

# Independent Comprehensive Evaluation of the Scaling Up Nutrition Movement, 2014-15

## Statement by the Evaluation Quality Assurance Advisors

---

Julia Compton, Chris Gerrard, and Eva Lithman

May 1, 2015

### Main Points

**Evaluation independence was adequate:** In the judgement of the Quality Assurance Advisors (QAA) contracted for the SUN-ICE, the evaluation has been adequately independent. It has conformed to the four widely recognized criteria of evaluation independence: organizational independence, behavioural independence, protection from outside interference, and avoidance of conflicts of interest. The evaluation originally had no independent Evaluation Manager, which raised concerns about organisational independence, but one was engaged after the evaluation team was selected, following a strong recommendation from the QAA.

**Quality of evidence was generally sufficient:** In the judgement of the QAA, the evaluation was generally of good quality, objectives-based and evidence based — presenting credible links between the findings, conclusions and recommendations.

**Unrealistic expectations for what could be achieved in the time frame had some adverse effects on quality of process, despite impressive efforts from the evaluation team.** In the view of the QAA, the evaluation Terms of Reference (ToR) was too demanding given the time and resources allocated. The rushed nature of the process adversely affected the depth of the analysis of the extensive and rich material assembled by the evaluation team. The restricted timeline also meant that broad stakeholder consultations were not undertaken at the draft report stage, as is normal practice. We suggest that for future such exercises quality advisors be engaged earlier, to provide comments on the evaluation ToR before it is finalized.

**Further work is needed on gender and equity aspects:** It was not possible to address gender and equity aspects in adequate depth in this evaluation, although they have been shown to be critical to overall nutrition outcomes. The QAA have therefore suggested that a separate study or studies be carried out on gender and equity as part of the preparatory work for the next SUN strategy.

## 1. Introduction

The Scaling Up Nutrition movement (SUN) has recently undergone an independent comprehensive evaluation (SUN ICE), which was planned in 2013 and undertaken in 2014 to early 2015. The evaluation terms of reference (ToR) made provision for the appointment of independent quality assurance advisors (QAA) and the three of us were appointed in April 2014 through a competitive process. Our formal accountability has been to the SUN Visioning Sub Group (VSG) — an international group of senior technical staff chaired by the Executive Director of UNICEF and appointed by the Lead Group of the SUN Movement to oversee the SUN ICE. Our principal role has been *“to aid the VSG in assuring that both the process and the product of the evaluation are credible and independent”* (Evaluation ToR, p. 12).

The purpose of this note is to provide an overall assessment of the independence and quality of the SUN ICE in relation to international evaluation standards, and to make some comments on the process.

The main audiences for this note are:

- The evaluation clients and evaluation team, as a public statement that we consider the evaluation independent and good quality.
- SUN governance structures, and those of other complex international initiatives, for some lessons that they may find useful for the management of similarly complex evaluations in the future.
- Other evaluators employed as Quality Assurance Advisors.

## 2. Inputs of the Quality Assurance Advisors to the SUN-ICE

The QAA were contracted after finalisation of the SUN-ICE ToR and before the recruitment of the evaluation consultants. Our main inputs have been:

- Technical appraisal of the bids for prospective teams of independent evaluators.
- Reviewing the Inception Report, the Interim Progress Report (including the assessment of the SUN Movement Secretariat, SMS), and the draft final report *“with regard to their adequacy, methodological rigor, application of good practice in comprehensive evaluations, soundness of evidence, and independence”* according to international evaluation standards (Evaluation ToR, p. 13). We also reviewed the discussion paper on future options for the SUN movement and the main survey instrument — neither of these being in our original ToR.
- At each stage, we provided comments and advice both to the VSG and to the evaluation team, and participated in several meetings of the VSG during which our comments and advice were considered. These comments were purely advisory; decisions on evaluation products were taken by the VSG.

Two additional contributions of the QAA to the SUN-ICE process have been:

- **Recommending recruitment of an independent evaluation manager:** At the time of reviewing the bids from prospective evaluation teams (May 2015), it was not clear who was actually managing the evaluation. In order to promote evaluation independence, we recommended that an independent evaluation manager be hired by the VSG. This recommendation was accepted by the VSG and an evaluation manager was hired by the Bill and Melinda Gates Foundation on behalf of the VSG. The QAAs provided a sample ToR for the evaluation manager and commented on the final ToR, but was not involved in recruitment.
- **Recommending that a formal management response** to the final evaluation report be elaborated, involving broad consultations with SUN stakeholders.

### 3. QAA Comments on the Evaluation Process and Substance

#### 3.1 Evaluation Independence

We find that the evaluation process has been independent according to four widely recognized criteria for independence in evaluation: organizational independence, behavioral independence, protection from outside interference, and avoidance of conflicts of interest.

**Organizational independence:** The SUN ICE was commissioned by the Lead Group (the highest level governing body in the SUN Movement) and overseen by the VSG — both of which were appropriate. The QAA had initial concerns about the organizational independence of the day-to-day management of the evaluation, but these concerns were addressed following the recruitment of an independent evaluation manager accountable to the VSG (see above). The evaluation manager became responsible for managing all aspects of the evaluation process, and for liaising among the VSG, the QAA, the evaluation team, and the SMS.

**Behavioral independence and protection from outside interference:** The evaluation team comprised experienced evaluators who have produced a candid and uncompromising report whose findings, conclusions, and recommendations are being reported transparently. We are not aware of any outside pressures to influence their evaluation findings, conclusions and recommendations. The evaluation manager also played an important role in protecting the team from any potential interference.

**Avoidance of conflicts of interest:** All the teams who submitted proposals at the outset of the evaluation process were required to indicate any conflicts of interest and how they proposed to manage such conflicts which existed. The winning team did not indicate any such conflicts; some other proposals scored lower due to the existence of such conflicts.

The QAA have also been free from conflicts of interest in relation to the SUN Movement and the evaluation team. We have not experienced any interference in our work.

**QAA Comment:** It is important to pay attention to all four aspects of evaluation independence during the planning stages of an evaluation. Putting in place institutional arrangements for organizational independence can be particularly challenging when the bodies that are

commissioning and facilitating the evaluation (in this case the Lead Group and the SMS) are also subjects of the evaluation.

### 3.2 Evaluation Scope

Evaluating the SUN Movement was always going to be challenging, given that:

- SUN is a “movement”, rather than a typical global partnership program with well-defined activities that the program itself chooses and implements. SUN has been more of an advocacy program that is dependent on its members — particularly national governments — to choose and implement activities.
- SUN has a complicated and changing structure now embracing more than 50 countries and four global networks: the Donor Network, the Business Network, the UN System Network and the Civil Society Network.
- The objectives of SUN are complex and changing over time as the nature of malnutrition changes (Global Nutrition Report, 2014).

The evaluators were faced with two further major challenges: the evaluation timeline and the ambitions of the evaluation.

**Evaluation timeline:** The Lead Group made its decision in September 2013 to seek an independent evaluation of the SUN Movement by the end of December 2014 that would feed into a visioning process for its next phase that would take place in 2015. But the call for proposals from prospective evaluation teams did not take place until March–April 2014, and the evaluation team was not contracted until June 2014, so that the team had only six months to complete the evaluation. This very tight timeline has been a major constraint for the evaluation team. It also put pressure on the QAA, who were asked to provide comments within very short timescales, including over weekends and holidays.

**Ambitions:** As implied by the evaluation title, the evaluation ToR attempted to be “comprehensive”:

*“The Independent Comprehensive Evaluation (ICE) of the SUN Movement is to consider all aspects of SUN — its institutional structure, objectives, working model(s), decision processes, role within the wider architecture of international development, relevance, value-added, efficiency and effectiveness. It will address how effective SUN has been in carrying out its objectives — concerned with accelerating the reduction of undernutrition — and to pose options for evolution of the SUN movement to build on strengths and address weaknesses. It will provide an independent assessment of what SUN has accomplished and is accomplishing, the efficiency and effectiveness of its different components (its governance, networks and secretariat), its current functioning and to the extent feasible, its contribution at country, regional and global levels. It will examine the extent to which SUN is helping national governments, and other stakeholders, to contribute to transformations in the way nutrition is being addressed. And it will assess the role of SUN in increasing attention to women’s empowerment and gender equality and in catalyzing nutrition-sensitive approaches in*

*agriculture, health care, water and sanitation and other sectors” (Evaluation ToR, pp. 2-3).*

This was a very broad and ambitious agenda for a single evaluation, and would have been very challenging to manage even if ample time and resources had been available. The evaluators were requested to define options for the future of the SUN Movement — the latter being more usually the role of management consultants than evaluators. The evaluation team was also expected to make a judgement on the legitimacy and effectiveness of the SMS. Furthermore, due to the compressed timetable, both of the above were expected to be completed at an early stage of the evaluation, before most of the fieldwork had been undertaken.

**QAA comment:** In multistakeholder evaluations, it is common during the planning stages to find that the evaluation ToR becomes more and more demanding and complex, as an attempt is made to include everyone’s comments and wishes for the evaluation. Often there are few countervailing forces to the drive for increased complexity. While stakeholders often ask for things to be kept simple, this rarely includes sacrificing their own pet evaluation issues. Those commissioning evaluations need to start planning well in advance of the actual conduct of the evaluation. **For future evaluations, we would suggest that quality assurance advisors be contracted at an earlier stage in the evaluation process and be given a chance to comment on the ToR before this is finalised and issued.**

### 3.3 Evaluation Design and Questions

As mentioned above, the evaluation team faced two major challenges: defining the boundaries of the SUN ICE, and conducting the evaluation in a short timeframe with a limited budget. The evaluation team embraced these challenges and defined an ambitious approach in the Inception Report. This elaborated a set of evaluation questions based on a preliminary theory of change with underlying assumptions that was constructed, among other things, from the SUN Movement’s four strategic objectives and its approach to monitoring and evaluation and accountability. The evaluation matrix of the Inception Report contained 6 main evaluation questions and 26 subquestions.

Addressing all these questions in sufficient depth would have been a major challenge for any evaluation team, even given unlimited time. As it was, the evaluation team had to make trade-offs between questions to address and with different degrees of depth.

The final evaluation report does not provide a matrix which enables a direct comparison with the plans set out in the Inception Report, or explicitly state which evaluation subquestions were addressed in what depth. However, the impression of the QAA is that while overall answers were provided to the six main evaluation questions, many of the subquestions feeding into these were not addressed in much depth.

In the view of the QAA, the evaluation team performed well in difficult circumstances. Nevertheless, there were some areas which received less attention than would be ideal. These included:

- **Overall analysis and reflection:** The team invested a lot of resources in collecting data, including visiting eight SUN countries, but was then left (in our view) short of time for overall reflection and analysis, since the time between the final country visits and the submission of the draft evaluation report was less than two months. This reduced the depth of the summative analysis and assessment of the SUN Movement, including its theory of change. Many readers of the evaluation may well want to reflect and draw some of their own conclusions from the wealth of information presented.
- **Consultation and feedback:** Some important aspects of the usual evaluation process were missed out, due to the pressure to deliver a final report on time. In particular, there was no time to follow the common practice of circulating the draft evaluation report and recommendations widely among stakeholders for comments. (This has however occurred after the final evaluation was issued.)
- **Gender and equity:** Gender and equity issues are both recognized as key to improving nutrition (Global Nutrition Report 2014). The evaluation presented some evidence that *"SUN's aspirations for ensuring a systematic focus on gender empowerment have not yet been moved from global-level rhetoric into country-level practice"* (p. viii), but was not able to investigate the practical issues, particularly for gender and nutrition. The QAA have therefore suggested that a separate study or studies be carried out on gender and equity as part of the preparatory work for the next SUN strategy.

The QAA did raise the issue of the broad scope of the evaluation questions, the high expectations, and short timeframe at the inception phase, saying that *"Serious consideration should be given to eliminating some subquestions ... and reducing expectations of others ... and 'managing down' stakeholder expectations at this inception stage, given the very tight timeframe and resourcing for this evaluation."* The evaluation team agreed and did indeed make efforts to reduce the scope of the work. However with hindsight, both the evaluation team and the QAA could have done even more at the inception phase toward getting agreement on a reduced scope of work that would have given the team more time for analysis and consultation, and to address acknowledged key issues such as gender in more depth. Once the Inception Report was agreed, the evaluation felt like a roller-coaster, with the team working flat out to collect a lot of data in a short period of time. It was not possible to go back and try to renegotiate the scope.

**QAA comment:** It is common for the scope of work of an evaluation to grow to barely-manageable or unmanageable proportions in relation to the time and resources available. It is often difficult for an independent evaluation team to agree to a reduced set of evaluation questions at the inception phase, partly because of native human interest and optimism, and partly because most evaluators have a contract that can be terminated at the inception phase for inadequate performance, and are therefore inclined to try to please evaluation stakeholders rather than confront them. In these circumstances, both the evaluation manager (who in this evaluation was recruited late) and the QAA potentially have important roles to play in helping manage expectations and reducing the scope of work to reasonable proportions. A lesson for us

as QAA is that **we should have pushed harder for a reduced and more balanced scope of work in the SUN ICE**. More broadly, making available information on the resources and timeline that was needed for different evaluations would have enabled decision makers to be more realistic about planning future work.

### 3.4 Evaluation Methodology

Evaluating complex global initiatives such as the SUN Movement presents many methodological challenges. The "Note on Approach and Methods" that the evaluation team is providing on its experience with this evaluation will therefore constitute a welcome and valuable input to learning for future evaluations.

The evaluation team applied a mixed methods approach: individual and group interviews, document reviews, an extensive literature review, a visit to the SMS, participation in various SUN meetings, an electronic survey, and eight country case studies. A large number of interviews (over 250) were conducted. The main sources of evidence underlying the conclusions and recommendations appear to have been the interviews, the document and literature reviews, and participation and observation at the various SUN meetings. The electronic survey, which was intended to validate preliminary findings, obtained a response rate of 26 percent (186 respondents out of 711 recipients provided by the SMS, SUN global network coordinators, and team members). Each of the eight country case studies included a one-week visit to each country. These illustrated the diversity of country contexts, and the potential and the difficulties in achieving impacts at country level.

We consider the mixed methods approach adopted by the evaluation team to be fully appropriate and necessary to capture the complexity of the task. Overall we are impressed by the wealth of information collected by the evaluation team in a short period of time, as evidenced by the annexes to the report. However, while recognizing the considerable time pressure, we think that some modifications to the design and implementation of the survey and the country case studies could have enhanced their contributions even more to the evidence base of the evaluation. The QAA made a number of suggestions that (in our view) would have made the survey more inviting and accessible. For example, rather than asking questions about the SUN Movement as a whole, we suggested asking recipients about their own experiences and interactions with the SUN Movement, and then aggregating their responses to get an overall picture.

To triangulate and analyse a large amount of interview data require a systematic and iterative approach for successive testing and validation to avoid undue influence from operational pressures. Overall, we found that the evaluation team did manage to present a credible analysis of the evidence assembled to respond to the evaluation questions. Because of the tight timeline, however, we had to assume that their selection of quotations from the interviews and survey were representative views of the stakeholders consulted.

We note, for example, that the evidence in Annex I on "Issues and Lessons in Multi-Sector Planning for Nutrition" does not appear to demonstrate that a single national plan with all stakeholders (as opposed to something lighter touch, for example temporary multistakeholder



task teams to tackle particular identified nutritional problems) is necessarily the most effective way to make progress. From the examples and the references to political economy, it appears that a number of countries have made major progress on nutrition without a single national nutrition plan (and it also seems from other evidence not presented in the evaluation on the implementation of the Paris Declaration that it is very challenging to get ministries and funders to cohere around such cross-sectoral plans for long periods of time and to adhere to agreements on harmonization and alignment). Therefore, the evaluation highlighted and applauded a move towards accepting that there may not need to be a single comprehensive nutrition plan if a country results framework serves to adequately align the plans of different stakeholders. The redesign of SUN's guidance notes should pay particular attention to the lessons in Annex I.

**QAA Comment:** It is appropriate to use mixed methods to assemble evidence in a comprehensive evaluation of an initiative like SUN, and then to triangulate the evidence derived from the various sources to draw findings and conclusions. Thus, document reviews and stakeholder interviews are standard features of all such evaluations. For additional instruments such as electronic surveys and country visits, which can be costly and time-consuming, it is important to pay particular attention to their design and implementation for them to make a cost-effective contribution to the findings and conclusions.

### 3.5 Evaluation Findings, Conclusions and Recommendations

We find that the findings and conclusions are evidence-based and that the evaluation has established reasonably credible links between the main findings, conclusions, and recommendations as documented in Annex Q of the main evaluation report.

The evaluation *"did not uncover any fatal conceptual weaknesses in SUN's theory of change. . . [but] identified some serious weaknesses and bottlenecks in its implementation which will need to be addressed as the SUN strategy is revised and updated"* (p. 25). While the rushed nature of the SUN ICE and youthfulness of the SUN Movement did not permit an in-depth probe and validation of all the basic assumptions underlying the theory of change, we nevertheless think that the conclusions point to several critical issues for the successful continuation of the SUN initiative. A particularly important one is how SUN can develop its support for successful country-level action.

This having been said, we found that some of the recommendations could have been more specific as to whom they were addressed, and some elements of recommendation 5 on the revised strategy were poorly substantiated by the evidence given in the report, including the historical evidence presented in Annex I, as well as the country case studies. Experience with the Paris Declaration, Accra, and Busan (barely mentioned in the main body of the report, although these appear in the ToR) also raise questions. This is a collective action situation in which stakeholders are participating on a voluntary basis. Incentives for "good behavior" are often outweighed by other considerations — for example domestic politics in donor countries — and enforcing accountability may be very challenging.



We have been given access to the draft summary of responses to the evaluation prior to writing this note. These reflect a basic endorsement of the conclusions and recommendations, although some express disappointment with a perceived lack of clarity and detail.

Many issues regarding the future of SUN remain to be defined during the visioning process as suggested by the evaluation. We think the expectations for the SUN ICE may have been too high to provide concrete suggestions and options given the complexity of the subject matter and the constraints of time and resources. This being a program evaluation, it could not be expected to provide solid evidence, for example, on which types of country-level policies are likely to lead to the most rapid improvements in nutrition. That is a task for researchers, not program-level evaluations.

**QAA Comment:** It is important for evaluators to establish credible links between the main findings, the conclusions, and the recommendations. The SUN ICE did a credible job of this, particularly in Annex Q on the "Matrix of Findings, Conclusions and Recommendations." It is important to challenge, not simply reiterate current ideas of "development correctness" where the evidence warrants.

### 3.6 Monitoring and Evaluation

We found section 5.4 on monitoring, learning, and evaluation to be a strong section based on solid evidence, as well as recommendation 9 on strengthening monitoring, evaluation, and learning. In our experience, almost every single global initiative has struggled with setting up a workable and effective system for monitoring results that meets the expectations of contributors while not imposing unrealistic expectations on recipients. This takes much more time and effort to establish and maintain than initially expected, as evidenced by the efforts of UNAIDS to facilitate the establishment of a global monitoring and evaluation system for HIV (see <https://www.globalhivmeinfo.org/>).

We have also noted the skepticism of some stakeholders with respect to SUN ICE conclusion 4 and recommendation 9 on monitoring and evaluation, and the apparent lack of consensus on how to proceed in this area. There clearly exist tensions and a need to strike a balance between national-level monitoring of nutrition indicators, project-level M&E requirements of donors and other development partners, and a desire to track global progress while facilitating comparisons among countries. Given how slowly nutrition indicators change over time, we find the evaluation's emphasis on monitoring output-level results and country-level policy and institutional outcomes to be appropriate.

We would emphasize that monitoring is important not only to assess progress in implementing individual activities but also to facilitate the next cumulative assessment of the SUN Movement's performance in achieving its objectives — presumably towards the end of the 2016–2020 strategy period. Monitoring is also important to enhance policy dialogue by identifying issues that require policy and institutional responses beyond the scope of individual nutrition interventions.

**QAA Comment:** Putting in place a workable and effective monitoring and evaluation system has proven to be a considerable challenge for almost all global initiatives like SUN. It takes more time and effort than initially expected to strike an appropriate and realistic balance between the expectations and desires of contributors and recipients.