuc, E., Ouerghi, A. (2008), *For Protection and Promotion – The Design and Implementation of Effective Safety Nets*, Washington D.C., World Bank.
- Hall, A., Blankson, B., Shoham. J. (2011), *The Impact and Effectiveness of Emergency Nutrition And Nutrition Related Interventions: A Review Of Published Evidence 2004-2010*, Oxford UK, Emergency Nutrition Network.
- Harvey, P., Proudlock, K., Clay, E., Riley, B., Jaspars, S. (2010), *Food aid and food assistance in emergency and transitional contexts: a review of current thinking*, HPG – ODI, London.

- Headay, D. (2012), “Developmental Drivers of Nutritional Change: A Cross-Country Analysis”, *World Development*, Vol. 42, pp. 76-88.
- Hidrobo, M., Hoddinott, J., Peterman, A., Margolies, A., Moreira, V. (2012), *Cash, Food, or Vouchers? Evidence From a Randomized Experiment in Northern Ecuador*, IFPRI, Washington D.C.
- Hodgson, A., White, Ph., Greenslade, M. (2011) *Guidance for DFID country offices on measuring and maximising value for money in cash transfer programmes - Toolkit and explanatory text*, DFID, London
- Lentz, E.C., Barrett, C.B., Gomez, M.I., Maxwell, D.G. (2013a), “On the Choice and Impacts of Innovative International Food Assistance Instruments”, *World Development*, in press.
- Lentz, E.C., Passarelli, S., Barrett, C.B. (2013b), “The Timeliness and Cost-effectiveness of the Local and Regional Procurement of Food Aid”, *World Development*, in press.
- Levin, H.M., McEwan, P.J. (2001), *Cost-Effectiveness Analysis*, second edition, Sage Publications.
- Little, I. M. D., Mirrlees, J. A. (1974), *Project Appraisal and Planning for Developing Countries*, New York, Basic Books
- Michiels, J. (2012), *Literature Review on Efficiency and Value for Money*, unpublished internal WFP document, Rome.
- Omamo, S., Gentilini, U. Sandstrom, S. (eds.) (2009), *Revolution: From Food Aid to Food Assistance - Innovations in Overcoming Hunger*, Rome, WFP.
- Rossi, P. H., Lipsey, M. W., Freeman, H. E. (2004), *Evaluation : a Systematic Approach*, Thousand Oaks CA, Sage.
- Sandström, S., Tchatchua, L. (2009), “Do Cash Transfers Improve Food Security in Emergencies?”, in: Omamo et al. , pp. 75-88.
- Ravallion, M. (1999), “Appraising Workfare Programs”, *World Bank Research Observer*, Vol. 14(1): 31-48.
- Regnault de la Mothe, M., Molinas, L. (2012) *The School Feeding Investment Case*, Rome, WFP.
- Ryckembusch, D. (2012a) *Egypt Cost Efficiency Effectiveness Voucher Transfer Programming*, 1st draft, Rome, WFP.
- Ryckembusch, D. (2012b) *Transfer Modality Selection*, powerpoint, Rome, WFP.
- Ryckembusch, D., Frega, R., Guilherme Silva, M., Gentilini, U., Sanogo, I., Grede, N., Brown, L. (2013), “Enhancing Nutrition: A New Tool for Ex-Ante Comparison of Commodity-Based Vouchers and Food Transfers”, *World Development*, in press.
- Palenberg, M. (2011), *Tools and Methods for Evaluating the Efficiency of Development Interventions*, Evaluation Working Papers, Bonn, Bundesministerium Für Wirtschaftliche Zusammenarbeit und Entwicklung.

Efficiency Analysis for WFP – Definitions, Concepts and Methodology

Savedoff, W., Levine, R., Birdsall, N. (co-chairs) (2006), *When Will We Ever Learn? Improving Lives through Impact Evaluation - Report of the Evaluation Gap Working Group*, Center for Global Development, Washington D.C.

WFP (2012a), *Draft Concept Note for Discussion - Inputs to the current discussion on Middle Income Countries*, internal document, Rome.

WFP (2012b), *School Feeding Cost Analysis – Final Report*, Powerpoint presentation, Rome.

WFP (2012c), *VfM/E&E brainstorming discussion*, Powerpoint presentation, Rome

White, Ph. (2006), *Cost comparison of cash, food and agricultural input transfer schemes in Malawi and Zambia - Summary of conclusions*, unpublished paper commissioned by DFID.

Williams, A. (1999), “Calculating the Global Burden of Disease: Time for a Strategic Reappraisal”, *Health Economics* 8: 1-8.

World Bank (1993), *World Development Report 1993 – Investing in Health*, Washington D.C.

Acronyms

BCR	Benefit-Cost Ratio
BMI	Body Mass Index
BMZ	Federal Ministry for Economic Cooperation and Development (Germany)
CBA	Cost-Benefit Analysis
CEA	Cost-Effectiveness Analysis
CER	cost-effectiveness ratio
CTR	Cost-transfer ratio
CO	Country Office
COMPAS	Commodity Movement, Processing and Analysis System
CPE	Country Portfolio Evaluation
CUA	cost-utility analysis
DAC	Development Assistance Committee (of the OECD)
DALY	Disability adjusted life year (lost)
DFID	Department for International Development (UK)
dr	discount rate
DRC	Domestic Resource Cost
DSC	Direct Support Cost
EMOP	Emergency Operation
FAO	Food and Agriculture Organization
FBF	Fortified Blended Food
GDP	Gross Domestic Product
IFAD	International Fund for Agricultural Development
ISC	Indirect Support Costs
LTSH	landside transport, storage and handling
M&E	Monitoring and Evaluation
MADM	multi-attribute decision model
MCHN	Mother and child health and nutrition
MT	metric tonnes
NGO	Non Governmental Organisation
NPV	net present value
NVS	Nutritional Value Score
ODOC	Other Direct Operational Costs
OE	Office of Evaluation
OECD	Organisation for Economic Cooperation and Development
OpEv	Operations Evaluation
P4P	Purchase for Progress
PRRO	Protracted Relief and Recovery Operation
QALY	quality-adjusted life year (gained)
RB	Regional Bureau
RMI	Business Innovation and Support Division, WFP
RMP	Performance and Accountability Management Division, WFP
SPR	Standard Project Report
TOR	Terms of Reference
UN	United Nations
VSL	Value of statistical life
WFP	World Food Programme
WHO	World Health Organisation
WINGS	WFP Information Network and Global Systems

Glossary

FOOD CONSUMPTION GROUP: A standard classification of food consumption scores into poor; borderline, and acceptable.

FOOD CONSUMPTION SCORE: A weighted diet diversity score calculated from the frequencies of consumption of different food groups by a household during seven days before the survey

FOOD DIVERSITY INDEX: The unweighted number of food groups consumed over the seven-day reference period

NUTRITIONAL VALUE SCORE (NVS): a composite score of the nutritional value of a specific ration of a food commodity, based on the quantities of 14 macro and micro nutrients it contains. On this basis the nutritional values of a typical food basket, cash vouchers or cash transfers can be computed. Software programme available at WFP

Annex A : A selective review of rigorous WFP efficiency studies

Authors	Hidrobo et al. (2012)	Audsley et al. (2009)
Country	Ecuador (Colombian refugees)	Malawi
Comparison	Food, vouchers or cash	Food, cash, and mixture
Methodology	Ex post evaluation, randomised experiment, 2 surveys >2000 households, regression analysis, differentiated for poverty group (tertiles)	Ex post evaluation, randomised experiment
Effectiveness analysis	<u>Output</u> : Food consumption per cap <u>Outcome</u> : dietary diversity index, household dietary diversity score, food consumption score	<u>Output</u> : not discussed explicitly <u>Outcome</u> : food diversity score, food consumption score, food consumption group
Efficiency analysis	<u>Output</u> : cost per food transfer in each modality <u>Outcome</u> : cost of increasing each food security measure by 15%, household dissatisfaction with modality provided	<u>Output</u> : inverse of cost per 1\$ equivalent of food received (alpha value) <u>Outcome</u> : cost of increasing each food security indicator by 1%
External validity limitations	Urban areas of MIC	Lean season food insecurity of farming households in remote region with poorly working markets; problems with cash delivery
Comments		Outcome efficiency leads to different ranking than output efficiency

Annex A (continued)

Authors	Sandström and Tchatchua (2009)	Ryckembusch (2012a)
Country	Sri lanka	Egypt
Comparison	Food, cash	Food, cash, vouchers
Methodology	Ex post evaluation, randomised experiment (at community rather than individual level), 2 surveys >2000 households, difference-in-difference	Appraisal, 2 surveys, nutritional value tables
Effectiveness analysis	<u>Outcome:</u> Consumption patterns, diet quality, household livelihood-related decisions, benefit perception and assessment of modalities	<u>Output:</u> nutritional value transferred <u>Outcome:</u> nutritional improvement observed (food consumption group)
Efficiency analysis	<u>Output:</u> cost per ration (or its cash equivalent)	<u>Output:</u> CER _{o,i} : cost per beneficiary of identical ration, with different food rations compared on basis of nutritional content = cost per nutrient transferred CER _{o1} / CER _{o2} (omega value): cost per nutrient transferred modality 1 / cost per nutrient transferred modality 2 <u>Outcome:</u> CER _{o1} / CER _{o2} (omega value): cost per nutrient improvement observed modality 1 / cost per nutrient improvement observed modality 2
External validity limitations	Well integrated food markets, experienced local bank branches	
Comments		VAT tax means that part of cash transfer is captured by government, which is not the case with food-in-kind

Annex A (continued)

Authors	Regnault de la Mothe and Molinas (2012)	
Country	Based on 8 country studies (which ones?)	
Comparison	School feeding compared with no intervention	
Methodology	Ex ante economic evaluation of 4 categories of benefits from school feeding: real income transfer, productivity gains, decreased mortality and morbidity, cost savings from government	
Effectiveness analysis	<u>Impact</u> : based on available data from COs and from scientific literature	
Efficiency analysis	<u>Impact</u> : full CBA	
External validity limitations	8 country studies	
Comments	Based on optimistic assumptions about the effect of school feeding on learning and productivity, on life expectancy, and on real income and investment, culled from the literature. However many empirical studies are pessimistic about these effects.	

Annex B : An analysis of 13 selected OE evaluations

What	Where	Comments	PS terminology		Input	Activity	Output	Outcome
			Effectiveness	Efficiency				
		1. MALAWI 2009 CPE				√		
cost per metric ton LTSH	§76-78	discussion between WFP and donors on the cost of transport		√		√		
cost per metric ton LTSH	§168-173, §192	comparison of cost per MT with WFP average		√		√		
resourcing ratio	§30, §162	erratic and unpredictable resource flows from donors						
outcome measurement	§59	alignment of HIV/AIDS: The weakness of the Ministry of Health's monitoring systems. was a challenge for WFP which does not have its own systems and depends on counterparts for outcome information. This has hindered WFP ability to develop a clear analysis of outcomes, and enable its own strategic planning.						
baseline information	§94	school Feeding: no baseline information collected per district						
output measurement	§100	school Feeding Programme: there was no exact data on the WFP food commodities delivered through school feeding related programmes						
outcome measurement	§106-107	school Feeding Programme: comparison of project and non-project schools, referring to earlier evaluation						
outcome measurement	§110	school Feeding Programme: the mission compared the project school with the averages for the districts or national averages						
monitoring and reporting	§128	implementation and reporting of nutrition programmes: There were challenges in programme monitoring and reporting, resulting from limited government capacity						
outcome measurement	§141, §144, §205	community Based HIV, FFW, FFA: The M&E system is weaker, focusing on a limited number of output indicators. Programmes were not designed to assess outcome in a systematic manner. At present WFP's monitoring systems are inadequate to that task (outcome measurement).						
cost measurement	§175	management: no effective cost control system is yet in place and management has no possibility of obtaining the overall cost of running the fleet and the final cost per ton of food-aid delivered.						

Efficiency Analysis for WFP – Definitions, Concepts and Methodology

What	Where	Comments	PS terminology		Input	Activity	Output	Outcome
			Effectiveness	Efficiency				
		2. LAOS 2009 CPE				√		
accessing remote areas	§91, 127	discussion on the low population density and limited number of people in need. This increases the costs and lowers efficiency. No ratios.	v	v		v		
process efficiency	§92, 127	discussion on the optimisation of the process of FFW, No ratio	v			v		
timeliness of food distribution	§93	discussion on the timeliness. Response time available for some programmes, discussion on reasons for delay,	v			v		
unit cost comparison	§94, 127	comparing unit costs of FFW with other donors.		v			v	
cost per child fed	§96	comparing the cost per child fed with other programs		v			v	
programme to overhead ratio	§97	comparing programme and overhead at organisational level with other UN departments		v		v		
output monitoring	§93, 103	monitoring data for FFR does not allow determining whether there was a delay in FFR distributions and by how						
output monitoring	§42, 94, 109	weak monitoring system did not record the length of roads constructed.						
monitoring system	§69	generating and using analytical information: analytical information on portfolio performance, that is monitoring data, is not collected systematically across years						
unreliable data	§78	performance, outreach, people fed: the data indicates several years where the number of participants is larger than the number of beneficiaries, which is an indication of the unreliability of the recorded data.						
monitoring system	§138	also recommendation on an improved monitoring system						

Efficiency Analysis for WFP – Definitions, Concepts and Methodology

What	Where	Comments	PS terminology		Input	Activity	Output	Outcome
			Effectiveness	Efficiency				
3. CHAD 2010 CPE								
resourcing ratio	§23 ex. sum., §159, annex 4	73,5% of the financial requirements were met	v		v			
beneficiaries reached	§24-27 ex. sum., 117	75-200% of beneficiaries reached	v				v	
efficiency at country level	§28-29 ex. sum., §168	efficiency was hampered by the fact that activities are designed and implemented within the context of individual operations						
cost per MT	§164	food commodity cost per MT, total cost per MT, comparing against WFP corporate average		v		v		
complete logistics assessment	Annex 4	comparison of cost components with WFP corporate averages		v		v		
recruitment	§32 ex. sum., §170	some key positions remained vacant for periods up to one year						
monitoring system	§20 ex. sum., §169	the M&E system was weak and limited the evaluation's assessment of the performance and results of programme activities. The focus was only on initial data collection and output indicators, not outcome indicators.						
attribution problem	§ 36 ex. sum							
monitoring system	§74	WFP M&E system has generally been weak, focussed on output indicators						
cost per school	§140	anecdotal evidence					v	
outcome measurement	§146	lack of baseline information or measurement of indicators defined in the country programme						
timeliness	§166	distribution organised in a hurry to "beat the expiry date", given the difficult logical situation						
monitoring system	§168	monitoring of activities was difficult either as a result of security (at various times in several parts of the country and currently in eastern Chad), dispersion and long distances over which activities are taking place, or insufficient staff or financial resources (most acute in the						
monitoring system	§200	recommendation to put in place a M&E system						
ODOC as % of total costs	Annex 4	compared to WFP average						
cost per MT per km	Annex 4	suggestion p 73, possible with new software						
funding ratio	Annex 4							

Efficiency Analysis for WFP – Definitions, Concepts and Methodology

What	Where	Comments	PS terminology		Input	Activity	Output	Outcome
			Effectiveness	Efficiency				
		4. NEPAL 2010 CPE				√		
factors having an impact of efficiency	§269, §225-228, §264	several factors had an impact on efficiency (conflict, insecurity, delays in food delivery...), no ratio	v			v		
timeliness	§231, 269	a government direction to WFP to purchase food outside the country resulted in shortfalls and delays in food delivery						
pipeline break	§234, 269	pipeline break in 2009 as a result of funding constraints						
beneficiaries reached	§270, §145, §147, §149, §151, §152, §154, §156, §157, §164, §270, Annex 16, 19, 20, 22	number of beneficiaries reached and amount of food distributed	v			v	v	
cost/MT	Annex 17	LTSH/MT				v		
quality of assets created	§187 vn 107	assets were sub-standard, reference to other evaluation, anecdotal evidence		v			v	
monitoring system	§56	the country has developed a strong assessment and analytical capacity						
monitoring system	§80, 107, 127, 210	the vulnerability analysis and mapping system (VAM) currently manages NEKSAP, the most comprehensive food security monitoring system in the country						
monitoring system	§128	description of a new improved M&E system that will measure outcome						
monitoring system	§132-134	WFP has established an impressive network of M&E field-level officers (also monitoring efficiency of food aid distribution). But, while it generates useful information, it only indirectly captures the impact of WFP Nepal's interventions. Surveys also collect a lot of data, but improvements can be made in the collection of data.						
monitoring system	§170, 173, 174	poor reporting on outcome						
monitoring system	§289	lack of M&E of MCHC programmes, recommendation to improve the M&E system						
monitoring system	§294	specific recommendation on M&E, recomm 5 and 10						

Efficiency Analysis for WFP – Definitions, Concepts and Methodology

What	Where	Comments	PS terminology		Input	Activity	Output	Outcome
			Effectiveness	Efficiency				
		5. HAITI 2011 CPE					v	
effectiveness and efficiency	§140, §152, §163, §179, §192, §238	comparing planned and actual beneficiaries	v					
effectiveness and efficiency	§239	WFP has been able to distribute an unprecedented amount of food aid quickly and efficiently. No ratio		v		v		
commodity cost per MT	§219, 220	comparing to WFP corporate average		v	v			
different cost	§220, §221			v		v		
timeliness	Annex 4	annex 4 gives the evaluation matrix. Under efficiency and effectiveness it suggests an indicator transit time (not elaborated in the report).						
monitoring system	§47, §111, §189, §240, (§58 of annex 1 ToR evaluation)	WFP's M&E system during the evaluation period was weak, providing a limited amount of data useful for assessing impact and improving programmes. The system measure inputs to its programmes, tracks basic output data, but even the basic beneficiary data is variable and inconsistent. Without this data, it is difficult for WFP to assess where its contributions have been the most effective and efficient. WFP's inability to show relative impact of their interventions in such a competitive donor landscape is a critical disadvantage.						
monitoring system	§103, §131	the VAM unit is a strong contributor to the food security body of knowledge. This information is however not necessarily used at an operational level.						
monitoring system	§223	It is difficult to go deeper into the cost analysis given the absence of an analytical accounting system						
monitoring system	§245	first recommendation on strengthening the M&E system						

Efficiency Analysis for WFP – Definitions, Concepts and Methodology

What	Where	Comments	PS terminology		Input	Activity	Output	Outcome
			Effectiveness	Efficiency				
		6. UGANDA 2001 PRRO				√		
performance against assessed need based on rations provided	p8, Annex E	actual distribution compared to needs	v				v	
security	p 39	security is a considerable constraint on the efficiency and effectiveness of operations. No ratio						
pipeline rupture	p 41	as a result of a pipeline rupture, rations of refugees have been halved.						
efficiency by combining two projects	p 47	combining two projects has resulted in some rationalisation of management and other efficiency savings						
monitoring system	p vii	more attention should be paid to verifying refugee and IDP numbers, improving final distribution mechanisms, and post-distribution monitoring						
monitoring system	p14-15	field staff regularly visited FFW sites but results were not assessed and it is not clear from reports what has really been achieved.						
monitoring system	p31	the project document includes a list of indicators but they are essentially activity level indicators and not helpful for monitoring achievements at the output and objective levels. In practice monitoring has focused on tracking commodities and commodity utilisation. There has been little monitoring of the effectiveness of relief and recovery						
monitoring system	p 33	a lot of quantitative data is generated, particularly commodity tracking, logistics and finance. There is however little qualitative reporting and even the quantitative reports do not record performance against						
monitoring system	p 33	recommendation to invest in the capacity of field staff to conduct post-distribution monitoring.						

Efficiency Analysis for WFP – Definitions, Concepts and Methodology

What	Where	Comments	PS terminology		Input	Activity	Output	Outcome
			Effectiveness	Efficiency				
7. SOMALIA 2002 PRRO								
service cost	p 23, 27-28	analysis of costs per MT, analysis of the different components (DOC, DSC, ISC, LTSH), evolution over time		v		v		
timeliness	p24-25	discussion of the response time needed to deliver food						
monitoring system	p 19	as a separate consultancy focuses on M&E, this evaluation did not go into this subject in depth. Reporting is generally excessively lengthy and uses indicators which are not always seen to be relevant by WFP field staff.						
monitoring system	p 33	recommendation to enhance post-distribution monitoring (e.g. redistribution within the community).						
monitoring system	p20	M&E is often hindered by various factors like distance, security, limited time...						
What	Where	Comments	PS terminology		Input	Activity	Output	Outcome
			Effectiveness	Efficiency				
8. TAJIKISTAN 2006 PRRO							v	
cost per MT	p11	average total cost delivered to WFP warehouses for in-kind contributions in 2005 was US\$15/MT compared to US\$221/MT to buy the same goods in Kazakhstan. This means that cash contributions would buy 2,3 times more food for the same amount of money		v	v			
cost per MT	p 11	comparison of local market prices versus costs of WFP assistance by procurement source in 2005 Somoni/kg		v	v			
cost efficiency by activity	p 15-16	calculating Alpha value. When WFP is able to purchase regionally, it can bring in the food cheaper than it would cost the beneficiary to buy the same ration in the local market. For in-kind contributions, all alpha values are below 1. Overall comparison when to buy regionally, buy in kind, give cash		v			v	
Alpha value	p 24	on the basis of the Alpha value, cash-for-work would be more cost-efficient than FFW.		v			v	
monitoring system	p 41	the operation has made a commendable effort at outcome monitoring. A challenge to the general functioning of the system is the lack of a database to effectively capture information. Recommendation to earmark funding for the completion of a country office monitoring and reporting system						

Efficiency Analysis for WFP – Definitions, Concepts and Methodology

What	Where	Comments	PS terminology		Input	Activity	Output	Outcome
			Effectiveness	Efficiency				
		9. HAITI 2007 PRRO						
resourcing ratio	§235	level of financing received: 70%	v		v			
Personnel	§239-248	a section on the adequacy of human resources: 87 personnel, sufficient number; but too few monitors to visit project sites and no "pipeline officer"						
Partnership with national government	§249-256	how the partnership with national government, donors and local partners affects the functioning of WFP						
Capacity strengthening of implementing	§257-268							
Operational mechanisms	§269-286	explanation of the selection procedure of eligible health centers and beneficiaries, and distribution mechanism at centers...						
Adequacy of food assistance	§287-296	adequacy compared to the usual food basket of beneficiaries, WFP food directives						
pipeline breaks	§297-301	late arrival and empty stocks made the delivery of planned food baskets impossible, no ratio. Food baskets have been adapted to maintain the nutritional value as much as possible.						
Transport	§304-309	Accès maritime et portuaire						
Entrepot	§310-315							
Distribution mechanism	§316-324	distribution at final distribution point						
Transport	§325-335	national transport in Haiti						
Lack of stock	§336-339	analysis of the reasons of a lack of stock and solutions found by the country office						
Cost analysis	§340-341	comparison of costs in project document, revised and effective						
Rationality for food	§348-353	explaining the necessity of food assistance						
Cost analysis	§342-347	comparison of costs/MT (Transport Terrestre, Entreposage et Manutention, Coûts d'Appui Directs et Coûts Opérationnels Directs) with WFP average		v		v		
monitoring system	§354-360	the monitoring system is not only focused on output monitoring but also on outcome. Different indicators are however not measured. Monitoring staff do not have enough time to visit projects sites regularly enough.						
monitoring system	§361	information collected by implementing partners is not always comparable or in line with the logical framework making systematic analysis difficult.						
monitoring system	§366	the M&E system developed has not been implemented.						

Efficiency Analysis for WFP – Definitions, Concepts and Methodology

What	Where	Comments	PS terminology		Input	Activity	Output	Outcome
			Effectiveness	Efficiency				
		10. TIMOR LESTE 2009 PRRO				√	√	
comparing target and actual (food and capacity development)	§30-66	overall, later per project component	v				v	
logistics	§67-76	discussion on capacity building of WFP staff, government, commodity managers...						
adaptation to changing needs	§77-83	discussion of the port of Dili, warehouse in Dili, Timor Leste road infrastructure,						
external institutional arrangements:	§87-89	to what extent has WFP been succesful in adapting its approach to the chaning conditions in Timor Leste.						
internal institutional arrangements	§90-92	discussion of the partnerships with government ministries, other UN partners, NGOs and private sector						
pipeline breaks	§94-95	discussion of the staff positions and their impact on functioning						
cost and funding of the operation	§96	insufficient funding and late arrival of commodities caused pipeline breaks, to which the CO responded through advocacy with the government and donors and suspending FFA activities for seven months in 2009.						
cost to WFP per meal	§96-100	discussion of pipeline breaks and how the problem was solved, how this could be avoided,						
M&E	§101	cost to WFP per meal, per student in 2008 was US\$0,11 and in 2009 with a smaller ration was US\$0,07.						
	§84-86, §151	the Indicator matrix includes baseline data, output and outcome indicator data for 2008, targets for 2009, data sources and collection method, frequency of collection, responsibility for collection and use of information. However, M&E raised some problems as some monitors lack necessary skills and the evaluation team saw some reports from communities that were clearly inaccurate.						

Efficiency Analysis for WFP – Definitions, Concepts and Methodology

What	Where	Comments	PS terminology		Input	Activity	Output	Outcome
			Effectiveness	Efficiency				
		11. NIGER 2011 EMOP				√		
cost per metric ton	§128	comparison of actual cost per MT (food cost, ITSH, LTSH, ODOC, DSC) with WFP average and project documents		v			v	
cost ratio	§86	cost to transfer cash transfers: 14% or 0,16 US cents to transfer 1 USD		v			v	
timeliness	§134	calculating the number of days it takes to deliver. Comparing international (90-120 days) and regional purchase (30-50 days). The difference in number of days is important in emergency operations.			v			
evaluation diagramme	§170	In the evaluation diagramme they score 10 different evaluation criteria. They do not explain how they got to this number (Annexes are not available)						
monitoring system	§53, 90, 157	important discrepancy between the logical framework and indicators effectively measured.						
monitoring system	§53, 156	Weak monitoring system and compilation of data						
monitoring system	§158	WFP uses implementing partners for M&E. The request of WFP limits to quantitative aspects (beneficiaries and rations) to fill in the COMPAS and WINGS database. Even here discrepancies exist in the provided information						
	§11-12	12. COLOMBIA 2010 PRRO					√	
Conversion of contributions to products	§56-57	conversion of contributions to products: which amount of planned financial resources was received and what amount of planned products has been distributed.	v				v	
Timely delivery	§58	% of products delivered within 30 days of date programmed: 79%	v		v			
Quality of products	§59	anecdotal evidence: during project visits, the team only received one comment on lack of quality						
Pipeline breaks	§60	major pipeline breaks could be avoided. Some breaks in the less important products like salt or sugar						
cost analysis of non-alimentary costs	§62	LTSH, ODOC and DSC costs as a % of total costs, compared with different countries	v		v			
different efficiency indicators per project component	§64-72	number of beneficiaries and metric tons (compared to planning), number of project sites, incidents of security, timely delivery, number of participants	v		v		v	
monitoring system	§154-158	during the project a monitoring system has been set up (Sistema de Informacion Operativa). Project monitors are highly satisfied with the system.						

Efficiency Analysis for WFP – Definitions, Concepts and Methodology

What	Where	Comments	PS terminology		Input	Activity	Output	Outcome
			Effectiveness	Efficiency				
2. B Productos y proceso de ejecución: elementos de eficiencia	§ 52 - 98	13. GUATEMALA 2010 PRRO				√	√	
comparison of actual with planned	§viii, §52	43,6% of beneficiaries have been assisted	√				√	
comparison of actual with planned assisted	§viii, §53-54	the program has been implemented in 109 instead of the planned 83 municipalities	√				√	
financing ratio	§ix, §59-60	43% financed	√		√			
cost analysis	§63-64, annex 8 p69	food cost, external transport, LTSH, ODOC, DSC, ISC: comparison actual versus planned	√		√			
cost per beneficiary	§x, §69	31 USD per beneficiary		√			√	
% lost	§78	0,019% of food was lost < 2% tolerated by WFP						
Alpha value	§70, annex 9 p70	the alfa value confirms the efficiency of local purchase, Detailed costing in analysis but the actual calculation of the alfa value is not included.		√	√			
provision of Vitacereal (the product provided)	§71-77	amount purchase, amount received, amount distributed						
tendering process	§80-81	explanation of the tender process and information in the WINGS system						
financial panorama	§82-83	a discussion of supportive projects and the mobilization of additional financing						
staff input	§84	discussion on the importance of staff						
mechanisms of coordination	§85-87	discussion of the coordination with the Ministry of Health, NGOs...						
participation	§88	discussion on the participation of beneficiaries						
monitoring system	§24, 91	as no indicators have been identified to monitor the impact of the project, evaluation is difficult. Information is limited to the number of beneficiaries.						

***Annex C : Fortified Blended Food Factory Project Timor Vita –
Comments from an efficiency perspective***

In what follows we do not question the desirability of acquiring fortified blended food (FBF) for distribution in nutrition programmes. In terms of the logical framework of the intervention, this means that we do not question that a positive output (FBF rations delivered to pregnant and lactating mothers in health clinics) will transform into a positive outcome (short term nutritional improvement) and a positive impact (lasting positive effect on health and well-being). Neither do we analyse whether that output is effectively secured. We instead focus on an activity that is necessary to achieve that output. This activity can be described as: appropriate FBF is available in central warehouses in TL for distribution to health clinics. When there is more than one way to achieve this activity-level goal, as is the case here, efficiency analysis can help choose the most desirable one. Five logical steps can be distinguished. (1) Appropriate alternative courses of action are identified. (2) The technical feasibility of each alternative is established (can it be done, and what is required to do so). (3) The financial profitability of each alternative is studied. (4) The economic profitability of the each alternative is studied for society as a whole. (5) Relevant non-economic considerations are brought into the analysis. We discuss each in turn.

Stage 1: project identification

The two obvious alternatives to achieve the activity-level goal are, on the one hand, importing FBF, and on the other hand, procuring agricultural commodities from local farmers, importing vitamin and mineral nutrients, and organising the production (milling, extrusion-cooking, mixing, packaging,...) in an industrial facility in the country. The import alternative may consist of importing a corn-soya blend (CSB) and other ingredients such as oil and sugar separately, store them separately in warehouses, and leave it to health staff to do the scooping out and mixing at the moment of final delivery to the beneficiaries. This is what has been happening in practice in the past, and is understandably the alternative mostly referred to in WFP discussions and documents. There is however a serious methodological issue here. There are reportedly important differences in taste and nutritional quality between the two alternatives so defined, and also a difference in wastage during final storage and distribution at the health centres. In efficiency analysis, it is important to compare the project under consideration with an alternative that is a close substitute. In this case it would have meant comparing the project with an imagined but feasible alternative such as importing fully prepared and packaged products that are similar in taste, nutritional content, and ease of use to the local product. This would of course have involved some additional market research. This was not done, and as a consequence the comparison becomes unwieldy. In WFP documents this lack of comparability is subsequently used as an argument in favour of the project, as we will see below, but this is methodologically questionable.

Stage 2: technical feasibility study

The feasibility of the import alternative that is retained in WFP thinking did not have to be researched, for it was directly observed in the past: WFP did import and distribute CSB and

complementary products in TL for some time. In 2005 a technical feasibility study ³⁹ was performed by WFP of the local procurement alternative. It came to a negative conclusion, essentially because of an underdeveloped local agricultural sector, characterised by high prices compared to the world market, the difficulty of securing sufficient quantity, and quality concerns. A further reason was that no private sector partner was at that time willing to invest in the project. The report suggested to wait until the agricultural sector was more developed. Some years later, a private sector partner showed interest, and MoH and WFP and the private company, Timor Global, in 2008 agreed to collaborate on the setup of a factory. Additional feasibility analysis was performed but did not address the main reservation expressed in the 2005 report about agricultural supply capacity. Yet the reality in agriculture had not changed, and the 2005 report has been vindicated by the facts: the project has not until now reached its anticipated purchases of domestic agricultural commodities, and has resorted to importing most of the commodities. Contract farming or targeted advice to a group of farmers suppliers could have been an interesting alternative to explore, but was not considered. During the visit to the project in December 2012 the manager told us that there had been several recent proposal by farmer groups to enter in contracts for the sale of maize. A number of other problems occurred during implementation that may be causally related to the lack of a sufficiently comprehensive feasibility study and the related failure to set out realistic operational timelines, training needs and production schedules. A recent WFP review ⁴⁰ lists a number of problems that occurred. Here is an overview.

- The project changed its mind about the mixture of products to be included in the FBF after machines for a different mixture had already been ordered and installed, leading to the need to order additional parts and send a team to adapt the machinery, and further leading to the loss of commodities of the discarded mixture that had already been donated to start up production.
- Electricity consumption was not estimated realistically, and generators had to be acquired later on.
- The project bought local agricultural commodities without testing them, as a result of which a whole batch of final product was found to be contaminated and had to be discarded.
- It proved impossible to recruit enough local staff with the necessary technical expertise.
- The absence of in-country testing of food quality was not taken seriously.
- The time between decision point and operation was seriously underestimated.

³⁹ “Local production of Fortified Blended Food in Timor Leste – Challenges and Potentialities”, WFP Timor Leste, 22 – 30 November 2005

⁴⁰ “Review of a Locally Fortified Blended Food Factory Project: A Timor-Leste Case Study” (undated draft)

Of course projects always run into unexpected difficulties⁴¹. But a more careful preparation would at least have shown that the project was only viable with imported commodities, at least for the foreseeable future.

Stage 3: financial analysis

For WFP the idea of contributing to a national capacity in the production of FBF was attractive. A problem was that the project as designed would be uncompetitive in the world market, and that by buying the Timor Vita product WFP would not procure from the cheapest source. The CO had to request a waiver from Rome HQ. The request was argued and granted on the basis of considerations of product quality and contribution to economic development of the country (see below). In addition WFP and its bilateral donors were happy to contribute generously to the investment and recurrent costs of the factory (feasibility studies, oversight, advise, TA).

For the MoH the project is also attractive, again apparently mainly on economic grounds, which would justify the subsidies that are granted to the project directly (contributions to the physical and working capital) and indirectly (buying the product at a much higher price than the world market dictates).

Beneficiaries obtain FBF for free, but there is an opportunity cost of the time spent to come to the health clinics. And there is a broader picture to consider. Free food rations act as a financial incentive to overcome the reluctance of pregnant women and lactating mothers to attend health clinics. If more women attend compared to the counterfactual situation where no FBF is distributed, this indicates that the project is sufficiently attractive to them.

For Timor Global financial considerations are of more obvious concern than for the other actors. A private company strives to ensure a sufficient return on the capital it invests in a project. No detailed information on the financial costs and benefits for Timor Global is available, but if the project is not financially profitable it is highly unlikely that the firm will continue to participate in it. WFP and the MoH have been generous. The machinery was paid by the MoH, WFP paid for the generators and a large tent acting used as warehouse, and the company does not have to pay compensation for any of these. There was an intention to include a clause in the MoU with the company stipulating that it would reimburse the cost of the machinery (\$250,000) over a period of ten years by a discount on the sales to MoH. However the MoH has resisted this inclusion because of the possible implications for the ownership of the machinery, which the MoH wishes to retain for reasons that are not entirely obvious. There is also recurrent financial support through the provision of a food technologist who is working on-site and additional advice. A long list of donations from different sources are mentioned in WFP documents, but are not described in detail. The most important concession is however well documented: the commitment of WFP to buy

⁴¹ Hirschman has called this the principle of the “hiding hand”. Even successful projects tend to encounter all sorts of unexpected problems that, if anticipated, could possibly have led to the decision not to do the project. But once encountered, management often finds the energy to overcome them and be successful. The fact that not all problems are recognised during appraisal may therefore be something of a blessing. Hirschman, A.O. (1967), *Development Projects Observed*, The Brookings Institution, Washington D.C.

2400 MT/yr at an agreed price calculated on the basis of a 10% profit margin above costs, and without any compensation for in kind contributions by the MoH or donors⁴².

Stage 4: economic analysis

For Timor Leste and its citizens, the project is worth doing if the costs outweigh the benefits. The economic benefits consist of the foreign exchange saved by not having to import an equivalent tonnage of FBF from the world market and bring it to a central warehouse. The costs consist of the foreign exchange cost of the machinery (whether donated or not), imported inputs, UN and government staff time (whether or not paid for by Timor Global), the economic cost of electricity (which may be higher than the cost paid by Timor Global), the cost of the buildings and the use of the land, the cost of local agricultural commodities and other products bought in the market (expressing indirectly the cost of producing them), the opportunity cost of local skilled and unskilled labour employed by the factory (which may be different from wages paid by the company), and the like. As the project has a lifespan of about ten years, the economic cash flow is spread out over the same time period. A discounting procedure need to be applied to acknowledge the opportunity cost of the resources invested in the project. On the basis of all the above the net present value (NPV) of the project can then be calculated. The project is worth doing if the NPV is positive. If it is negative, importing is the better alternative.

Unfortunately no such economic CBA was performed, not even in simplified fashion, although it would have been possible to do so at reasonable cost. The data available during the field mission did not permit to reconstruct the financial and economic tables. What is available is a cost per MT or per ration comparison between the project and the imported alternative. This is a simplified but nevertheless interesting version of financial CBA. From this comparison it appears that the project is about 90% more expensive than the alternative of importing⁴³. This calculation is incomplete from an economic perspective however, for it does not consider a series of additional costs, such as the contributions by donors and the government. If such costs are added to obtain a simplified economic comparison, the project would in all likelihood look even less unattractive. Nevertheless the project is defended by WFP as the better alternative. First it is argued that the two alternatives that are discussed are not quite comparable: the FBF produced by the project is more tasty, easier to cook, easier to carry, and leaves time to health centre staff to do more important work. This may be correct, but is the consequence of making the wrong comparison. We just do not know how the project compares with a closer alternative until we actually do the calculations. The lack of such a calculation cannot be used to defend the project. A sloppy methodology is here turned into an argument in favour of the project.

Secondly, two economic arguments are offered. The first argument is that local procurement will stimulate development in the country . The second arguments is that the factory employs around 40 local staff, which has a direct economic impact in the region . Unfortunately the strength of these arguments cannot be established without a proper economic CBA. If for instance a more elaborate economic CBA had revealed that the project has a negative NPV, it would mean that the project stimulates the development in the country less than when

⁴² \$1378/MT for November 2012, data obtained from Antonio Avella of the CO.

⁴³ Based on the same data obtained from Antonio Avella

resources are left in their best alternative present use. The fact that it would employ 40 staff would not be an argument in that case either. For the benefits for those 40 workers, and the further effect on the region, would be less than the benefits elsewhere, including possibly to workers, in the case of not doing the project and leaving resources to their best alternative use.

All this does not mean that the project is not efficient. But if it is, it is not the project as described (local procurement) and not on the basis of the arguments that are provided in the documents that we have read. We also add that the fact that the project is apparently not financially viable (or at least that is what the company suggests) is not in contradiction with it being economically efficient. It may be a case of the “infant industry argument”, where a project is economically viable in the long term, but goes through a learning phase where it is not competitive. This justifies that it is temporarily subsidised (or protected against imports, which is not the case here, but is how the case is known in the literature). But again, these effects can be calculated. This is what CBA is for.

Stage 5: non-economic considerations

In defending the project, WFP offers a more political argument. It is stated that the government of TL prefers to buy FBF locally rather than import, and that therefore the future of the project, especially after the possible departure of WFP at the end of 2013, is better guaranteed when the project is realised. This may well be the case, but the question is whether this is a good enough reason to embark on a project that may well have a negative economic NPV.

Conclusion

Some not too complex efficiency analysis could have been usefully performed to clarify the decision making for this project and to set out a clear plan for its implementation and level of subsidy to the private company involved. That it was not done cannot be blamed on the CO alone, or even in the first place. It is not realistic to assume that a small CO has the economic expertise to do such an analysis. What is more surprising is that at the regional and HQ level the decision was endorsed on such flimsy arguments.