**MEL Community of Practice Session 15: Learning from Failure Documented in Evaluations**

**Description of the Session**

CARE recently completed a meta-analysis to draw lessons from what evaluators reported as challenges, gaps, and lessons learned across 114 project evaluations from 2015-2018. In this session, we will learn more about how the analysis was conducted (using qualitative analysis methods/software) and what it’s telling us about the implications for future work.

This meta-analysis complements the ongoing effort at CARE to focus more on learning from failure—including our [Failing Forward podcast series](https://eur01.safelinks.protection.outlook.com/?url=https%3A%2F%2Fcareinternational.podbean.com%2F&data=02%7C01%7CEcheverria%40careinternational.org%7C22119d047c4741fb585f08d6ee67bdd6%7Ce83233b748134ff5893ff60f400bfcba%7C0%7C0%7C636958526669231389&sdata=F1LtXwS4o6G4B8w6CE4KmK3QRWH9NOnHFAaw4huBJNM%3D&reserved=0) and country, project, and team-specific conversations about learning from what went wrong. This meta-analysis goes beyond case studies that look at challenges in one project or context and learn from the details of that specific experience. It provides a more holistic look at trends across projects, countries, and sectors to see where we may have structural challenges that contribute to the challenges people and projects across CARE face.

**Notes from the Session**

* Today’s focus is on a systemic view of what is going wrong in our projects via our evaluations-how might we need to change our global support
* Reviewed evaluations from 2015-2018 and evaluators told us what went wrong-focus on final evaluations (n=114, English and French)
* Used qualitative coding in MAXQDA
* IMPORTANT: Projects themselves were not failures-there were failures documented within a project. Only about 10% of projects did not meet their goals or closed early (could be considered they failed).
* Large categories of failures
	+ Implementation shows up in almost 9 of 10 projects
	+ MEAL shows up in 73% of cases
	+ Partnership-not involving the right partners
	+ Scale-limited scope of implementation or lack of consideration for sustainability
	+ Gender-encompasses many different aspects, but mentioned as a failure half of the time
	+ HR-almost half of the time
* Top 3 MEAL failures
	+ Bad MEAL planning
	+ Data Quality
	+ Lack of context analysis
* How can we learn from these data? Who can use it? What do we do
	+ Can assess regional breakdowns and compare regional to global
	+ Can break down by outcome area
	+ Can break down by CARE Member Partner
* Limitations and Lessons Learned
	+ **Evaluator bias-**Evaluation reports identified many failures related to evaluation-this may be because this is what they are skilled in and have more experience identifying
	+ **Coding bias-** coding quantitative vs. qualitative evaluations-based on evaluator observation
	+ **Coder bias-**There was disagreement between the 3 coders who were reviewing and coding these reports
	+ **Refine the Codebook-**more than 60 codes to keep track of
		- Created initial codes, then built upon base
		- Gender codes could capture more nuanced information
		- Design codes were not created initially, but came up after review of 10-20 evaluations

**Proposed Actions**

* Use of the categories and codes of this analysis to reinforce the MEAL modules and evaluation tools
* Upon request: provide colleagues with summaries for analysis (e.g. report for LAC, MENA, Norway).
* Provide the main "codes" the evaluation scored against back to the CO and CMP, for discussion/review. Or give country-specific graphics to those countries with more than 5 evaluations, so they can really look at their data.
* Colleagues to review codes and reduce or add relevant codes of interest
* Maybe we should be pulling all management responses into one place, just like we now do with all evaluation reports? Or have annexed them to the actual reports?. Can the evaluation library be re-designed so that evaluations and management response are clear and accessible?
* Greater use of management response to evaluations to document actions, recommendations, and follow-up. Explore practices from different CARE offices. Here a few:
	+ CARE Policy report-using the data to integrate into practice
	+ Head of MEAL in Burundi has done an analysis of trends from external evaluations
	+ Gender: we’ve observed over the last few years that the MEL sub-category of the gender marker is where we see the most projects NOT ticking the box for having met that criteria. So we’ve started there in developing a mini-guide to support project teams in actualizing/operationalizing the MEL criteria. Stay tuned it’s currently with the graphic designer/translators!
	+ GEWEP - management response to the end line reports. Mainly focused on recommendations the team didn't agree with.

**Questions-Comments raised in the session**

* Does ‘implementation’ also capture program design pieces that didn’t work out? So maybe the implementation for that piece was sound, but the actual choice to use that approach or those activities wasn’t a successful/appropriate fit?
* There is a long lag on these indicators...I.e. if we fix MEAL planning today, we won't see change here for three years
* Is there a way MAXQDA software could be available? is there possibilities for capacity building on its use?
* Were there failure themes are sub-themes that came up that you were surprised by (e.g., you hadn’t had a code for them before starting the exercise but added them as you started to see more failure themes around a certain area)?
* Did this include only evaluations performed by external evaluators? and if not, were there differences between those done internally and those done by an external evaluator?
* What about the timing of the final evaluations taking place? Was that taken into consideration and would that make a difference? Most have to take place while the project is still ongoing and there is a lack of some outcomes to be reflected well.
* Evaluations vary a lot in quality. Should we give the same weight to issues raised in really bad evaluations? bad as in poorly done, not critical...
* Another element that may add to evaluator bias is that evaluation designs may not be requesting the evaluator to look at all the dimensions we are looking in this analysis (HR, gender, implementation, MEAL, etc.). It also depends what type of evaluation it is... a "formative" evaluation may look at HR, budget, etc but a summative one may focus on results/MEAL.
* Did you get any sense of what CARE's management response was to evaluations? Our (2008) CARE Evaluations Policy commits us that: "Recommendations from evaluations are to be followed up with action plans, and these action plans, in turn are to be followed up by relevant supervisors. In the case of single-country projects or programmes these will usually be Assistant Country Directors for Programming. In the case of multi-country programmes, the responsibilities for follow-up will be the persons or units responsible for those programmes.”
* If a country team does not have a ACD P role in place...who would have the role of looking at evaluation findings and acting on them?
* What about the donor's specific requirements around the evaluation design, some have templates to be followed as well as the recommendation management plan that needs to follow
* Oxfam does publish management responses in their effectiveness reviews - click on the downloads button on this, for example - <https://policy-practice.oxfam.org.uk/publications/accountability-review-in-kenya-evaluation-of-the-improving-access-to-water-sani-620536>
* On the budgeting side — what resources can we share with design/proposal writers for costing out some of these things (e.g., robust gender or context analysis, KM systems. etc.)? In working on a few bids recently it’s one thing to write it into the narrative, but if we can’t share some helpful guidance on what that piece will cost it’s hard to operationalize it once awarded.