What works for gender-norm change?

Enhancing gender-inclusive agricultural development programming

2018, CIMMYT
Addis Ababa, Ethiopia

What works for gender-norm change?
Enhancing gender-inclusive agricultural development programming

2018
CIMMYT[1]
Addis Ababa

CIMMYT
MR
International Maize and Wheat Improvement Center

Federal Ministry
for Economic Cooperation
and Development

Abstract:
This Research for Development (R4D) Policy Report summarizes findings from a meta-analysis of external evaluations of Ethiopian agricultural development projects, while also highlighting best practices around gender programming. It explores: What works for gender norm change in agricultural development projects? Intent on building a body of evidence, the inclusion criteria required documents to be: external; methodologically rigorous; incorporate gender in the evaluation; and demonstrate social norm change. With this strict inclusion criteria, external evaluations were then assessed for best practices (n=2), while the disqualified evaluations (n=24) were analyzed for areas in need of improvement. The findings show that the CARE Ethiopia office is producing the most rigorous and successful projects around gender norm change. Agricultural development project evaluations are currently a lost opportunity for learning 'what works' for gender norm change. The findings outline both what to do and highlight what to avoid in undertaking gender transformative development.

TABLE OF CONTENTS

CIMMYT
International Maize and Wheat Improvement Center

Federal Ministry
for Economic Cooperation
and Development

1.   EXECUTIVE SUMMARY

Gender equality has proven important for sustainability, to decrease unintended consequences, and to improve agricultural productivity[2]. Social norms lie behind gender inequality in endowments and agriculture productivity but have not been adequately studied. This R4D policy report enables evidence-based gender programming, project management, and implementation recommendations. It provides the CGIAR system and the broader agricultural development sector with guidelines for enhancing gender-inclusive project design and subsequent evaluations to iteratively strengthen the evidence base around gender norm change. This research was motivated by an interest to better understand the available evidence regarding the gendered dimensions of agricultural development projects. The focus on gender outcomes purposefully pushes project boundaries to the end-users of agricultural technologies, helping shift from a lab-centered design to a human-centered design practice.

This report has identified areas in need of improvement based upon the 24 disqualified documents and best practices from two final performance evaluations: the Women's Empowerment: Improving Resilience, Income and Food Security (WE-RISE) program implemented by CARE Ethiopia and the Graduation with Resilience to Achieve Sustainable Development (GRAD) project implemented by CARE Ethiopia and consortium. Both projects met the inclusion criteria of rigorous methods, inclusion of gender, and evidence of social norm change. They had women in their Theory of Change (TOC) and incorporated women at the highest level of the project design, and women were involved with the strategies and indicators that addressed the needs of women and the social relationships in which they live.

The meta-review findings converged around successful moments of gender-sensitive project design and implementation. Regarding project design, multi-level, multi-component projects which directly target gender transformation or women within their communities achieve the best results, and project Theories of Change should incorporate and address gender. Regarding evaluations, collecting data on women and gender relations is crucial for learning. Focus should be placed on outcome and impact indicators and not only output. Regarding methods, mixed methods are important, with qualitative research offering key insights into the contextualized aspects that characterize women's lives, relationships, and communities. Regarding programming, programs should contextualize gender relations, engage a variety of community stakeholders, including traditional leaders, men and boys. Regarding the timeline, projects should invest in longer-term projects between five and ten years or not expect social norms to change within a shorter project. Other lessons include:

- creating community-based knowledge linkages between women and their relevant legal protections;
- targeting and improving women's access to finances and buttress with literacy and numeracy capacity-building where needed;
- facilitating household and community discussions on gender relations;

---

[2] FAO (Food and Agriculture Organization of the United Nations). (2011). State of food and agriculture 2010–2011: Women in agriculture - closing the gender gap for development. Rome: FAO; World Bank. (2011). World Development Report: Gender Equality and development. Washington, DC; IMF (2011). Inequality and Economic Outcomes in sub-Saharan Africa, Chapter 3 in Regional Economic Outlook: sub-Saharan Africa).

- designing and implementing multi-component programs which incorporate gender norm change;
- targeting gender relations and work with husband/wife teams;
- facilitating discussions on gender-related issues and women's empowerment;
- pairing support for women's financial contributions to the household with gender sensitization;
- recognizing that male and female-headed households are likely to require different programming;
- placing women at the center of project design, as does CARE Ethiopia, directly targeting women or undertaking differential investment to ensure equal outcomes for women, gathering sex-disaggregated data, and ensuring that evaluations capture gender dynamics.

There is rigorous evidence from Ethiopian agricultural development projects to encourage incorporating gender norm change at the highest level of project design, pairing income-generating activities or savings with community discussions and training around gender relations, allotting differential budget and resources to engage women as participants and decrease their opportunity cost, and to ensure that male and female-headed households receive differential programming or consideration.

The disqualified evaluations largely fail to capture gender-related data or mainstream gender in evaluations. Agricultural development project evaluations are currently a lost opportunity for learning 'what works' for gender norm change. The findings outline both what to do and highlight what to avoid in undertaking gender transformative development. In this case, what we are *not* learning appears to be as important as what we *are* learning in order to enhance project efficacy for all beneficiaries. The findings identify what works to accomplish and capture gender norm change, and areas for improvement from eliminated documents. A substantive body of learning is found in internal documentation where public access is restricted, and documentation is considered less credible.

In sum, the evidence base for gender norm change is currently undermined, rather than enhanced, by external evaluations. This represents a loss of knowledge around gender transformative agricultural development projects and reproduces women's poor engagement with agricultural development programming in Ethiopia. We conclude with guidelines for enhancing gender-inclusive development projects and building the evidence base.

CIMMYT MR
International Maize and Wheat Improvement Center

Federal Ministry
for Economic Cooperation
and Development

2.   FIGURES AND TABLES

CIMMYT
International Maize and Wheat Improvement Center

Federal Ministry
for Economic Cooperation
and Development

3. ACRONYMS

ACDI/VOCA	Agricultural Cooperative Development International/Volunteers in Overseas Cooperative Assistance
AMDe	Agribusiness Market Development
CARE	Cooperative for Assistance and Relief Everywhere
CFIW	Chronically Food Insecure Women
CIMMYT	International Maize and Wheat Improvement Center
DFID	Department for International Development
FEED	Feed Enhancement for Ethiopian Development
GoE	Government of Ethiopia
GRAD	Graduation with Resilience to Achieve Sustainable Development
IFAD	International Fund for Agricultural Development
IFPRI	International Food Policy Research Institute
IGA	Income Generating Activities
LNWB	Leave No Woman Behind
NGO	Non-governmental Organization
OECD	Organization for Economic Cooperation and Development
PSNP	Productive Safety Net Program
R4D	Research for Development
SAA	Social Action and Analysis
SNC	Social norm change
SOW	Statement of Work
TOC	Theory of Change
UNFPA	United Nations Population Fund's
USAID	United States Agency for International Development
VESA	Village Economic and Social Association
VSLA	Village Savings and Loan Association
WE-RISE	Women's Empowerment: Improving Resilience, Income and Food Security
WEI	Women's Empowerment Index
WEIA	Women's Empowerment in Agriculture Index
Woreda	Local Districts

CIMMYT
International Maize and Wheat Improvement Center

Federal Ministry
for Economic Cooperation
and Development

4.   INTRODUCTION

Smallholder agriculture is a family affair. Yet men are often perceived of and seen as 'the farmers'. Men have access to technical training and knowledge, make decisions, and control resources, while women look after the children. The reality is quite different with women often contributing significantly to agricultural labor, in addition to completing all childcare and household labor. In a global economy where migration and sicknesses mean the head of the household cannot always be a man, there is a need to recognize and enhance women's roles in agricultural development. Technologies are fast changing, innovation is constant, and markets are volatile. Keeping up requires considerable clout and the effort of whole families, which is why the gendered division of labor is more porous. Aiding communities and individuals to embrace such changes through development projects is the role of implementing partners. However, the evidence below suggests that many projects and partners do not know how to assist with the social norm changes necessary to ensure women can benefit from development equitably.

> **Social norms** refer to how group beliefs outline appropriate behavior and expected actions for members. In this manner, social norms refer to both the ideas held by and the actions they elicit in individuals, and how individuals are sanctioned by their surrounding community if they break a social norm (Boudet et al., 2012; Mackie et al., 2015). Individuals engender and conform to norms due to the expectation of sanctions by their friends, family, neighbors, and community members.
>
> **Gender norms** typically operate on an ideological level and an individual level. In other words, gender norms represent perspectives on what *gender relations* 'should be like' and how individuals of particular genders 'should behave' through their *gender role* (Marcus, 2014). Changing social norms then requires altering what is, and is not, acceptable at a community level and the presence of individuals willing and able to 'break' social norms through their actions. Gender norms are often connected to broader social norms.

In a 20-country qualitative study completed by the World Bank (Boudet et al., 2012), gender norms were found to have just as much impact on the actions and opportunities of men and women as the conditions of their communities and countries. While development has thus far focused on ensuring opportunities are available to individuals in developing societies, social norms can act as barriers *or* catalysts to individual action. Furthermore, social norm change is found to be most effective when multiple levels are targeted in an integrated and sustained manner: individual, relationship, community, and societal levels (DFID, 2012; Haider, 2017). Recent reviews of gender norm research (Marcus and Page, 2014) find that social norm change is best buttressed by changes in macro-level, contextual factors such as the economy, education, political and social change, as well as exposure to new ideas. Yet an understanding of country-level and project-level efficacy of gender norm change is lacking.

In the agricultural development sector, there is less research on gender norm change than in sectors such as health and education. The Food and Agriculture Organization (FAO, 2011) argues that, due to the social roles and constraints that women farmers face, they are unable to achieve outcomes equal to their male counterparts. See Table 1 for examples of differential outcomes for men and women due to social norms (WE-RISE example), how social norms embed in 'ownership' over particular crops (AMDe example), and how gender norms serve as a barrier to participation in agricultural development projects (FEED II). Given the dearth of research on this topic, learning from evaluations about successful project interventions is a prudent starting place.

CIMMYT MR
International Maize and Wheat Improvement Center

Federal Ministry
for Economic Cooperation
and Development

Table 1. Why should we care about gender: Examples of how and why gender norms are a key component to agricultural development outcomes

What do we mean when we talk about gender? Differential impacts for men and women? Women's empowerment? Failure to include women? Here are three examples from recent projects in Ethiopia:

In WE-RISE, implemented by CARE Ethiopia, chronically food-insecure women were successfully engaged, yet differential outcomes were observed in beehive and honey-processing activities (TANGO 2016). Select households were provided beehives and trained in improved honey processing practices. Honey processing is traditionally a male activity in Ethiopia and although the project did successfully engage women, men overwhelmingly took on leadership roles. Men comprised 66% of the leadership committee, while being only 4% of the members. Data such as this demonstrate how difficult altering gendered practices can be, despite investment in women. One informant stated, "We can change the name but maybe not the attitude" with respect to gender (28).

The AMDe project, implemented by ACDI/VOCA, followed a USAID investment strategy to target smallholder farmer poverty reduction through value chain enhancement activities, with a target for export dollars. At the same time, women were meant to be specifically targeted and engaged as beneficiaries. However, the majority of the crops the project chose to address are commonly for export and dominated by men. As a result, women were marginalized through the selection of the export value chains. Here, the theory of change and accompanying activities, although with a stated interest in women, were unable to reach women due to local gender norms and practices (Tufts University 2015b).

The FEED II project, implemented by ACDI/VOCA, had a substantive stand-alone section on gender and positioned gender as a cross-cutting issue (ICOS Consulting PLC, 2016). This evaluation was eliminated because there was no evidence of social norm change, yet norms are recognized as a barrier to women's participation in the program. Women's economic power, decision-making with their husbands, differential land ownership, household workload, and neighbors' lack of respect for women all undermined their participation:

> "**Husbands have remained dominant** and, as a result, women have little role in important decision-making processes such as selling fattened animals, entering into loan agreements and use of income generated by the household… **Neighbors consider women's participation in the project a waste of time. This attitude of neighbors has also challenged knowledge and skills transfer as their followers, who are often their neighbors, don't take them seriously.** This may suggest the importance of wider community awareness. Limited economic capacity of women is also another factor that hindered their participation in the project as **they have weak resource ownership and financial capacity to buy necessary agricultural equipment and improved livestock breeds**… **Male dominated culture seems also to have influenced development agents' decision in selecting beneficiaries of the project as they sometimes unintentionally incline towards men.** The place in which the training is delivered has also challenged the participation of women in project activities. [Emphasis added] (2016:34).

These selections highlight the different ways in which norms interact with project objectives, targets, and programming, ultimately minimizing women's participation and outcomes. In the Ethiopian context, norms typically result in women's decreased participation, failure to implement training, or to sustain adoption of new technologies.

CIMMYT
International Maize and Wheat Improvement Center

Federal Ministry
for Economic Cooperation
and Development

Several documents were disqualified because they did not mention women or gender in any substantive way. This was common for evaluations using the Organization for Economic Cooperation and Development (OECD) DAC Evaluation Guidelines[3] which assess projects based on relevance, effectiveness, efficiency, impact, and sustainability. Several projects claimed an impact on gender norms, and yet did not offer empirical proof. If evaluations failed in any of these areas, they were disqualified.

Due to the small sample which meets the inclusion criteria, this review elicits lessons about what not to do. First, qualitative data was often presented as anecdotal 'cases', 'stories' or 'selected beneficiaries' without an exploration of how this individual story may, or may not, be representative, suggesting false generalizations. Furthermore, many evaluations had underdeveloped data to support conclusions. This is likely a result of varying capacities to conduct evaluations, familiarity with research-based claims, and cultural norms around writing style. Ultimately, 'methodological rigor' served as a quality check to international standards of evaluation. In short, the review also creates a body of evidence which critiques the broader landscape of evaluation as a learning device, while including 'where not to step' based upon the shortcomings of the disqualified evaluation documents[4]. Furthermore, it offers an empirical analysis of the breakpoints of how evaluations fail to address women and gender concerns.

This research offers a snapshot in time of gender-norm documentation and learning in Ethiopian agricultural development. It is nested within greater trends by development agencies to address the social context in which women live, ensuring they benefit from development projects as well as catalyze sustainable improvements in food security and poverty reduction.

The reports structure is as follows: methodology and explanation of the inclusion criteria; a deep dive into the two selected evaluations to understand best practices to build the evidence base around capturing gender-norm change in evaluations, and what works for gender-norm change in project design; areas for improvement from the eliminated documents covering ways to more strongly present qualitative data, to ensure the OECD-DAC evaluation criteria are gender sensitive, and how to build gender into an evaluation to ensure lessons are captured (even in projects that do not have a strong impact on gender relations). We conclude with guidelines for enhancing gender-inclusive development projects and building the evidence base.

5.   METHODOLOGY

The sampling frame for evaluation documents was constructed during the 2017 CIMMYT stakeholder interviews and updated during 2018. Stakeholders were originally sampled from the Directory of Development Organizations of Ethiopia and the Ethiopia Network for Gender Equality in the Agriculture Sector, as well as identified through snowball sampling[5]. All stakeholders were asked if they had a current program in the agriculture sector with a gender component and if they were willing to be interviewed. Stakeholders self-selected to participate and then nominated a contact person for the interview. During the interview, stakeholders were asked to describe the best program they had seen on gender equality in agriculture, if they had any evaluations to share, and to provide

---

[3] See http://www.oecd.org/development/evaluation/dcdndep/41612905.pdf for additional information regarding DAC Evaluation Guidelines.

[4] The sampling process identified for this review is similar in process to the IFAD Gender Equality and Women's Empowerment Assessment (2017). However, IFAD evaluators explicitly applied a "less stringent filter of criteria" and chose to drop the impact assessment or counterfactual evidence requirement in order to raise their sample size. Since the goal of this R4D project was to identify programmatic interventions which have met with success in Ethiopian agricultural development, it was determined to maintain a strict focus on exemplary evaluations while identifying areas in need of improvement from disqualified evaluations or document types.

[5] Gaining additional recommendations from interviewed stakeholders.

organizational documentation around what works for gender-norm change. The resulting initial sample of 26 documents serves as the focus of this report.[6] Internet research during 2018 was used to determine if a more recent evaluation was available.

The stakeholder-provided documents were then analyzed and categorized into three types: evaluations, learning documents, and gender assessments[7]. Documents were eliminated from consideration if an updated evaluation is present, no original data collection had taken place, multi-country aggregate data was presented rather than disaggregated Ethiopia-specific data, or the document took on a promotional tone. The first inclusion round required the document be an external evaluation. See Appendix B for first round results. In order to develop a dataset appropriate to address 'what works for gender norm change?' additional inclusion criteria became necessary to ensure rigor.

To qualify during the second round of inclusion, an external evaluation must demonstrate methodological rigor, incorporation of gender in the evaluation, and demonstrable gender norm change. Since the research question is focused on the outcome variable of gender norm change, evaluations which *claimed* projects addressed gender norms and projects *which did not aim* to address these norms were disqualified. Through these two rounds of qualifying criteria, the final sample focuses strongly on 'what works for gender norm change?' See Appendix C for second round results. By selecting for these criteria, this meta-analysis analyzes and synthesizes existing data on the phenomenon of gendered social norm transformation from exemplary external evaluation sources.

Table 2. The second round of inclusion/exclusion criteria	
Rigor of methods	
Quantitative sample	What was the quantitative sampling strategy? Did strata include women or female-headed households? Was it statistically representative? Are sampling methods and survey procedures explained?
Qualitative treatment	How was their qualitative sample taken? Did it involve women? Is qualitative data presented in a systematic and rigorous manner?
Evidence for findings	Is their analysis and presentation of evidence credible? Are the results disaggregated by sex? Are evaluative claims supported by empirical data?
Gender included in evaluation	
None	No substantive mention or exploration of gender or women in the document
Section on gender	Is there a section in the evaluation on gender? Is this the only place where women and/or gender is discussed?
Woven throughout	Is the differential impact of the project on men and women continuously disaggregated throughout the document? Is terminology gender sensitive (e.g. "women" used as an

---

[6] Not all stakeholders had documents they were willing or able to share and some stakeholders shared more than one document. Approximately ten documents were immediately eliminated because they were promotional in nature.

[7] By sharing a variety of document types, organizations demonstrate there is not a consistent nor single location for capturing learning around gender.

CIMMYT MR
International Maize and Wheat Improvement Center

Federal Ministry for Economic Cooperation and Development

	adjective to farmers only in the gender section)? Are results explored by different head-of-household type?
Social-norm change (outcome)	
None	No mention of social norms in the document
Claimed	Did the project claim to change social norms but struggle to provide credible evidence of change that can be linked to the program?
Present	Did the project change social norms and provide credible evidence of social-norm changes that are linked to the program?

6.   DEEP DIVE INTO BEST PRACTICE EVALUATIONS

Both of the final evaluations selected for a deep-dive analysis were designed by CARE Ethiopia[8]. As an organization, CARE takes an explicit long-term approach to poverty reduction and develops projects which assist three target populations which they have identified at an organizational level: pastoralist girls, chronically food-insecure rural women, and poor young girls living in cities and on the outskirts of urban areas. The two final performance evaluations which met the criteria of rigorous methods, inclusion of gender, and evidence of social-norm change include the Women's Empowerment: Improving Resilience, Income and Food Security (WE-RISE) program implemented by CARE Ethiopia and the Graduation with Resilience to Achieve Sustainable Development (GRAD) project implemented by CARE Ethiopia and consortium. Both projects had women in their Theory of Change (TOC) and incorporated women at the highest level of project design, and into strategies and indicators that addressed the needs of women and the social relationships in which they live.

However, two other evaluations which directly targeted women and gender relations in their design failed to qualify: CARE Ethiopia's ABDISHE Final Performance Evaluation (TZBMC, 2016) and the United Nations Population Fund's (UNFPA) Leave No Woman Behind (LNWB) Final Evaluation (Kabuchu, 2013). While the methods were fairly rigorous, ABDISHE was eliminated because the evaluation failed to report statistical significance findings, despite surveying a sample size capable of assessing results with 95% confidence. Rather than noting the changes were not significant, there was no comment. The LNWB project successfully envisioned a multi-component grassroots program which both targeted women beneficiaries and addressed multiple areas of harmful traditional practices, income generation, literacy and numeracy. However, (as noted several times throughout the evaluation report) the project monitoring and evaluation system was underdeveloped, baseline data was still being collected during the mid-term evaluation and, therefore, the final evaluation was unable to provide any quantitative evidence of social-norm change.

6.1   "WE-RISE" FINAL PERFORMANCE EVALUATION: Case Study

The WE-RISE Evaluation (TANGO, 2016) represents best practice lessons around gender in evaluations for two core reasons: (1) creating a gendered project *in its design*; and (2) by including programmatic details that others can learn from.

---

[8] CARE began working in Ethiopia in 1984 during the severe drought that claimed nearly one million people. CARE's work addresses the root causes of poverty and vulnerability in the areas of livelihoods and food security, sexual and reproductive health, HIV/AIDS, education, governance, water and sanitation, and emergency preparedness and response. CARE Ethiopia Country Factsheet 2010.

CIMMYT
International Maize and Wheat Improvement Center

Federal Ministry
for Economic Cooperation
and Development

6.1.1 PROJECT OBJECTIVES AND THEORY OF CHANGE

At the design level, WE-RISE is based on a gender-sensitive theory of change. The WE-RISE hypothesis states that food security increases when women have greater capacity, are served by gender-sensitive institutions, and agriculture-related markets are more competitive. This can be accomplished through advances at multiple, reinforcing levels, including: women's agency, social relations, and supportive community and governance structures.

Table 3. WE-RISE Objective (Obj) and Theory of Change (ToC)	
Obj	CARE's program, Women's Empowerment: Improving Resilience, Income and Food Security (WE-RISE), focuses on improving household food security and resilience by empowering women, particularly through increased agricultural productivity…WE-RISE is designed to improve the quality of life for chronically food-insecure rural women (CFIRW). The program seeks to increase agricultural productivity through income-generating activities, support environments promoting women's rights and gender-sensitive agricultural programming, and increase institutional capacity for improved gender-equitable programming at the global level[9].
ToC	CARE's previous work on the Women's Empowerment Strategic Impact Inquiry provides the basis of the WE-RISE Theory of Change (ToC), which includes three domains of change: a) women's agency (i.e., skills, knowledge and aspirations), b) formal and informal structures, and c) social relations that women engage in on a daily basis (i.e., cultural and social norms and attitudes)… Thus, the program theorized that marginalized, chronically food insecure rural women would be more productive and their families more food secure when: women have increased capacity (skills, knowledge, resources), capabilities (confidence, bargaining power, collective voice), and support; local governance and institutions have/implement gender-sensitive policies and programming that are responsive to the rights and needs of poor women farmers; agricultural service, value chain, and market environments of relevance to women are more competitive, gender-inclusive and environmentally sustainable[10].

6.1.2 EVALUATION METHODS AND RIGOR

The quantitative portion of the evaluation was designed as a randomly sampled 'beneficiary-based' survey, meant to capture the same participating households from baseline to endline. The project originally targeted 15,441 households but worked instead with around 11,000 households, effectively dropping entire *kebeles* from the project and subsequent final evaluation. Due to this, there were many households on the endline sampling frame that should have been removed prior to selection, rather than counted as attrition or non-response. The baseline reached 921 households, and the desired endline size was 890. Accounting for project participation, the endline sample size target was 578, and 551 households were reached. Household demographics at baseline and endline were assessed for similarity. While the number of widows and divorced women increased, trending to a greater number of female-headed households at the endline, yet there are slightly more female-headed households in the baseline. The authors note that because Village Savings and Loans Associations (VSLA) were the access point to project resources and women were the target beneficiary, households had an incentive to self-identify as female headed.

The qualitative component was purposively selected for diversity along project activities, geography, topography, and agro-ecological area. The qualitative study was designed to provide complementary data on women's participation in income generation, VSLA activities, and agricultural production.

---

[9] TANGO (2016:vii)
[10] TANGO (2016:2)

CIMMYT
International Maize and Wheat Improvement Center

Federal Ministry
for Economic Cooperation
and Development

Unfortunately, annexes were not provided with additional detail on the methodology and interview guides.

The evaluation report-built credibility through strong descriptions of metrics, including a laymen's definition of the metric and disclosure of desired directionality, an acknowledgement of when baseline data did not exist, conflicting findings across triangulated data sources which were disclosed and explored, and relatively detailed descriptions of programmatic interventions (see below). They acknowledge differences and aim to explore them. Survey data was contextualized with local practices. For example, capturing poor women's access to high protein foods, such as meat, eggs, and fish, may be underestimated by data collection processes which require skipping all consumption questions if respondents report that the previous day was a festival, wedding, or funeral[11]. However, these events are primary avenues for poor women to eat protein-rich foods. Understandings and explanations such as these, build credibility for the evaluation.

6.1.3 FINDINGS

The evaluation itself breaks down the project objective of food security, resilience, and income into indicators commonly used, yet keeps a focus on understanding women's outcomes, both within male-headed households and female-headed households. When analyzing performance against the highest-level objective, the evaluation primarily utilizes 'gender neutral' measurements, while considering head of household types (male or female): food security is measured through dietary diversity and intra-household access; income is measured through mean asset and income diversity, livelihood resilience is measured through Consumption Coping Strategies and savings; and lastly, women's empowerment is measured through an adapted Women's Empowerment in Agriculture Index. By ensuring that respondents are both men and women, the evaluation is able to provide a gendered analysis of rigorous food security measures, successfully building a body of evidence around the engagement of women in agricultural development projects.

When comparing baseline and endline survey results, using random samples of projects' beneficiaries, nearly all measures found statistically-significant differences in beneficiaries' lives. See Table 4. This is a strong accomplishment considering the target beneficiary (chronically food-insecure women) is one of the hardest to reach and achieve outcomes for.

Table 4. Performance Outcomes for WE-RISE		
Outcome	Indicators	Statistically significant
Food security	Dietary diversity Intra-household access	Yes Yes
Income	Mean asset Income diversity	Yes Yes
Livelihood resilience	Consumption Coping Strategies Savings	Yes Yes
Women's empowerment	Women's Empowerment in Agriculture Index	Mixed results

---

[11] This is common procedure for Food and Nutrition Technical Assistance (FANTA) recommended data collection tools around Household Dietary Diversity. More information and tools available via FANTA III.

CIMMYT MR
International Maize and Wheat Improvement Center

Federal Ministry
for Economic Cooperation
and Development

6.1.4   WHAT WORKS FOR SOCIAL NORM CHANGE?

In total, the activities described in the evaluation demonstrate WE-RISE tackled gender norms both directly through paralegals and Social Action and Analysis (SAA) and indirectly through targeting women-only or women-skewed programming which is input-based or access to loans (including giving goats, sheep, and chickens to select households, while providing supporting training). Here, the programmatic interventions which more directly address gender relations or social-norm change are described:

Create community-based knowledge linkages between women and their relevant legal protections
The WE-RISE team created and trained 26 paralegal groups, one for each *kebele*. Paralegals received training in how to facilitate discussions around harmful practices, such as early marriage, rape, polygamy, female genital mutilation, and gender-based violence (GBV). In doing so, paralegals publicized women's rights to local women. Focus group discussion data found that paralegals were credited with translating government policies down to the local level through a variety of activities not uncovered through survey data: decreased conflict between husbands and wives and GBV, assistance to pregnant women, and awareness-raising around harmful practices and laws protecting women's rights. It is unclear how frequently paralegals facilitated discussions or where they took place, yet this intervention was highly appreciated by the women focus group participants.

Target and improve women's access to finances and buttress with literacy and numeracy capacity-building where needed
Village Savings and Loans Associations (VSLAs) are informal savings and loans groups initiated by the WE-RISE program. They specifically targeted women for membership and provide an alternative to high-interest moneylenders. VSLAs appear to foster a savings culture and train women in business skills and leadership. Most information in the evaluation report describes the reported impact of VSLAs rather than the specific model or intervention. In other words, programmatic information about how to run a VSLA, such as how often they meet or the management model, is absent. The only detail included is that the moneybox is secured in a member's house. However, VSLA members ranked VSLA formation and participation as the most impactful activity of WE-RISE to their lives. Where leadership training of women members was successful, VSLAs are operating well[12]. Yet, in other areas, women lacked the needed foundational skills in literacy and numeracy in order to properly manage the group savings and the VSLA suffered.

Facilitate household and community discussions on gender relations
Social Action and Analysis (SAA) was the key programmatic element which directly addressed gender relations by getting WE-RISE participants to discuss gender and patriarchal norms, becoming more aware and transforming gender relations in the process. SAA is an approach used and developed by CARE.

An SAA group was created in each *kebele* and led by a core group. The local implementing partner in each area created a group of 52 people representing a stakeholder cross-section of society, such as police commander, *kebele* manager, health extension worker, homemaker, etc. This group of 52 then selects approximately 14 core group members as "model gatekeepers." The role of the gatekeepers is to promote "progressive attitudes, behavior and practice concerning gender relations, gender roles, and ensconced negative cultural and social practices" (TANGO, 2016:29). This group then utilizes an SAA manual to select discussion topics. Two topics are selected per month by the

---

[12] Details regarding the leadership training and how or why it may have been more successful in some locations was not detailed in the evaluation.

CIMMYT
International Maize and Wheat Improvement Center

Federal Ministry
for Economic Cooperation
and Development

core group, which then forms sub-groups to talk to men and women in the community. Some of the topics include women's participation in public, women and men walking side by side, household division of labor, girls' education, inheritance, female abduction for marriage, and female genital cutting. Focus-group discussion data demonstrates that the topics discussed are akin to a 'revolution' in Sidama culture and that, while change is slow, even talking about these topics is progress. Focus-group data shows that members are aware there is a difference between discussion and practice, with the practice lagging behind.

The evaluation noted that this activity was implemented relatively late within the project cycle and, therefore, did not have enough time to demonstrate effectiveness. Although this activity was ranked lower by participants, one explanation may be that relatively few (~10% of total WE-RISE VSLA beneficiaries) participated in an SAA group. The evaluation stated that because Sidama culture is very conservative around gender, a longer project timeline is needed for cultural change activities. The evaluation report (TANGO, 2016:30) noted that women, even when selected, were absent from meetings since "women continue to find it much more difficult to obtain sufficient time to participate and contribute to the discussion than do men," meaning that women's participation in activities is undermined by their duties elsewhere.

6.1.5 BEST PRACTICES TO BUILD THE EVIDENCE BASE

The following represent the best practices within the WE-RISE evaluation which help to build an evidence base around gender-norm change.

The Theory of Change should explicitly address women or gender relations
By building the TOC around addressing CFIW, the project monitoring and evaluation system and the evaluation survey were designed to measure gender-specific outcomes around food security, resilience, and income as well as cultural beliefs about women. This attention to women and gender norms is further manifested in the high-level objectives and the project outcomes. This ensures data is collected around gendered outcomes, resulting in a useful gender-sensitive project dataset.

Place of women or gender relations in the highest levels of the project's monitoring and evaluation system
The strong evaluation focus on CFIW can be further seen in the indicators used to assess project performance at the outcome level.

Table 5. WE-RISE Outcomes and Indicators		
	Outcome	Associated Indicators
1	CFIRW have increased household productive assets and resources and control over them, and are more resilient to climate shocks.	▪ Women's access to and control of loans ▪ Diversification of sources of income ▪ Agricultural production, diversification, and improved practices ▪ Access to agricultural inputs and markets
2	Formal and informal institutions are more responsive to women's priorities and accountable to upholding their rights.	▪ Women's access to agricultural financial and extension services ▪ Women's participation in formal and informal groups ▪ Self-confidence speaking and expressing public opinions
3	Cultural and social norms and attitudes better support the individual and collective aspirations and improved opportunities for CFIRW.	▪ Women's control of income, expenditures, and asset decisions ▪ Women's control of health care and reproductive health decisions ▪ Attitudes about gender equality in family life ▪ Women's mobility

CIMMYT MR
International Maize and Wheat Improvement Center

Federal Ministry for Economic Cooperation and Development

Gender programming is an experiment

While the evaluation found mixed results around women's empowerment, without measuring and including the module as part of their endline survey, there would be no learning around women's empowerment. In addition to assessing the technical outcome areas, the external evaluation team created and conducted a Women's Empowerment Index (WEI)[13], modeled after the Women's Empowerment in Agriculture Index (WEIA)[14] – a survey which assesses women's empowerment along five domains—production, resources, income, leadership and community, and autonomy—and a gender parity index[15]. See Figure 1 below, from the WE-RISE evaluation (19), which demonstrates that the project has *statistically-significant increases and decreases* in women's empowerment. Out of 12 indicators included in the WEI, statistically-significant improvement occurred in only three indicators, at p-values of .01, .05, and 0.1 (See highlighted indicators in Figure 1). Without this detailed data collection, learning what works and what does not work, would be lost.

Figure 1: WE-RISE Presentation of Women's Empowerment Index Results

Table 16: Domains of empowerment

Domain	Indicator	Point Estimate BL	Point Estimate EL		Sample Size BL	Sample Size EL
Production	With decision-making input for all HH productive decision domains	66.5	71.2	*	501	548
	With autonomy in one or more HH production domains	38.3	28.8	***	501	548
Resources	With sole or joint ownership of 75% of household assets	78.4	62.0	***	476	548
	With sole or joint control over purchase or sale of 75% household assets	79.1	68.3	***	478	548
	With access to and decisions on credit	63.4	67.0		257	303
Income	With control over household income and expenditures in 50% of HH decision-making domains	55.5	71.7	***	510	551
Leadership & community	Participating in formal and informal groups	96.0	89.2	***	525	499
	Confident speaking about gender and other community issues at the local level	82.6	70.7	***	534	543
	Demonstrating political participation	^	^		^	^
	Who express self-confidence in 5 of 7 statements	50.6	52.9		534	534
Autonomy	Satisfied with the amount of time available for leisure activities	64.4	71.3	**	533	543
	Achieving a mobility score of 16 or greater	46.7	46.8		533	543
	Expressing attitudes that support gender equitable roles in family life	74.1	68.7	**	534	543

Statistically different from baseline at the 10% (*), 5%(**) or 1%(***) levels.

What works for male-headed households, may not work for female-headed households

When looking at the overall empowerment of women, there is only a slight level of empowerment significant at the p=0.05 level. However, when disaggregating by household type, women within male-headed households' empowerment increased by 25% at the p=0.01 level while women within female-headed households decreased (significant at p=0.05).

---

[13] The Women's Empowerment Index (WEI) is similar to the Women's Empowerment in Agriculture (WEIA) index, yet with modifications, including dropping a political participation question due to sensitivity in the Ethiopian context and an indicator for workload.

[14] The WEIA is an index created by IFPRI, USAID, and Oxford Poverty and Human Development Initiative (OPHI) to better capture agricultural development project impacts on women's empowerment, agency, and inclusion. For more information, see IFPRI's resource page on WEIA.

[15] The gender parity index compares the differences in answers between a man and a woman in a household. Female-headed households are not measured using the WEI because the Gender Parity Index is unable to be completed.

CIMMYT
International Maize and Wheat Improvement Center

Federal Ministry for Economic Cooperation and Development

Tip for improvement: Qualitative data from focus group discussions can provide meaningful insights. However, this data is often presented as a single voice rather than as representative of the group discussion. For example, the WE-RISE evaluation cited a single man in a focus group discussion as saying there is a local oral history of the Sidama queen Fura. This history continues to engender fear of women's leadership in Sidama society. This would serve as stronger data if it were contextualized as a shared concern among the men's focus group. If the data is from a key informant, then the particular expert perspective and why (or why not) it carries weight should be explained.

Households respond to programming in dynamic ways
For male-headed households, the largest gains from the WE-RISE program are in women's participation in income and household expenditures. Although men retain control over these decisions, women are now participating in the decision-making process (as seen through a statistically-significant increase from 36% to 66% during the life of the project). These same men reported a large decrease in decision-making autonomy (from 52% to 28%), meaning 72% of men feel less autonomous than at baseline. However, this finding was not mirrored by women. Instead, women reported a decline in autonomy from 8% to 6% at endline. Fewer women feel they have sole or joint ownership over assets, while men increased their feelings of sole or joint ownership. The evaluation report (TANGO, 2016:22) states, "These findings may be an indication of women's frustrations that through the project, men, talk about the desirability for increased joint household production and asset decision-making, the reality of changing household production and decision-making has yet to match the talk." The findings acknowledge that shifting gender dynamics to equitable outcomes, male backlash and women's frustrations with the project are a reality.

Do not assume women believe in gender equality
Men and women alike expressed hostile views towards gender equality and women's empowerment, even at the endline evaluation for WE-RISE. Project designers need to understand that no man or women exists outside of gender socialization and even women may hold themselves or other women back in order to maintain or conform to expected social norms.

Ensure women are half, or more, of targeted beneficiaries
As a result of the direct targeting of women beneficiaries and inclusion of gender-norm change at the highest levels, the evaluation did not require a specific "gender" question in order to assess the gender dimensions of the project. It was inherently gender-sensitive while completing a rigorous evaluation. Throughout the technical areas of the evaluation regarding accomplishment of indicator targets and project efficacy, women's outcomes were included in the results. This is not demonstrative of social-norm change, but a gender-sensitive analysis using male- and female-headed households, as well as driving up the percentage of female beneficiaries by explicitly targeting chronically-food-insecure women (CFIW).

Consider organizational dimensions, staff ratios, and staff understanding about gender
While gender is commonly presented as a 'cross-cutting issue' and typically siloed into a single section, since the project and corresponding evaluation report mainstream gender in all manners, the cross-cutting issues section addresses organizational concerns around staff and project management. Here, project management concerns were raised around the burn rate of project resources and the implementing partner staff reported knowing 'nothing' about gender before being trained to implement the project. In this manner, staff capacity-building around gender was a core component of success. Yet, his important aspect would have gone unrecognized without a gender-sensitive evaluation team.

CIMMYT
International Maize and Wheat Improvement Center

Federal Ministry
for Economic Cooperation
and Development

Use participatory methods that allow beneficiaries and broader stakeholders to share their perspectives of 'what works'

Stakeholder Ranking, a participatory method which allows different stakeholder groups the ability to share their own perceptions of which WE-RISE activities had the most impact, is an exemplary participatory method. Women, community leaders, *woreda* technical teams, and field staff all participated in ranking 16 different WE-RISE activities. Through the discussion of the differences in ranked items between stakeholder groups, a variety of perspectives emerged: how some activities were ranked low by beneficiaries because only a few people benefitted, yet staff saw large improvements in their lives; other activities started later and did not have time to rate more highly.

Create space in evaluations to share details on programs

Unlike other evaluations, the WE-RISE evaluation includes a description of programmatic interventions. This serves to help others learn about possible successful program designs. Three successful norm-change interventions were identified: VSLAs to improve access to savings and loans; the SAA to facilitate discussions around gender roles; and the use of paralegals to reinforce a positive environment for women's rights and translate government policies down to the local level. Programmatic interventions targeting women's economic empowerment were identified, including shoat[16] rearing, shoat fattening, day-old chicks to select households, among many other activities. In these programs, women were specifically targeted to receive freely-distributed livestock and related training to improve their ability to engage in income-generating activities. The underlying theoretical intervention in this latter set of programs is that by operating input-based activities, women are able to have greater control over resources and provide income to the household. The WE-RISE evaluation described a successful multi-component initiative addressing both economic and social empowerment for women.

6.2   GRAD FINAL PERFORMANCE EVALUATION: Case Study

The GRAD final evaluation (Social Impact, 2017) serves as an exemplary gender-inclusive evaluation document. This is accomplished through: (1) structural inclusion of gender in the report and methodology; (2) programmatic innovations which directly address women and gender relations; and (3) attention to gender in data collection, analysis, and presentation of findings.

6.2.1   PROJECT OBJECTIVES AND THEORY OF CHANGE

Unlike WE-RISE, which targeted chronically-food-insecure women at the objective level, GRAD seeks to graduate households from a Government of Ethiopia program titled Productive Safety Net Programme (PSNP) without mentioning gender. However, the theory of change highlights that women's exclusion in agriculture is a core contributor to poverty. This is followed by mention of livelihoods, households, and communities – in which women are integral contributors.

Table 6. GRAD Objective (Obj) and Theory of Change (ToC)
Obj   The Graduation with Resilience to Achieve Sustainable Development (GRAD)… was designed to support and enhance livelihood options of chronically-food-insecure households by promoting and supporting on- and off-farm income-generating activities, facilitating output and input market linkage, and increasing access to micro-finance services. GRAD's activities complement Ethiopia's Productive Safety Net Program (PSNP) to accelerate the graduate of targeted beneficiaries from PSNP[17].

---

[16] In the Ethiopian context, shoat is a general term for either a sheep or a goat.
[17] Social Impact (2017:1)

ToC	The approach followed by GRAD implementing partners (IPs) was based on a global theory of change that addresses the underlying causes of poverty and particularly women's exclusion in agriculture. GRAD's theory of change was founded on three pillars: (1) enhancing livelihood options; (2) improving household and community resilience; and (3) strengthening an enabling environment to increase GRAD's impact and sustainability. Actually, GRAD complemented the GoE's PSNP by supporting village-level savings/lending and market-driven agri-business value chains (VCs) for selected commodities[18].

6.2.2 EVALUATION METHODS AND RIGOR

Quantitative data was collected via a survey in the same villages as the baseline. The endline survey matched the baseline methodology which purposively sampled by *woreda* and implementing partners. However, while the baseline used simple random sampling from household registers, the endline used the baseline *kebeles* as clusters, then randomly sampled from the *kebele*-level GRAD beneficiary lists. The baseline sample was 1,584 and the endline sample was 1,602; this size is capable of detecting statistically significant changes in proportion at the *woreda* levels (ranging from 9.2% to 10.6%) and in the entire sample overall of 5% or above. Changes less than these percentages will not be statistically significant.

Qualitative data collection was done in the same locations as the survey. Methods included 55 key informant interviews and 41 focus-group discussions. Careful attention was paid to the gender ratio for focus-group participants to properly represent the overall beneficiary gender ratio (which was approximately 40%). Detailed annexes are provided including interview and focus-group discussion guides.

Methodologically, the evaluation was transparent and clear in its methods. Quantitative and qualitative methods are presented as corroborating data, or contradictions are acknowledged with possible explanations included. Qualitative data is presented with limits to generalizations and the source for qualitative data is disclosed. The evaluation data builds credibility by acknowledging when evidence are not available or robust enough to make a claim.

6.2.3 FINDINGS

GRAD made substantial progress on graduating households from PSNP support, reaching 78.6% of the overall goal of 50,000 graduate households. Households in different *woredas* had different outcomes, likely because of the 2015 drought and its impacts. Gender played a stronger role in the food-insecure *woredas*: it appears as if traditional barriers to women's livelihoods and having few adults in the household disproportionally and negatively impacted female-headed households.

---

[18] Social Impact (2017:1)

CIMMYT
International Maize and Wheat Improvement Center

Federal Ministry
for Economic Cooperation
and Development

6.2.4   WHAT WORKS FOR SOCIAL-NORM CHANGE?

Design and implement multi-component programs which incorporate gender-norm change
VESA is the key delivery mechanism for GRAD program components (See Figure 2), offering integrated multi-channel interventions which reinforced one another. By targeting economic empowerment through loans, savings, and agricultural productivity, and then layering on gender-norm change, VESA's appear to have successfully stimulated discussion and norm change between genders. This builds the evidence base that economic models, when paired with explicit gender-sensitivity programming, can enhance women's, men's, and household outcomes from development projects.

Tip for improvement: When describing a successful intervention as measured by project indicators, be sure to include a description of the program so others can learn from the good work. During GRAD, "innovation programs" were developed which specifically targeted women for a micro-franchising intervention, yet too few programmatic details were provided to build a body of evidence.

Figure 2: GRAD's VESA program components
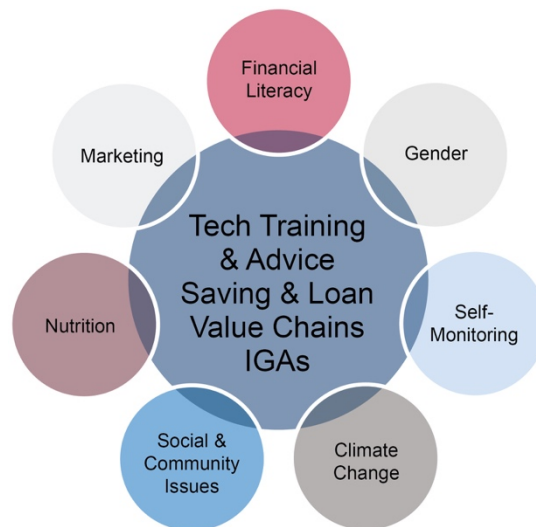
Target gender relations and work with husband/wife teams
Programmatically, the evaluation report presents project beneficiaries, not as isolated individuals, but *in relation to one another*, while also targeting women to ensure they benefit from the project. This was made possible by the programmatic intervention of the Village Economic and Savings Association (VESA) which sought to address *households* as a unit—inviting both husband and wife to attend VESA meetings, a space normally reserved for the men. GRAD took specific measures to target women as beneficiaries, creating more equitable beneficiary results across genders, and sought to address the social relations the couple exists within by working with husbands in the household setting. For male-headed households, ranking the lead woman's contribution to household decision making increased significantly (p=0.000) from 2.9 to 3.5 out of a 4-point scale.

Facilitate discussions on gender-related issues and women's empowerment
The VESAs facilitated discussions on gender issues, allowing husband and wife teams, women and men, and communities to begin grappling with gender-related practices and views. Focus-group discussions and key informant interviews credit the gender training in VESAs as crucial to encouraging dialogue about deeply-held views.

Pair support for women's financial contribution to the household with gender sensitization
Focus-group discussions and key informant interviews credited the statistically-significant change in women's role in household decision-making to women's increased economic contribution to the household. In doing so, women's social capital within the household was strengthened. By pairing income-generating activities with gender trainings for the household, women's contributions are better able to be recognized and valued for more sustainable project outcomes.

Male- and female-headed households are likely to require different programming
Through data disaggregation by household head type, the GRAD evaluation found differential outcomes for male- and female-headed households. Expenditures are a proxy indicator for household economic well-being, as rising income correlates with rising expenditures, and GRAD planned to improve household income through participation in value-chain activities. In three out of

CIMMYT
International Maize and Wheat Improvement Center

Federal Ministry
for Economic Cooperation
and Development

the four surveyed areas for the evaluation, the gap between male- and female-headed households was statistically significant. Female-headed households typically have 1.2 fewer individuals in the household, members which would normally be able to contribute some labor to benefit the household. Despite controlling for household size, *woreda,* and survey period, female-headed households had $77 less in annual income than their male-headed household counterparts. Although both household types experience growth, male-headed household expenditures grew at faster rates and were able to maintain higher expenditures than female-headed households. The evaluation concludes that female-headed households' may face barriers that the GRAD program did not address.

6.2.5   BEST PRACTICES TO BUILD THE EVIDENCE BASE

The following represent the best practices within the GRAD evaluation which help to build an evidence base around gender-norm change.
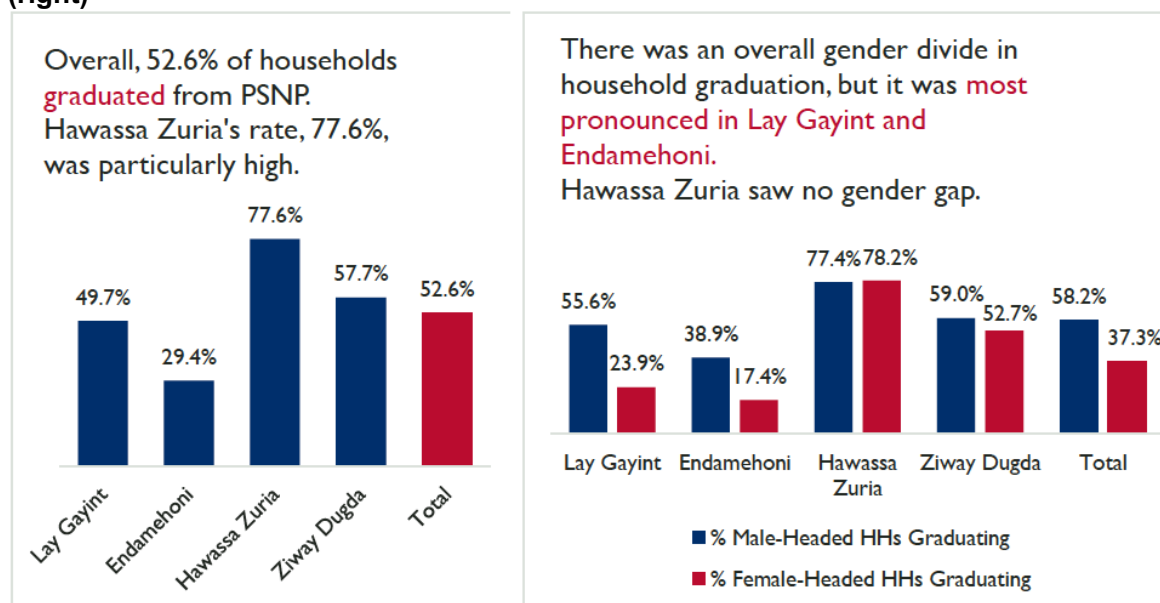
Incorporate and explore gender throughout the evaluation
The evaluation report excelled at addressing gender *both* in a stand-alone section, in which the project's explicit contribution to women's empowerment was assessed *and* mainstreamed throughout the report. The section on gender was made possible by the evaluation Statement of Work (SOW) which included a specific question on gender: to what extent has the activity contributed to gender equity and women empowerment? In several report sections assessing accomplishment on the technical project components, the findings explore changing gender relations, different project outcomes for men and women, and present sex disaggregated data. The report not only explores sex-disaggregated beneficiary ratios, but also reports on implementing organizations' staff ratios.

Explore differential impacts for men and women and by household type
For example, in assessing whether or not the project objective of graduating households from the PSNP program was successful (a supposedly gender-neutral project objective), the evaluation team disaggregated by both head of household and location in order to uncover differential project outcomes for men and women. See Figure 3. Without this disaggregation, the project would have appeared to have more equitable outcomes for both male- and female-headed households. However, through disaggregation, the inability of female-headed households to graduate from the PSNP program is better seen. By including gender consistently throughout the report, the report builds a body of knowledge around project effectiveness for both men and women.

Figure 3: GRAD's graduation rates by *woreda* (left) and disaggregated by household type (right)

Overall, 52.6% of households graduated from PSNP. Hawassa Zuria's rate, 77.6%, was particularly high.

There was an overall gender divide in household graduation, but it was most pronounced in Lay Gayint and Endamehoni. Hawassa Zuria saw no gender gap.

■ % Male-Headed HHs Graduating
■ % Female-Headed HHs Graduating

Use qualitative data to add depth and meaning to quantitative findings
As seen above in Figure 3, disaggregating during analysis uncovered a gender divide in project outcomes. Key-informant interview data provided nuance to quantitative findings by noting differences in women's roles in the household, reproduction, normative restrictions on their mobility, and reduced access to finances. When capturing these 'possible reasons why' from people on the ground and pairing it with survey data, a learning opportunity is made possible.

Critically consider and contextualize moments that appear as 'gender successes'
Although Hawassa Zuria was the only area to have equitable impacts between male- and female-headed households, it was also the only area to *fail* to see expenditures increase. GRAD households in the three other surveyed areas saw average expenditures increase from $329 in 2012 (baseline) to $617 in 2016 (endline), whereas in Hawassa Zuria male-headed household expenditures stayed the same, while female-headed households decreased their expenditures (though the difference was not statistically significant). During the time of GRAD, this area was hit particularly hard by a drought and likely all households struggled to maintain their livelihoods. While Figure 3 demonstrates that Hawassa Zuria had higher graduation rates from the PSNP, the evaluation simultaneously recommends not using PSNP graduation rates as an indicator because it is locally defined. While Hawassa Zuria may first appear as a gender success, it is important to then ask, what outcomes were experienced for the men, women, and families in question?

Investigate interesting outcomes with qualitative case studies to understand how and why
While the Hawassa Zuria case of equitable outcomes for male- and female-headed households may not be as successful as it first appears, it does raise the question if there is a unique configuration occurring in that location which others could learn from. During projects and evaluations, when areas, households, or programs appear uniquely successful in quantitative studies, a qualitative exploration of how and why could help build a body of knowledge about what works to secure equitable developmental outcomes for female-headed households.

CIMMYT MR
International Maize and Wheat Improvement Center

Federal Ministry for Economic Cooperation and Development

Incorporate high-level indicators on men, women, and their relationships to ensure lessons are learned from collected and analyzed data

Gender was also mainstreamed in data collection and analysis processes. Sex disaggregated data was collected at all levels and presented in tables. The GRAD evaluation (Social Impact 2017:6) identified six key outcome indicators, with two indicators explicitly targeting women's roles in decision making: (1) percent of men and women reporting meaningful participation of women in decision making regarding productive resources and income and increased access to productive resources; and (2) percent of women and men reporting an increase in women's influence over household decision making. By including changes in decision-making authority as outcome indicators, GRAD built a monitoring and evaluation system which tracked data in order to measure these outcomes, building data on gender relations.

Check gender bias in your qualitative sample and have a rationale for sample

As seen in the Figure 4, disaggregating the focus group discussions and key-informant interviews allows evaluators and readers to better assess gender bias in the data and findings. By including a percentage column, rather than only raw counts, the relative voice of female beneficiaries becomes easily comparable across reports, stakeholders' groups, etc. The representation of female participants as evaluation data sources largely matches the percentage of women beneficiaries' participation in VESA, value chain groups, and/or cooperative membership.

Figure 4: GRAD disclosure of gender participation in qualitative research methods

Category	Total	M	F	% Female
KII	55	48	7	13
FGD	41	186	128	41
21-VESAs, 16-VC Groups & 5-Coop members	42	222	161	37
Totals	138	456	296	39

Source: November 2016 Endline Qualitative Interviews Results

Build the credibility of qualitative research and use it to explore women's lives and relationships

The GRAD Final Evaluation included an Annex detailing a matrix of evaluation questions against eight different stakeholder groups, such as the Federal Government of Ethiopia, targeted households, and VESA focus groups. All eight stakeholder groups were consulted in order to answer the gender-specific evaluation SOW questions: "To what extent has this activity contributed to gender equity and women's empowerment, specifically in addressing the role of gender in decision making on the use of resources? To what extent has GRAD addressed gender gaps identified among women, men, girls, and boys?" No other evaluation question utilized all eight stakeholders.

Armed with evidence, make context-driven lessons that recognize areas of intervention for future projects

As a result of these gender-sensitive practices, the GRAD evaluation (Social Impact, 2017:55) was able to assess the extent of the activity's contribution to gender equity and women's empowerment:

> GRAD's influence on gender was reported as positive. One of the most important contributions GRAD made to advancing gender equity and women's empowerment was by creating a safe space for husbands and wives, other men and women, and the community in general to begin exploring and discussing gender issues. Some of these changes, however, go against deeply held attitudes and traditions. Although VESA savings groups have women's access to financial resources, women still struggle against unequal access to financial capital and resources. Female-headed households, particularly, have seen slower gains than male-headed households.

CIMMYT
International Maize and Wheat Improvement Center

Federal Ministry
for Economic Cooperation
and Development

Here, GRAD is recognized for creating a 'safe space,' through the VESA programmatic intervention, for both husbands and wives to explore gender issues. Women are not presented as the sole target of these initiatives and other men, their husbands, and the community broadly are also targeted by the intervention. The evaluation recognizes that changes in this area are linked to gender norms and beliefs and acknowledges that not all women experience the project similarly, with the most notable difference being between women in male-headed households and women leading households.

Include gender-related recommendations as part and parcel of project recommendations
With these findings the evaluation report is able to make four gender-related recommendations, which are *interspersed* throughout technical recommendations rather than siloed. These help to iteratively build an evidence base around what works for gender-norm change.

Table 7. GRAD Final Performance Evaluation Recommendations – Gender-Related	
#	Recommendations (n=23)
2	Gender of household heads needs to be more explicitly taken into consideration in future project design to ensure that female-headed households' unique barriers to livelihood opportunities are addressed
10	Balance the gender of staff and community volunteers to reflect the gender balance of project participants, and advocate with GoE to achieve a similar gender balance among their extension workers
20	Include a women's economic empowerment approach, coupled with explicit gender sensitization, in future USAID projects
21	Conduct a sustainability study of gender relationships to identify challenges to maintaining changes after the end of project support

7. KEY LEVERS FOR IMPROVEMENT

Since gender-norm change is a new area within agricultural development projects, there are many ways to learn from the work that has been done, either through best practices or in areas in need of improvement. While reviewing the evaluations which were disqualified during the second round, three key levels for improvement emerged: (1) stronger presentation of qualitative data; (2) ensuring the OECD-DAC criteria are gender-sensitive; and (3) building learning around gender into evaluation even when the project may not have been successful.

7.1 Stronger presentation of qualitative data

The World Bank, IFAD, and DFID agree – qualitative research is an important method for understanding women's lives and the realities they face while participating in, and sustainably benefiting from, development projects. Qualitative research is undervalued in development evaluation, further deepened by the lack of rigor around qualitative research methods. IFPRI (2013:10)[19] argues qualitative research has an important role to play in assessing agricultural development projects:

---

[19] Please see IFPRI (2013) available here, pages 24-27 for methods and pages 123-126 for a substantive example of how to present rigorous qualitative data about differential gender impacts.

The main motivations for use of qualitative research is the recognition that (1) processes of program implementation, and how these are received at the local level, involve complex interactions, and may be different than anticipated by program planners; understanding these interactions and their effects on outcomes requires the use of qualitative research methods that develop trust and rapport between researchers and respondents, allow for open-ended responses, and triangulation through multiple respondents and participant observation; and is the recognition that (2) qualitative methodologies can provide insights into the reasons and causes of certain impacts that will not necessarily be uncovered by quantitative approaches alone; and that (3) views, opinions, and interpretations of the program held by beneficiaries are important, credible, and worth listening to.

Here are some tips to strengthen the rigor of qualitative research for development evaluation:

- Develop a strong sampling method for qualitative research, including key informant interviews and focus-group discussions. Ensure sampling method includes women beneficiaries of different household types, organization staff, and government partners. Share descriptive tables on qualitative data sources, breakdown sources by gender and include percentages for easy cross comparison.
- Include annexes which detail focus-group strata and rationale and focus group and interview guides. In interview guides, ask questions which probe *how and why* a project intervention was successful, adding value beyond confirming the *what and how many* offered from survey findings.
- Create a naming strategy for interview and focus-group data so that respondents may remain anonymous.
- Interweave qualitative data and survey data, using qualitative data to *make sense of* what the numbers mean. This presents qualitative data as adding value above and beyond data triangulation.
- Analyze focus-group discussions for themes *across* participants, rather than selecting single quotations to serve as sound-bites.
- Present quotations framed by statements that highlight their representativeness. Include and explore divergences. If possible, include multiple examples or quotations from beneficiaries to demonstrate generalizability or demonstrate divergences.
- Use participatory research methods, such as stakeholder ranking to best capture project impacts from the perspective of beneficiaries[20].
- If survey data identifies high-achieving areas or households related to gender concerns, include recommendations for qualitative case studies to follow up and build knowledge based on what works.

7.2   Ensuring the OECD-DAC criteria are gender-sensitive

Some evaluations which had strong methods, yet did not include gender, used the OECD-DAC criteria for evaluation.[21] The OECD framework assesses performance in relevance, effectiveness, efficiency, impact, and sustainability—all of which are able to report and evaluate gender-related progress. However, from the sample here, when evaluations utilize the OECD-DAC criteria, gender appears to be sidelined or forgotten. This may occur even if the project has successfully addressed gender programmatically.

---

[20] See TANGO (2016) for details on stakeholder ranking.
[21] For more information please see the OECD-DAC criteria online.

CIMMYT
International Maize and Wheat Improvement Center

Federal Ministry
for Economic Cooperation
and Development

The CARE Ethiopia's ABDISHE Final Evaluation (TZBMC, 2016) serves as a strong example of how to ensure the OECD-DAC criteria works for women and gender-related issues. In the ABDISHE evaluation, the project's performance indicators were turned into questions and then matched to OECD-DAC areas, resulting in 17 questions across the five OECD criteria. Although not all questions appear gendered, the ABDISHE project incorporated women at the highest level of project design, working to graduate chronically-food-insecure women from the PSNP program. This is further enhanced by CARE's request that the evaluation also respond to the effectiveness of CARE's women's empowerment framework of agency, structure, and relations.

Relevance:
- ABDISHE: Are the project's activities well-suited to the needs and priorities of the target groups?
- Recommended: How were women and/or gender relations targeted by this project? Did the programs address the everyday realities for women?

Effectiveness
- ABDISHE: What evidences are there to demonstrate women's graduation from a chronically-food-insecure situation?
- Recommended: What evidence is there to demonstrate women's participation resulted in desired outcomes and impact? What evidence is there to demonstrate changes in gender relations?

Efficiency
- ABDISHE: Were the objectives of ABDISHE achieved on time?
- Recommended: Was women's meaningful participation supported early on through differential investment in items such as childcare? Were delays or issues in engaging women taken seriously and remedied quickly?

Impact
- ABDISHE: What were the positive and negative changes to beneficiaries that, directly or indirectly, intended or unintended, were due to the ABDISHE's project's intervention?
- Recommended: How were women differentially impacted by the program than men? How were the different types of households impacted—did the project have equitable outcomes irrespective of household type? Were women in male-headed households changed? What social norm changes are observed in household and community gender relations?

Sustainability
- ABDISHE: How did the community/ local government have a sense of ownership during the project (development of outcomes, design and implementation)?
- Recommended: What structures exist to provide ongoing support to women and/or changing gender relations beyond the project? Has women's status in the household and/or community grown to a point where intervention is no longer needed (and be clear about who should determine this)?

7.3 Build learning around gender into evaluations, even when the project may not have been successful

The Agribusiness Marketing and Development (AMDe) mid-term evaluation (Tufts University, 2015b) serves as a strong example of how evaluations can successfully mainstream gender and build an evidence base, even when not accomplishing the desired results[22]. The mid-term did not include a survey, but was informed by a review of project documents and 215 interviews with key

---

[22] As of June 2018, the final AMDe evaluation report was not available. To read the AMDe Midterm Evaluation report in full, access it here.

CIMMYT
International Maize and Wheat Improvement Center

Federal Ministry for Economic Cooperation and Development

stakeholders, including farmer organizations, primary cooperatives, farmer cooperative unions, partner organizations, industry associations, and the private sector.

The evaluation SOW required the evaluation address the question: "To what extent has this project contributed to gender equity in terms of access to credit, capacity-building support, improved inputs and technologies resulting in an increase in sales of agricultural commodities? Is there evidence supporting positive changes in the aforementioned areas?" The evaluation had a single section dedicated to gender equity and women's empowerment *and* discussed gender and male/female beneficiaries in several locations throughout the report. Because of this dedication:
- The evaluation recognizes differences among women, explicitly mentioning women-headed households and women in male-headed households.
- Women's differential participation is presented as intersecting with poverty. Noting that poor households, the target beneficiary, are not currently being reached by the AMDe program strategies which focus on the middle- and upper-tier of the value chain.
- The evaluation detailed a *gendered* value-chain analysis, noting that if the project objective is to assist smallholder women in value-chain participation, the selection of domestic-product value chains would have been more advantageous given that men dominate export-oriented value chains.
- The evaluation mentioned sex-disaggregated data throughout the report, allowing for differential impacts to be better seen.
- The evaluation documented that budget was not allotted to support the engagement of women.
- The report disclosed the number of male and female staff within the implementing organizations.

The evaluation (2015b:15) was able to draw out lessons about the gendered nature of the program components.
- It raised concerns about the use of incentives to drive women's membership in farmer cooperatives, citing concerns about sustainability, and instead emphasized the meaningful engagement of women to drive demand for cooperative membership:
- AMDe utilized a gender quota system to ensure women's participation in all training sessions, exchange visits, and investment support. The quota was set at 30%. The report (2015b:14), however, argues this number underestimates women's role in agriculture:

> *Women-headed households account for 28 percent of households in AGP* woredas*, and women in male-headed households typically constitute 50 percent of family labor. The 30 percent figure therefore under-represents women in agriculture.*

- The "main gender equity success" was the creation and implementation of the Women in Agribusiness Leadership Network (WALN). WALN programming included "business development training—negotiation, marketing, networking, financial planning, and communication skills— leadership training, mentoring and coaching support, and networking opportunities for women leaders operating in Ethiopia's agribusiness sub-sector." However, no evidence was provided for this claim other than the number of women (100 women) at mid-term receiving a total of USD$1.5 million and that a national conference was held.
- Gender is included in the recommendations for a follow-up project to AMDe with specific recommendations to improve targeting and programming for poor smallholder farming women, strengthen women's genuine participation in farmer cooperatives, and build value chains according to the project's Theory of Change. Budget allotment is again highlighted as important.

CIMMYT MR
International Maize and Wheat Improvement Center

Federal Ministry
for Economic Cooperation
and Development

As the AMDe evaluation demonstrates, it is possible to capture key aspects of interventions and how they have helped or hindered women's equitable benefits. This learning can now be used to inform future programming.

8. CONCLUSION

This Research for Development (R4D) Policy Report summarizes findings from a meta-analysis of external evaluations of Ethiopian agricultural development projects, while highlighting best practices around gender programming. Gender equality has proven important for sustainability, to decrease unintended consequences, and to improve agricultural productivity. Social norms lie behind gender inequality in endowments and agriculture productivity but have not been adequately studied. The report explores: What works for gender-norm change in agricultural development projects? The findings outline both what to do and highlight what to avoid in undertaking gender transformative development.

This report has identified areas in need of improvement based upon the 24 disqualified documents and best practices from two evaluations to strengthen the evidence base. The findings show that the CARE Ethiopia office is producing the most rigorous and successful projects around gender-norm change. Ethiopian agricultural development projects should incorporate gender-norm change at the highest level of project design; pair income-generating activities or savings with community discussions and training around gender relations; allot differential budget and resources to engage women as participants and decrease their opportunity costs; and ensure that male- and female-headed households receive differential programming or consideration.

Evaluations in agricultural development are written to evaluate a project's achievement to the project contract with an emphasis on effectiveness and efficiency. Although evaluation theory is meant to outline the lessons learned about 'what works', this element is currently being relegated to internal documents. In effect, learning 'what works' is rendered interior to the donor and implementer relationship, rather than more widely circulated through public final-performance evaluations. Evaluations serve an accountability function, rather than a learning function.

Areas to target for improvement:

- **Detail successful programs:** Descriptions of project activities and intervention should be detailed in the body of the evaluation or in an annex. This should include the identification of mechanisms that are effective, as well as an exploration of how and why, rather than silo this information in an internal learning report that is unlikely to be shared, and, if it is shared, unlikely to be trusted due to a lack of independence. Evaluations should pair empirical evidence (both qualitative and quantitative) with programmatic details about the mechanisms that produced the results.
- **Strengthen qualitative research skills:** Build analysis and presentation of qualitative data as a rigorous, empirical method.
- **Strengthen empirically-backed claims:** Many evaluations failed to develop findings from data, resulting in 'empty' evaluations that claim effectiveness without credibility.
- **To help build a foundation of data, build gender into project objectives and monitoring and evaluation systems:** Many evaluations did not have gender-related data upon which to base their claims. This was often caused by project monitoring and evaluation systems which did not capture gender dynamics or disaggregate by categories meaningful to women, such as household type (male- or female-headed, widowed, divorced).
- **Ensure the evaluation Statement of Work includes a question regarding differential impacts on men and women, household types, and gender relations:** Some evaluation SOWs did not request assessment around gender, nor mandate attention to gender. This

CIMMYT
International Maize and Wheat Improvement Center

Federal Ministry
for Economic Cooperation
and Development

results in final evaluations which do not address gender, even if the project substantively invested in gender.

- **If using OECD criteria, add a section which evaluates gender specifically or consider under each criteria:** OECD-DAC criteria focuses attention on the relevance, effectiveness, efficiency, impact, and sustainability of projects. This appears to only result in a gendered analysis when the project itself incorporates gender.
- **Theories of Change should include targeting of women and gender relations:** When women and/or gender are incorporated within the theory of change a ripple effect occurs: gender is included in the project's highest-level objectives and outcome indicators; project monitoring and evaluation systems are more likely to develop specific indicators to measure progress against social-norm change, especially around decision making; programs are more likely to concertedly address the needs of women and their households and communities; and the final evaluation will have data to inform the findings. This is a key lever for building strong programs and iterative contributions to the evidence base.
- **Incorporate gender in the design of the project:** Project design should account for the targeted and/or differential investments required to ensure equal participation of women beneficiaries. This may include basic literacy and numeracy courses, childcare, and addressing gender relations to ensure that women's benefits from the program are effective and sustainable.
- **Develop organizational definitions of gender equality or women's empowerment:** Organizations should articulate a vision of their contribution to women and gender through agricultural development. This provides a scaffolding for project designers.

Gender-norm change in agricultural development projects is a fledgling area that deserves investment and careful measurement to strengthen a body of evidence. Summative evaluations are currently a missed opportunity for learning. The agricultural-development sector should stop failing to learn about gender and quickly incorporate new project and evaluation practices to strengthen learning about successes in gender-related agricultural development. Agricultural-development project evaluations are currently a lost opportunity for learning 'what works' for gender-norm change.

CIMMYT
International Maize and Wheat Improvement Center

Federal Ministry
for Economic Cooperation
and Development

9.   Appendix A: State of Learning

State of Learning:
As the development sectors struggle to understand 'what works' for gender-norm change, synthesizing the lessons across research is key to establishing a foundation of evidence. Three large-scale reviews were analyzed to understand the state of evidence on 'what works' for social-norm change. The documents analyzed include:

Table x. Methodologies of Meta-Reviews

DFID PPA Learning Partnership Gender Group. June 2015. What works to achieve gender equality and women's and girl's empowerment? ActionAid UK & Christian Aid. Available [here](#).
Strong discussion of gender-norm change theory and essential vocabulary. Includes 24 case studies which highlight successful social-norm change at individual, household, community, and societal levels. From these, recommendations are given for designing gender-transformative projects.

Independent Evaluation Group (IEG). 2016. Women's Empowerment in Rural Community-Driven Development Projects: An IEG Learning Product. World Bank Group. Available [here](#).
This report is based on 20 Community-Driven Development (CDD) projects from 1999–2017 primarily in Asia and Africa with one Middle Eastern and South American project. Reports were analyzed for the following questions: "Do CDD interventions result in economic, social, and/or political empowerment of women, as well as men?" and "What are the conditions (including contextual elements) and the design elements that enhance or weaken these impacts?". From this assessment, recommendations are made to improve the empowerment of women.

IFAD. March 2017. What works for gender equality and women's empowerment? A review of practices and results. Evaluation Synthesis. Report No. 4390.
Undertaken to assess project performance against the IFAD Gender Equality and Women's Empowerment (GEWE) Policy. Reviewed 57 IFAD project documents selected from a sampling frame of 163 documents, all of which had an internal gender rating from 2011-2015. Beginning from the sampling frame of all GEWE-rated projects, their next sampling criteria required projects to report outcomes, strategies, *and* evidence on gender results. This resulted in only 17 documents, including those with a "partial gender mainstreaming" GEWE rating (4 out of 6 with 6 being "gender transformative). To raise their sample, IFAD loosened their original inclusion criteria on 'evidence' (defined as impact assessments or counterfactuals) to grow their sample to 57 documents. Recommendations are provided to enhance learning, reporting, and project design around reaching and empowering women.

Note: This is not a comprehensive list of gender-norm change meta-reviews.

CIMMYT
International Maize and Wheat Improvement Center

Federal Ministry
for Economic Cooperation
and Development

10.   Appendix B. Documents Considered for Inclusion – First Round

List of documents considered for inclusion - first round of selection criteria (n=26)

Project (Implementer)	Document Type			Independence	
	Evaluation	*Gender*	*Learnings*	*Internal*	*External*
ABDISHE (CARE)	x				x
AMDe MTE	x				x
Defar (Send a Cow)	x				x
Emergency Nutrition Response (GOAL)	x				x
ENGINE MTE	x				x
ESSPII EDRI Report PSNP 2013	x				x
FEED II MTE (ACDI/VOCA)	x				x
GRAD Final (not in yet)	x				x
LMD MTE	x				x
LNWB (UNFPA)	x				x
PRIME MTE	x				x
SEAES (Oxfam America)	x				x
WE RISE (CARE)	x				x
ABCD (Coady International Institute)	x			x	
AGP II Social Assessment (GoE) 2015	x			x	
Ethiopia Bee Keeping (Oxfam GB)	x			x	
GRAD (CARE) Outcome Mapping Process Report	x			x	
Sustainable Land Management Project II - Social Assessment		x		x	
AGP (GoE) Gender Analysis		x			x
FSF (CARE) Gender Analysis		x			x
GRAD (CARE) VC Gender Analysis		x			x
ABCD documentation			x	x	
GRAD (CARE) Outcome Mapping Report		x		x	
P4P documentation			x	x	
PASIDP-II Participatory Small-scale Irrigation Development			x	x	
Send a Cow - Household Methodologies			x	x	

11. Appendix C. Second Round of Inclusion Results

Project (Implementer)	Rigor			Gender included in evaluation			Social-norm change desired		
	Quant	Qual	Evidence	None	Section	Woven	None	Claimed	Present
ABDISHE (CARE)	x	x			x	x	x		
AMDe (ACDI/VOCA)		x			x	x	x		
Defar (Send a Cow)		x			x		x		
Emergency Nutrition Response (GOAL)	x	x	x	x			x		
ENGINE (Save the Children)		x			x		x		
ESSPII EDRI PSNP 2013 (GoE)	x	x	x		x		x		
FEED II (ACDI/VOCA)	x	x	x	x	x		x		
GRAD (CARE)	x	x	x		x	x			x
LMD MTE (CNFA)		x			x	x	x		
LNWB (WFP)	x	x				x	x		
PRIME (Mercy Corps)		x			x		x	x	
SEAES (Oxfam America)	x	x		x			x		
WE RISE (CARE)	x	x	x		x	x			x

Rigor: Quantitative sampling, Qualitative presentation, Substantiated evidence for claims
Gender Included in Evaluation: No treatment of gender, Included largely in a single 'gender section,' Woven throughout
Social-norm Change Desired: No stated desire for social-norm change, Claimed norm change (without evidence) or desired, present with substantial evidence
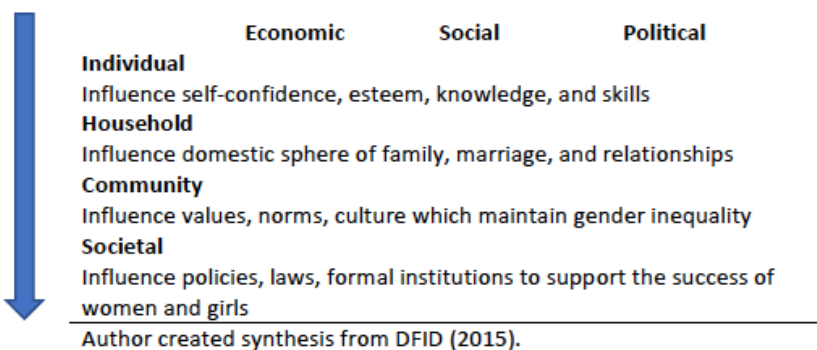
CIMMYT MR
International Maize and Wheat Improvement Center

Federal Ministry for Economic Cooperation and Development

12. Appendix D. Research-driven guidelines

In addition to the high-level guidelines identified through the review of meta-analyses, more specific lessons were generated from the review of Ethiopian agricultural development projects. To aide in their application and uptake, they are presented, broadly, across the project cycle.

During project design, be sure to answer each question thoroughly:
- Have you reviewed existing research to understand the social-norm context for women *before* designing the project?
- Does your project address multiple levels of women's lives? Including individual agency, household dynamics, community support, and societal interventions?
- Which of the following substantive areas does your project target: social, economic, political?
- Have you explicitly targeted different types of women? Women or girls broadly, women in male-headed households, women household heads (perhaps divorced or widowed).
- Have you incorporated a social-norm change element to reinforce and catalyze investments in income generation?

Figure 5: Multi-level and multi-component project design schema

	Economic	Social	Political
Individual			
Influence self-confidence, esteem, knowledge, and skills			
Household			
Influence domestic sphere of family, marriage, and relationships			
Community			
Influence values, norms, culture which maintain gender inequality			
Societal			
Influence policies, laws, formal institutions to support the success of women and girls			

Author created synthesis from DFID (2015).

- Have you set targets for women's participation to ensure their equal participation?
- Have you incorporated additional programming or strategies to ensure the participation of women?
- Have you taken programmatic steps to reduce the opportunity cost of women to participate? Have you created budget lines and timeline adjustments to accommodate these?

Project Design
- Multi-level, integrated programs are more likely to result in sustainable social-norm change. This includes developing programming to reinforce social-norm change at the individual, household, community, and/or societal levels.
- 'Gender norms' are locally-embedded and socially-constructed, meaning it will take different forms in each context. While some knowledge is transferable, research must be completed to understand each context.
- The root causes of gender inequality and power differentials should be analyzed and incorporated into the project design. Without gender-specific assessments, projects run the risk of perpetuating and reinforcing social norms and are unable to address gender-specific dimensions in their project and targets.
- Gender inequality should be understood as intersecting with other inequalities, such as class, race, religion, and skill-level. Gender analyses should describe the myriad of differences among and between women and studies across economic, social, and political dimensions.

- Project designs should be based on a Theory of Change which addresses transformative change for women and girls and/or gender norms. A clear description of empowerment, identification of social, economic, and/or political dimension, and how each is expected to be impacted should be articulated. Backlash against women should be considered and mitigated.
- Look at organizational gender dynamics and staffing: Are women represented at the beneficiary level? Within implementing partners? Data collectors or enumerators? Partnering government agencies?

Project Length and Expectations
- Changing women's social status and abilities takes time. It is simplistic to think you can 'empower women' in the short term. Long-term projects are required for changes in discriminatory norms.
- Empowerment should be understood as both a process and an outcome. Consider short- and long-term impacts.
- Long-term projects build a critical mass of women who then have the skills, capacity, and networks to maintain social change momentum.

Programming
- Women should be supported through programs to determine their own pathways for empowerment.
- Develop programmatic elements to address gender-norm change while incorporating components to address barriers to women's participation (child care, transportation, etc.).
- The most commonly-used project strategy to engage women is to introduce a quota system, yet there is no evidence which demonstrates that this is an effective strategy for women's active participation once present.
- Projects have successfully promoted female attendance, and sometimes female participation in meetings. However, the quality of participation and maintenance of this engagement is not well understood. Projects should focus on quality and sustainability.
- Projects are likely to target women and support their participation but rarely assess whether a women's participation has altered social and community standing.
- Allot budget for gender-related activities, including resources and staff. Proactively encourage women's participation in projects by addressing their opportunity cost through supplementary programs like provided childcare.
- Project activities should include and incorporate men and boys, including traditional leaders.
- Develop programs which address one or more dimensions: economic (market actor), social (social actor), and political (civic actor) domains of empowerment.

Project Monitoring
- The evidence base of 'what works' has made great strides but has much room to grow. Systematic assessment, reporting, and evaluation, including rigorous research, is needed in order to improve our learning on women's empowerment.
- Innovative programs should be encouraged; the sharing of best practices, and knowledge exchange should be an essential component of all projects.
- Consult local women's groups on relevant and needed research topics.
- Any mention of women as target beneficiaries or women's empowerment as a project objective or a cross-cutting theme must be matched with explicit gender-specific output and outcome indicators.

CIMMYT™MR
International Maize and Wheat Improvement Center

Federal Ministry
for Economic Cooperation
and Development

- Develop indicators to measure gender dynamics at the household and community levels and individual agency. Implement at the time of baseline and include in project monitoring systems.
- Enforce the collection of sex-disaggregated data for applicable indicators.
- Pay attention to the unit of analysis for beneficiaries. Is it individuals? Is the focus on the household? If yes, how is that defined and who is able to speak on behalf of the household?
- Projects which only measure output indicators will fail to capture gender-norm change and meaningful women's participation, resulting in a dearth of information on what works or does not work for empowering women.
- Measure long-term outcomes and impact. Do not focus solely on input and output indicators. Successful measurement of outputs should move forward into clear outcomes and consider the context beyond the project to best capture gender dynamics and the changing social status of women.
- Understand that quantitative data may provide strong outcome data, yet measuring processes, such as women's empowerment, qualitative and participatory data-collection methods are required.

Evaluation Methods and Considerations
- All evaluations (baseline, midline, and endline or summative evaluation) should include gender-related questions in their Statement of Work.
- Executive summaries should consistently include gender-related findings, documenting both successes and/or failures. If the project failed due to gender norms, document what they were and recommend future projects address them through programming.
- Incorporate participatory research methods to hear beneficiary perspectives. Ensure women have a clear voice and space during data collection, or local gender norms could silence women, resulting in biased data. Participatory methodologies may help empower women through the very act of determining and defining 'progress' in their lives.
- Mixed methods research should be included in all evaluations. Mixed methods and multi-dimensional indices best capture transformative change. Some elements that are significant in women's lives and process-based such as empowerment may not be best captured through quantitative data.
- Design the evaluation to test the project's Theory of Change.
- Design a sampling strategy that stratifies by men and women, or by head of household type. Results are typically skewed towards the male perspective due to sampling.
- Design rigorous qualitative elements and understand qualitative data to offer meaning and explanation in addition to the triangulation of findings.
- Be explicit about collecting data around gender relations and/or differential impacts.
- If using the OECD-DAC evaluation framework, add a custom question on gender which matches the overall intent of the project's Theory of Change.
- Ensure data collectors are trained in gender-related issues and gender is accounted for in matching beneficiaries with data collectors.
- Build the evaluation design so that the results will speak to programmatic effectiveness for beneficiaries, not only accountability to funders regarding performance on targets.
- Evaluations should include more programmatic information and identify successful mechanisms for change. Link survey results to programmatic interventions.
- In evaluations, present women's voices or any gender-related cases framed by statements of representativeness. Cases or 'anecdotes' should not only be focused on literal indicators the project was interested in, but address context.
- Gender should be considered in each of the technical evaluation questions and not just siloed into a 'gender section.'

CIMMYT
International Maize and Wheat Improvement Center

Federal Ministry
for Economic Cooperation
and Development

- Define and describe *who* smallholder farmers are; use language which presents 'farmers' as both men and women.
- Capture gender-related lessons and document them in the 'Recommendations' section to ensure learning is moved forward. Document barriers to participation if women were not well represented.

13. References

Boudet, Munoz, Ana Maria, Patti Petesch, Carolyn Turk, and Maria Angelica Thumala. 2012. *On Norms and Agency: Conversations about Gender Equality with Women and Men in 20 Countries (English)*. Washington, D.C.: The World Bank.

Climax Consulting. 2016. *Final Evaluation of Emergency Nutrition Response for Vulnerable Families - Increased Resilience to Food Security in Both Boricha (SNNPR) and Daro Lebu (Oromia) Woredas*. GOAL Ethiopia.

DFID PPA Learning Partnership Gender Group. 2015. *What Works to Achieve Gender Equality and Women's and Girl's Empowerment?* ActionAid UK & Christian Aid.

DFID 2012. *A Theory of Change for Tackling Violence against Women and Girls.* CHASE Guidance Note Series https://www.womankind.org.uk/docs/default-source/resources/vawg_guidance1_toc11.pdf?sfvrsn=2

FAO. 2011. *The State of Food and Agriculture: Women in Agriculture*.

Haider, H. 2017. *Changing Gender & Social Norms, Attitudes and Behaviours.* Brighton, UK: Institute of Development Studies.

HEDBEZ Business & Consultancy PLC. 2015. *End-Line Evaluation of Strengthening Ethiopia Agricultural Extension System (SEAES) Project (Final Report)*. Oxfam America Horn of Africa Regional Office.

ICOS Consulting PLC. 2015. *Mid-Term Evaluation of Abdishe/LINKAGES Project in East Hararghe (Final Report)*. CARE Ethiopia.

ICOS Consulting PLC. 2016. *Mid-Term Evaluation of the Feed Enhancement for Ethiopian Development Project Phase II (FEED II)*.

IFAD. 2017. *What Works for Gender Equality and Women's Empowerment – a Review of Practices and Results.* Report No. 4390.

Independent Evaluation Group (IEG). 2016. *Women's Empowerment in Rural Community-Driven Development Projects. An IEG Learning Product*. World Bank Group.

International Food Policy Research Institute. 2013. *Evaluation of Ethiopia's Food Security Program: Documenting Progress in the Implementation of the Productive Safety Nets Programme and the Household Asset Building Program*.

Kabuchu, Hope. 2013. *Final Evaluation UN Joint Programme on Leave No Woman Behind (LNWB)*. MDG Achievement Fund.

Mackie, Gerry, Francesca Moneti, Holly Shakya, and Elaine Denny. 2015. *What Are Social Norms? How Are They Measured?* UNICEF/University of California, San Diego, Center on Global Justice.

Marcus, Rachel. 2014. *Changing Discriminatory Norms Affecting Adolescent Girls through Communications Activities: Insights for Policy and Practice from an Evidence Review*. Overseas Development Institute (ODI).

Marcus, Rachel and Ella Page. 2014. *Drivers of Change in Gender Norms: An Annotated Bibliography*. Overseas Development Institute (ODI)

Raven-Roberts, Angela, Wolde Tadesse, and Fassil Gebeyehu. 2015. *Rising from the Hillsides: From Scarcity to Surplus in the Wolayita and Gamo Highlands of Ethiopia - Evaluation of Send-A-Cow Ethiopia Defar Project in Gamo Gofa and Wolayita Zones, SNNPR*. Send-A-Cow.

Social Impact, Inc. 2017. *Final Performance Evaluation: Graduation with Resilience to Achieve Sustainable Development (GRAD) Activity*. USAID.

TANGO. 2016. *WE-RISE Ethiopia End Line Evaluation*. CARE Ethiopia.

CIMMYT MR
International Maize and Wheat Improvement Center

Federal Ministry
for Economic Cooperation
and Development

Teruneh Zenna Business and Management Consultancy (TZBMC). 2016. *ABDISHE/Linking Initiatives, Stakeholders to Achieve Gender-Sensitive Livelihood Security (LINKAGES) Project: End Line Evaluation Report*. CARE Ethiopia.

Tufts University. 2015a. *Agricultural Growth Program - Livestock Market Development (AGP-LMD) Project EXTERNAL MID-TERM PERFORMANCE EVALUATION REPORT*. USAID.

Tufts University. 2015b. *Agricultural Growth Programme - Agribusiness Marketing and Development Project (AMDe), Ethiopia MID-TERM EVALUATION REPORT*. USAID.

Tufts University. 2014. *Empowering New Generations to Improve Nutrition and Economic Opportunities (ENGINE), Ethiopia EXTERNAL MID-TERM PEROFRMANCE EVALUATION REPORT*. USAID.

Tufts University. 2015c. *Pastoralist Areas Resilience Improvement Through Market Expansion (PRIME), Ethiopia Mid-Term Evaluation Report*. USAID.

CIMMYT ᴹᴿ
International Maize and Wheat Improvement Center

Federal Ministry
for Economic Cooperation
and Development