Agricultural Systems 240 (2027) 105062

Contents lists available at ScienceDirect

Agricultural Systems

journal homepage: www.elsevier.com/locate/agsy

Research Paper

GenderUp in practice: Barriers and opportunities for responsible scaling in agricultural research for development☆

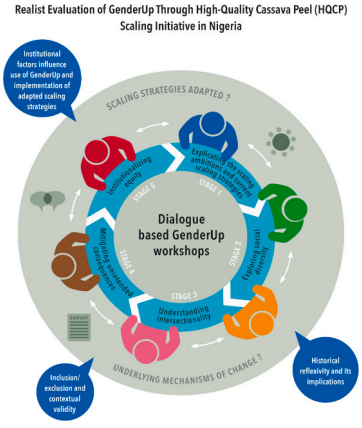
Ebenezer Ngissah [a,b,*], Esther Kihoro [a,b], Cees Leeuwis [a], Katharine Legun [a], Mareike Smolka [a,e], Comfort Freeman [d], Erin McGuire [c], Adebola Bunmi Adebayo [b], Hanna Ewell [a,f], Jasada Ijudai [b]

[a] Knowledge, Technology and Innovation Group, Social Sciences Department, Wageningen University and Research, Wageningen, Netherlands
[b] International Livestock Research Institute, Nairobi, Kenya
[c] University of California Davis, California, USA
[d] Department of Agricultural Extension, University of Ghana, Accra, Ghana
[e] Human Technology Center, RWTH Aachen University, Aachen, Germany
[f] Alliance Biodiversity and Ciat, Nairobi, Kenya

HIGHLIGHTS

- Process complexity makes socially diverse responsible scaling difficult to achieve
- GenderUp's structured reflection reveals unintended potentially negative scaling effects
- Donor and consortium dynamics enable participation but constrain implementation
- Limited historical reflexivity obscures key local barriers, like feed adulteration
- Diversity of workshop participants affects the contextual validity of findings

GRAPHICAL ABSTRACT



ARTICLE INFO

Keywords:
GenderUp
Responsible scaling
Realist evaluation
Social inclusion
Donor/partnership influence

ABSTRACT

Context: Transformative agri-food system change requires innovation and scaling approaches that are socially inclusive. Agricultural Research for Development (AR4D) organizations therefore emphasize the integration of gender and social diversity into innovation design and scaling, but translating these commitments into practice remains challenging. Although scaling decision support tools (SDSTs) look promising, research on their practical effects is limited.

Objective: This study assesses GenderUp, a dialogue-based SDST designed to identify relevant social diversity, anticipate differentiated consequences, and adapt scaling strategies for equity by (i) eliciting the underlying

☆ This article is part of a Special issue entitled: 'Responsible Scaling for Transformative Agrifood System' published in Agricultural Systems.

* Corresponding author at: Knowledge, Technology and Innovation Group, Social Sciences Department, Wageningen University and Research, Wageningen, Netherlands.
E-mail address: ebenezer.ngissah@wur.nl (E. Ngissah).

https://doi.org/10.1016/j.agsy.2026.105062
Received 30 January 2026; Received in revised form 3 November 2026; Accepted 5 November 2026
Available online 9 November 2026
0308-521X/© 2026 The Authors. Published by Elsevier Ltd. This is an open access article under the CC BY license (http://creativecommons.org/licenses/by/4.0/).

mechanisms of change through which it is expected to influence scaling processes and (ii) examining how and under what conditions it contributes to socially inclusive scaling.
Methods: We adopted a realist evaluation approach using document analysis, observations of a GenderUp workshop, follow-up observations of project inception meetings one year and six months after the workshop, and interviews and focus group discussions with stakeholders across the value chain to examine and test GenderUp's elicited mechanisms of change in the context of upscaling high-quality cassava peel (HQCP) technology in Nigeria.
Results and conclusions: GenderUp triggered critical reflection, enabling teams to identify diversity dimensions, articulate intersectional at-risk profiles, and revise scaling-strategy elements. However, donor requirements and consortium structures shaped the translation of insights into actionable strategies, enabling participation but constraining implementation of re-designed scaling strategies by prioritizing quantitative targets. Limited local representation in the workshop meant that outcomes reflected national-level assumptions rather than local realities; and insufficient historical reflexivity left key barriers e.g., feed adulteration undetected.
Significance: The study provides empirical insight into how SDSTs can advance inclusive scaling in AR4D and highlights key enabling conditions and methodological improvements, such as: enhancing participant selection guidance, incorporating local cross-checking and information gathering, and engaging donors and consortium partners more explicitly. These can strengthen the contextual validity and institutional uptake of findings generated through tools like GenderUp.

1. Introduction

Gender equality and social equity are global priorities central to achieving the United Nations' Sustainable Development Goals (SDGs). Agricultural Research for Development (AR4D) institutions, many of which operate with explicit mandates to contribute to the SDGs, increasingly recognize the need to integrate diversity, including gender, into the development and scaling of innovations (Lugalla, 2008; McGuire, Al-Zu'bi, et al., 2024; McGuire et al., 2022). Such integration ensures that innovations contribute equitably to development objectives while minimizing potential harm to vulnerable or marginalized groups (McGuire, Leeuwis, et al., 2024). Translating this commitment into practice however is difficult, because it requires scaling decisions that are attentive to social differences in complex, multi-actor settings (McGuire, Leeuwis, et al., 2024; Sartas et al., 2020). This study examines one attempt to support such decision-making: it assesses the application of GenderUp, a dialogue-based scaling decision-support tool (SDST) designed to help innovation and scaling teams identify relevant dimensions of social diversity, anticipate potential unintended negative consequences of scaling an innovation, and adapt scaling strategies to mitigate these risks and support a more equitable distribution of benefits (McGuire, Leeuwis, et al., 2024). Rather than assuming that the use of such SDST necessarily leads to socially inclusive outcomes, we examine what happens when GenderUp is applied in practice. Specifically, we elicit the underlying mechanisms of change through which GenderUp is expected to influence scaling processes, and examine how, why and under what contextual and institutional conditions these mechanisms are realised, enabled or constrained. Through this analysis, we seek to understand the potential contribution of dialogue-based SDSTs to socially inclusive scaling in AR4D.

1.1. From systems thinking to responsible scaling

This increased recognition for the integration of diversity into scaling processes builds on long-standing efforts to operationalize systems thinking in AR4D. Systems approaches to agricultural innovation have evolved considerably over time (see Klerkx et al., 2012). Currently prominent approaches include the thinking about agricultural innovation systems (AIS, e.g. Ekboir, 2003; Hall et al., 2003; Hall and Clark, 2010; Spielman and Birner, 2008) and perspectives drawing on the multi-level perspective on sustainability transitions (Grin et al., 2010; Darnhofer, 2015; Leeuwis et al., 2021). These approaches emphasize that agricultural development is a complex and multidimensional process (Douthwaite and Hoffecker, 2017), involving a web of interdependencies between multiple technical, social and institutional changes. Organizational set-ups, cultures and funding mechanisms in AR4D establishments, however, tend to pose barriers to working in a truly systemic manner from the outset[1] (Hall et al., 2003;Leeuwis et al., 2017; Woltering, 2024) and as a result many AR4D interventions are still focused on 'scaling' single solutions at a relatively late stage of AR4D trajectories. Taking this reality as a starting point, we currently see efforts to reconceptualize the very notion of scaling in AR4D, and to develop and operationalize more systemic perspectives on scaling (Wigboldus et al., 2016; Wigboldus and Brouwers, 2016; Sartas et al., 2020; Woltering et al., 2024; Barrett et al., 2022). A key characteristic of such approaches is that they look at packages or bundles of interdependent (technical, social and institutional) innovations that somehow need to scale together. The notion of responsible scaling, then, has been introduced to further broaden the scope of systemic perspectives on scaling, emphasizing the need to adapt innovation bundles and packages to diverse social groups and different ecological contexts. The aim is to anticipate, and mitigate possible unintended negative consequences of scaling (Gebreyes et al., 2021; McGuire et al., 2024a, 2024b; Sánchez Rodríguez et al., 2021). Responsible scaling therefore requires well-informed decision-making processes that continuously assess what-if scenarios, examine long-term implications, and ensure alignment with broader socioecological systems (Ewell et al., 2026; Ngissah et al., 2026).

Thus, a central tenet of responsible scaling is the capacity to anticipate how innovation and scaling may provide advantages or disadvantages for different social groups, and how these differences shape access, use, and benefits. Research demonstrates that the intersectionality of factors such as gender, education, age, location, and access to networks can determine who benefits from scaling initiatives and who risks exclusion (McGuire et al., 2022; McGuire, Leeuwis, et al., 2024; Ngissah, 2024; Sartas et al., 2020; Schut et al., 2020). Thus, innovation scaling may often produce uneven outcomes; for instance, male farmers with larger landholdings are more likely to benefit from digital agricultural tools than women with smaller plots (Ngissah et al., 2025).

1.2. Scaling decision-support tools and GenderUp

SDSTs have been developed within AR4D organizations including USAID, IDRC, World Food Programme and CGIAR to support these responsible processes (McGuire, Leeuwis, et al., 2024; Sartas et al., 2020; Ewell et al., 2026, Ngissah et al., 2026). In this paper, SDSTs refer to tools developed within the CGIAR system and its partners. SDSTs

[1] This will imply starting ideation with multi-stakeholder analysis, setting R&D agendas/priorities together, co-development of innovations etc. (Ngissah, Kihoro et al., 2026, Ewell et al., 2026)

provide structured methods and approaches that help innovation teams make systematic, context-sensitive, and socially informed decisions during the scaling process. These tools encompass a growing range of approaches; including diagnostic, methodological, and digital tools that assist innovation teams in assessing scaling potential from a systems perspective, identifying bottlenecks, and guiding strategic decision-making (McGuire et al., 2024a; Sánchez Rodríguez et al., 2021; Sartas et al., 2020; Woltering et al., 2024). For example, *Scaling Readiness* provides a structured methodology for assessing innovation maturity, diagnosing bottlenecks, and supporting adaptive management of scaling processes (Sartas et al., 2020). Similarly, the *Scaling Scan* offers a practical diagnostic framework that evaluates key 'scaling ingredients', such as demand, business models, and institutional support, to inform context-specific scaling strategies (Woltering et al., 2024). Moreover, emerging digital and AI-based tools increasingly complement these approaches by enhancing data integration, forecasting, and communication within innovation systems (for an overview, see Letty et al., 2022; Sánchez Rodríguez et al., 2021). Although these tools have strengthened attention for interdependencies among technical and social components in innovation packages, they have been critiqued for focusing primarily on the scalability and readiness of such packages, without sufficient consideration of social differentiation, gendered impacts, and unintended consequences across diverse user groups (McGuire et al., 2024a).

To direct attention toward such impacts and consequences, GenderUp (McGuire et al., 2024a, 2024b) was designed as a dialogue-based SDST to support responsible scaling. It seeks to provide a structured, conversational process through which innovation teams identify relevant dimensions of social diversity, examine intersectional vulnerabilities, and anticipate potential trade-offs. In doing so, it seeks to move beyond generic categories (e.g., women or youth) by identifying context-specific at-risk groups and iteratively adapting innovation packages and scaling strategies to mitigate exclusionary outcomes (McGuire et al., 2024a, 2024b). Such innovation packages and strategies may well include components geared toward changing discriminatory gender norms and addressing structural inequalities in society, and thus align to a considerable degree with principles advocated as part of Gender Transformative Approaches (GTA) in international development (MacArthur et al., 2022). Thus, while responsible scaling and GenderUp derive inspiration from both systems approaches and GTA, the chosen point of entry (i.e. particular scaling initiatives) is likely to narrow the scope to some degree. At the same time, they can be seen as attempts to operationalize systems approaches and GTA in specific settings.

1.3. Research gap, significance and approach

Despite the proliferation of SDSTs in the CGIAR including GenderUp, empirical assessments of their practical effects remain limited. Whereas considerable research has explored the conceptual value and methodological strengths of some SDSTs for intervening in complex natural, social, and economic systems (McGuire et al., 2022; Sánchez Rodríguez et al., 2021; Sartas et al., 2020), far fewer studies have examined their actual outcomes (Damtew et al., 2023). Understanding whether SDSTs influence scaling decisions is therefore important for assessing their contribution to responsible scaling. It is equally important to examine how and under what conditions they influence decision-making, including whether the mechanisms through which they are expected to work are activated in practice (Damtew et al., 2023). The gap between how decision-support tools are designed, used and achieve impact has long been documented (McCown, 2002; Rose et al., 2016), but has received limited attention in research; particularly for tools that support collective decisions about social inclusion. Addressing this gap can contribute to responsible scaling scholarship and agricultural decision-support research by providing insights into how the intended functions of SDSTs translate into actual use and influence decision-making.

This is particularly relevant to participatory and gender-transformative approaches, as participation and reflection alone do not guarantee changes in decisions, practices, gendered power relations or institutional arrangements (Cornwall, 2003; MacArthur et al., 2022). Understanding how these processes unfold is important for assessing whether tools intended to support social inclusion influence scaling decisions and practice.

This study addresses this gap through a realist evaluation (RE) of GenderUp. It investigates whether and how the tool contributes to developing and institutionalizing socially responsible and gender-responsive scaling strategies. Empirically, the study uses the scaling strategy development and deployment process of the high-quality cassava peel (HQCP) feed innovation in Nigeria as a case study. It draws on multiple data sources including stakeholder interviews, focus group discussions (FGDs), analysis of GenderUp documentation, and observations of a GenderUp HQCP workshop and of HQCP's implementation inception meeting. The evaluation combines empirical evidence about the HQCP project's scaling strategies before, during, and after GenderUp workshop outputs, with field-level perspectives from Oyo State, Nigeria, where the HQCP project is being implemented. The study addresses three research questions:

1. Through which mechanisms is GenderUp expected to foster reflexive and socially inclusive scaling decisions?
2. How, why and under what contextual and institutional conditions are these mechanisms realised, enabled or constrained in practice when GenderUp is applied to the HQCP initiative in Nigeria?
3. What does the HQCP case reveal about the factors shaping the uptake of GenderUp's recommendations, and what are the implications for the design, facilitation and governance of dialogue-based SDSTs supporting responsible scaling in AR4D?

In answering these questions the study make three contributions; First, we elicit and examine the underlying mechanisms of change through which GenderUp is expected to influence scaling processes. We show that GenderUp provides coherent mechanisms of change for fostering critical reflection and the co-development of socially inclusive scaling strategies. In doing so, it addresses the fact that these mechanisms in the case of many SDSTs including GenderUp are not formally articulated and demonstrates how such mechanism of change can be made explicit to support the evaluation of dialogue-based SDSTs. Second, it provides a case-based explanation of how these mechanisms were realised, enabled or constrained in practice within the HQCP initiative, showing how socially focused decision support interacts with stakeholder participation, learning and institutional conditions in responsible scaling. Third, it identifies implications for the design, facilitation and governance of SDSTs and the scaling initiatives in which they are used. We generate lessons for refining and broadening the application of socially inclusive scaling SDSTs to support sustained social transformation in AR4D and contribute to the SDGs. Together, these contributions offer empirically grounded insights for scholarship on responsible scaling and gender-transformative approaches, as well as for SDST developers and facilitators, innovation teams, donors and consortium partners.

The remainder of the paper is structured as follows. In section 2, we justify the use of RE as the analytical framework and outline our methodological approach, including the rationale for selecting the HQCP initiative as a case study and our data collection and analysis strategy. We then present the empirical findings, including both the elicited mechanisms of change underlying GenderUp and their empirical evaluation within the Nigerian HQCP case. These findings are discussed in relation to responsible scaling scholarship, while drawing key lessons and recommendations for strengthening socially inclusive decision-support tools in AR4D.

2. Analytical and methodological framework

This section outlines the analytical framework that underpins the study and guides the empirical approach. First, it presents the RE

analytical framework. Second, it describes the research design, including the methodology, its justification, and contribution to addressing the research objective. Third, it outlines the process of data collection and analysis.

2.1. Realist evaluation (RE) as analytical framework

The outcomes of agricultural innovation and/or scaling processes within AR4D rarely follow linear cause–effect pathways. Instead, they unfold within complex, adaptive systems shaped by multiple interacting factors—such as institutional arrangements, market incentives, partnerships, and sociocultural norms—that influence both decision-making and outcomes, and hence how SDSTs are appreciated and used, as well as the outcomes that they produce (Damtew et al., 2023; Douthwaite and Hoffecker, 2017; Douthwaite et al., 2017a, 2017b). This poses considerable challenges to the evaluation of SDSTs such as GenderUp. In response to similar challenges in other fields, Pawson and Tilley (2004) have developed RE as conceptual and methodological perspective for assessing interventions. This approach is rooted in critical realism (Parra, 2023) and has been successfully adapted to and applied within AR4D settings (Douthwaite et al., 2017a, 2017b; Hussain and Brown, 2018; Damtew et al., 2023). In contrast to conventional evaluations such as those conducted through randomized controlled trials, RE seeks not only to determine whether an intervention works, but also to explain *how*, *why*, *for whom*, and *under what contextual conditions* it does so (Pawson and Tilley, 2004; Kaboub, 2004). Central to RE is the concept of *mechanisms*, understood as underlying social and cognitive processes—such as sense-making, learning, motivation, or shifts in awareness and reasoning—that are triggered when actors engage with an intervention (in this case, a GenderUp process) under specific contextual conditions. Mechanisms are thus not understood as mechanical forces or linear causal drivers. Outcomes emerge not because an activity 'causes' change directly, but because it activates particular mechanisms within particular contexts.

These relationships are commonly expressed as context–mechanism–activity outcome (C–M—O) configurations (adapted to study SDSTs as activity-mechanism-activity outcome; AC-M-AO) (Damtew et al., 2023), which help explain why similar interventions may lead to different outcomes across settings. In AR4D contexts, RE has proven especially valuable for unpacking the social, institutional, and behavioral dynamics underpinning innovation and scaling processes (Esterhuyse, 2019; Kruijssen et al., 2021). Unlike traditional evaluation approaches that focus on quantifiable outputs or post-hoc assessments, RE accommodates change complexity and supports iterative learning. By developing and empirically testing *middle-range mechanisms of change*, RE enables evaluators to identify how mechanisms such as motivation, institutional incentives, and collective learning interact with contextual conditions to shape outcomes (Esterhuyse, 2019; Kruijssen et al., 2021).

RE can be characterized as a theory-driven approach as it requires that those intervening articulate a 'program theory' or 'theory of change' that species (ideally using the AC-M-AO formula) why and how and intervention program is supposed to work and result in outcomes (see e.g. Funnell and Rogers, 2011; Maru et al., 2018; Douthwaite et al., 2017). Research is then used to test and potentially refine the program theory in an iterative research cycle.

Evaluating programs or interventions such as SDSTs can be challenging because developers often do not explicitly articulate what the assumed mechanisms of change are (Maluka 2011; Hussain and Brown, 2018 Pawson, 2000). This is also true for GenderUp (McGuire et al., 2024a, 2024b). Without an explicit account of how specific activities are expected to influence decision-making and practice, learning and improvement remain limited as evaluators can neither determine nor demonstrate whether an SDST is producing its projected outcomes and under what conditions it does so, to inform improvements. As SDSTs do not operate through simple or linear causal chains, the generative processes through which such tools are assumed to influence actors and/or actions within particular contexts need to be reconstructed (elicited or drawn out) and investigated. Consequently, evaluators must elicit and empirically test plausible hypotheses about how SDSTs work in practice, rather than relying on pre-defined or deterministic causal models (Esterhuyse, 2019; Kruijssen et al., 2021). As is elaborated below, we will elicit GenderUp's underlying theory of change as a first empirical step in this study, and in doing so also provide a more detailed description of this particular SDST.

2.2. Study design and research methodology

To evaluate GenderUp from a RE perspective, this study adopts an in-depth qualitative case study approach. We examine the use of GenderUp in developing and deploying a scaling strategy for a HQCP innovation in Nigeria within an AR4D context. Drawing on qualitative social research traditions (Habermann et al., 2021; Ponelis, 2015), the study employs FGDs, workshop observations, document analysis, and semi-structured interviews to explore how stakeholders interpret and engage with the GenderUp process. Combining such methods enables a nuanced and multi-facetted understanding of the mechanisms that may be activated during the GenderUp process, the outcomes that result from these, and how both are shaped by contextual conditions. Qualitative social research is particularly suited to gaining insight in complex and evolving processes such as innovation scaling in AR4D, where outcomes are shaped by diverse perspectives and contextual dynamics (Habermann et al., 2021). The approach aligns with previous studies applying RE in similar contexts (Damtew et al., 2023; Esterhuyse, 2019; Kruijssen et al., 2021).

We selected Nigeria's HQCP initiative as an empirical test case for GenderUp because, prior to the workshop, it possessed a fully developed scaling strategy allowing for a comparative assessment of how GenderUp contributed additional insights or adjustments beyond the project's existing framework. The initiative forms part of a broader poultry-sector program implemented by a consortium of international organizations, providing a rich context to explore how institutional arrangements and partnership dynamics influence the uptake and implementation of GenderUp recommendations. The initiative's goals are to reduce environmental waste and strengthen women's economic position. By transforming cassava peels—an agricultural by-product often discarded by women processors—into livestock feed, the initiative addresses environmental waste challenges while creating new economic opportunities for women. The focus on social inclusion makes the initiative an ideal case to assess GenderUp's effectiveness in promoting equity-oriented and contextually responsive scaling strategies.

Multiple qualitative methods were selected for complementarity in examining different dimensions of the GenderUp process. Workshop observations enabled real-time analysis of interactions, deliberations, and decision-making processes during both the GenderUp workshop and HQCP consortium meetings. Given the dialogical and participatory design of GenderUp, direct facilitation and observation within these workshops provided the most appropriate means of examining how participants engaged with the method, negotiated interpretations, and collectively developed scaling strategies. Observations of consortium meetings further enabled analysis of how stakeholder interactions and institutional dynamics shaped implementation decisions. Semi-structured interviews provided in-depth insights into participants' interpretations and experiences, and FGDs facilitated collective reflection and validation of emerging findings. This combination of methods provides the analytical depth and contextual sensitivity required in RE to gain insight in the interaction between mechanisms, outcomes and contextual conditions. Although qualitative approaches do not aim for statistical generalizability, they provide the analytical depth and contextual sensitivity required to examine complex and context-dependent scaling processes in AR4D settings.

2.3. Data collection and analysis

Data collection and analysis followed an RE logic (Pawson, 2000), drawing on Damtew et al.'s (2023) approach to explain whether, how, why, and under what conditions GenderUp influenced responsible scaling decisions. Consistent with RE, we did not seek to establish linear cause–effect relationships; rather, we sought to examine how GenderUp activities interacted with contextual, institutional, and partnership conditions to enable or constrain mechanisms of change (Esterhuyse, 2019; Gilmore et al., 2019). Here, we describe in detail the three phases of the research.

Phase One: Eliciting GenderUp's mechanisms of change.

The research began by eliciting the mechanisms of change through which GenderUp is expected to influence scaling practices. This step was necessary because, as emphasized in the RE literature (Hussain and Brown, 2018; Esterhuyse, 2019), interventions and decision-support tools often do not explicitly articulate how their activities are assumed to produce change. Evaluators and researchers are therefore tasked to first elicit these to guide their assessments (Esterhuyse, 2019; Kruijssen et al., 2021). GenderUp documents (including a published paper, workshop facilitation documents, and a report on GenderUp), iterative interviews with a GenderUp developer, and workshop observations were therefore analyzed to identify the tool's stages, intended activities, underlying assumptions, and anticipated outcomes. The initial elicitation was conducted by the lead researcher through a review of the documents and subsequently refined through iterative discussions with the third author,[2] who was involved in GenderUp's development. The elicitation involved making explicit what was implicitly assumed to be the mechinism of change. This resulted in a final framework describing how its five sequential stages; (1) defining the innovation and current scaling strategy, (2) exploring diversity, (3) understanding intersectionality, (4) mitigating unintended consequences, and (5) institutionalizing equity and their corresponding mechanisms of change are expected to shape responsible scaling strategies and practices in AR4D settings. These elements were synthesized into AC–M–AO propositions. The elicited mechanisms presented in section 3.1 served as analytical heuristics for subsequent observation and interpretation.

Phase Two: Assessing GenderUp's change mechanisms in practice through the HQCP case.

The elicited AC–M–AO propositions informed the subsequent data collection and analysis by guiding attention to moments where GenderUp activities appeared to stimulate reflection, shift understanding, or influence decision-making processes. To evaluate how these mechanisms operated in practice, an in-depth case study was conducted using the HQCP feed initiative in Nigeria.

In realist evaluation, mechanisms are not directly observable; they are inferred from how participants reason and respond to an intervention within particular contexts (Pawson and Tilley, 2004). Observing the GenderUp workshop therefore provided an opportunity to examine these processes as they unfolded, including how participants engaged with the tool's prompts, reconsidered assumptions, negotiated interpretations and developed scaling strategies. This approach was particularly appropriate given GenderUp's dialogical and participatory design, in which interaction and collective knowledge production during workshops are central to the intervention. Workshops can also serve as a research method in their own right (Ørngreen and Levinsen 2017).

However, workshop discussions reflect the perspectives and social relations of those present, and participants' expressed intentions do not necessarily translate into subsequent action (Mosse, 1994; Cornwall, 2003). Workshop outputs were therefore triangulated with end-of-day interviews, comparisons of scaling strategy documents produced before and after the workshop, and a follow-up field phase conducted 1 year 6 months later (Phase Three). This combination of data sources enabled us to examine both participants' responses during the workshop and the extent to which the resulting recommendations were reflected in subsequent strategies and implementation.

Researchers trained as GenderUp facilitators conducted the GenderUp workshop for the HQCP project scaling team (see Table 1 for an overview). This workshop was spread over several days, two days at the start, then another day three weeks later as a follow-up. This follow-up was organized to embed the newly designed inclusive scaling strategies in project management and operational plans. It is assumed that, by this time, the team had produced a revised inclusive scaling strategy document that was systematically compared with the pre-GenderUp version to assess changes potentially stimulated by the first two days of the

Table 1
Profile of HQCP team members that participated in GenderUp workshop

Participant profiles	Number	Background of HQCP team members and their roles on the team
International Livestock Research Institute (ILRI) researchers involved in HQCP development	3 (1 female, 2 males)	They are core members of the HQCP research and scaling team with over 10 years' experience in product development, field trials, and piloting
Agricultural finance expert (Government agency)	1 male	He is a member of the HQCP scaling team, works for a government agency in Nigeria, and is an expert in financial service access and resource mobilization for scaling in Nigeria
Monitoring and evaluation specialist (ILRI)	1female	She works with ILRI in monitoring and evaluation; responsible for monitoring and evaluating innovation scaling within ILRI programs; member of the HQCP scaling team
Country Director (International NGO, experts Nigeria)	1 male	He led the establishment of initial HQCP pilot centers across Nigeria; active member of the HQCP scaling team
Scaling and innovation portfolio (ILRI)	2 males	They are system-level scaling experts with extensive experience across multiple ILRI innovations; members of the HQCP scaling team
Gender specialists (ILRI)	2 females	They provided technical guidance on gender integration and inclusive scaling; one served directly on the HQCP scaling team
Technology transfer officer (ILRI)	1 male	He is a long-term researcher involved in HQCP piloting and knowledge transfer activities; HQCP team member
National program coordinator (Cassava Producers Association, Nigeria)	1 female	She is a specialist in the cassava value chain with broad national experience in Nigeria; contributed contextual insights to the scaling process
Regional Director, Government extension	1 female	Partnered in establishing early HQCP processing centers nationwide; key institutional stakeholder in the scaling team

[2] The interviews with the GenderUp developer were conducted as part of the elicitation process. Following an initial review of GenderUp documentation and a preliminary reconstruction of the tool's proposed mechanisms of change by the lead author, the elicited mechanisms for each stage were presented to the developer for reflection, clarification, and critical assessment. At the time of the interviews, these mechanisms had not been formally articulated within GenderUp documentation and were therefore not treated as pre-existing constructs. Rather, the interviews served to examine whether the reconstructed mechanisms reflected the intentions and assumptions underpinning the tool. The first interview was recorded and transcribed and informed further refinement of the elicited mechanisms. An interview with the same developer was subsequently conducted to further clarify, refine, and validate the final set of mechanisms of change presented in this study.

workshop. All sessions were recorded, transcribed, and analyzed. During the workshop, several GenderUp materials were used, including an online whiteboard, which contained structured prompts and reflective exercises that guided participant discussions and reflections. Surveys designed as part of the GenderUp process were used to collect participants' prior knowledge on a topic and their reflection after each stage on this topic and its application to the innovation being scaled. Brief follow-up interviews were conducted with members of the HQCP scaling team at the end of each day to capture perceived insights and shifts in understanding.

During this phase, a structured analysis was conducted to examine whether and how the elicited mechanisms were activated within the HQCP case. Workshop observations, transcripts, whiteboard outputs, survey responses, follow-up interviews, and project documents produced before and after the GenderUp workshop were analyzed to trace sequences linking GenderUp activities to observed changes—or lack thereof—in scaling strategies and decision-making processes. Particular attention was paid to instances where participants articulated previously implicit assumptions, identified socially differentiated groups, recognized intersectional risks, or proposed adaptations to existing scaling strategies. Observed patterns were then compared with the anticipated outcomes embedded in the elicited mechanisms to evaluate the extent to which the mechanisms influenced change in the specific contexts. Data coding initially followed a deductive approach guided by the elicited mechanisms while remaining open to inductive themes and contextual dynamics emerging from the data. This enabled us to analyze how GenderUp's mechanisms interacted with institutional, organizational, and partnership conditions over time.

Phase Three: Field validation and contextual assessment.

To assess how GenderUp recommendations were translated into practice beyond the workshop setting, a final phase of data collection and analysis was conducted one year and six months after the last day of the GenderUp workshop. This phase helped assess the contextual validity of GenderUp outputs by examining alignment and divergence between workshop-generated recommendations and field realities in Oyo State. Particular attention was paid to how donor requirements, consortium governance structures, and role compartmentalization shaped the implementation of GenderUp recommendations in practice, because previous REs of AR4D projects and decision-support tools have shown that institutional and governance conditions may shape, constrain, or enable the implementation of tool-generated recommendations (Maluka et al., 2011; Hussain and Brown, 2018; Esterhuyse, 2019; Kruijssen et al., 2021). Additional data were collected through fieldwork in Oyo State, one year and six months after the workshop. This involved in-depth individual interviews and FGDs with project managers, researchers, and actors across the HQCP value chain—including poultry farmers, processors, fabricators, and cassava processors (Table 2). The interviews and FGDs were designed to explore participants' experiences with the HQCP scaling process, perceptions of barriers and opportunities affecting innovation uptake, the implementation of GenderUp recommendations, and the contextual, institutional, and partnership factors that enabled or constrained socially inclusive scaling outcomes. We attended, participated, observed, and took notes at the five-days long consortium meetings, where the implementation plans were finalized. These data enabled triangulation between the documented scaling strategies and field-level realities, revealing how GenderUp outcomes aligned or diverged from local socioeconomic and institutional contexts.

Table 2
Overview individual interviewees and focus group participants from stakeholder groups across the HGCP value chain

Stakeholder group	Number of participants and interviews	Description of participants and rationale for selecting them
Focus group discussions		
Female poultry farmer members of Poultry Farmers Association of Nigeria (PAN)	5 (1 FGD)	They are the primary intended beneficiaries of the HQCP technology. All interviewed participants were trained on, and previously used, HQCP but later discontinued. Their experiences and perceptions are critical for understanding barriers to adoption and the gender-differentiated vulnerabilities of association members.
Male poultry farmers (PAN members)	4 (1 FGD)	They were also trained users who discontinued using HQCP after some time. They provided comparative perspectives to women farmers, highlighting gendered differences in resource access, decision-making, and risk perception in feed adoption.
Individual interviews		
Individual non-member farmers (non-users)	2 (1 male, 1 female)	These farmers had heard of HQCP but never used it. They offered insights into non-adoption dynamics and the vulnerabilities of farmers outside organized associations.
Cassava processors (women and youth)	1 male	He was an active cassava processor and was also trained on HQCP processing but not currently processing HQCP. As a primary owner of cassava peel, he represented key upstream stakeholders in the value chain whose perspectives inform feasibility and incentives for process adoption.
Active HQCP processors	4 (2 male, 2 female)	They are trained and active HQCP processors. They are central actors in HQCP scaling. Their experiences and technical challenges directly influence the innovation's potential to scale across the value chain.
Trained but inactive HQCP processors	2 (1 male, 1 female)	They were trained and processed HQCP in the past but not engaged in current production. They provided insights into attrition and barriers preventing sustained engagement after initial training.
Fabricators	2 males	They fabricate the machines for converting cassava peel into HQCP and provide technical insights into HQCP equipment design and functionality. Both have extensive experience (7 and 10 years) building processing machines and understanding mechanical and maintenance constraints.
Researcher	1 male	He was involved in the research that developed HQCP, also in the first pilot testing of HQCP. He has had over

(*continued on next page*)

3. Results and analysis

This section first presents a description of the various stages of GenderUp and the elicited mechanisms of change underlying each stage of the GenderUp approach, followed by an assessment of the mechanisms in the HQCP initiative, focusing on how the GenderUp processes shaped reflection, decision-making, and revisions to the project's scaling strategy. Drawing on workshop data and comparative analysis of scaling

Table 2 (*continued*)

Stakeholder group	Number of participants and interviews	Description of participants and rationale for selecting them
		10 years engagement with stakeholders involved in the innovation and offers historical and technical perspectives on prior scaling attempts.
NGO representative	1 male	He is a member of HQCP scaling team and previously involved in supporting women-led processing centers and provides perspectives on gender mainstreaming and field-level facilitation.
Government extension officer (Director)	1 female	She is a member of HQCP scaling team with prior involvement in female-led HQCP centers and provides insights into public-sector coordination and extension linkages.
Director, Feed Regulation Institute	1 male	He has over seven years as director of Nigeria's feed regulatory body and offers perspectives on feed standards, quality assurance, and policy constraints affecting HQCP adoption.
Feed millers	2 (1 female. 1male)	They are active users of HQCP as a feed ingredient. Their perspectives inform quality requirements, farmer expectations, and concerns regarding feed adulteration and market acceptance.

documents before and after the workshop, we identify the new insights and recommendations that emerged through the GenderUp process. We show the extent to which these workshop-generated recommendations reflected field realities and how they were translated into early implementation processes. Particular attention is paid to how consortium arrangements, donor requirements, institutional structures, and power dynamics influenced the operationalization of GenderUp recommendations and their alignment with socioeconomic and institutional realities in Oyo State, Nigeria.

3.1. Description of GenderUp's stages and elicited mechanisms of change

Table 3 presents a summary of the GenderUp process and the elicited mechanisms of change through which the tool is expected to influence inclusive scaling strategies. Drawing on an analysis of GenderUp documentation, workshop materials, interviews with a GenderUp developer, and observations of how the process was structured and facilitated, Table 3 synthesizes the elicited AC–M–AO propositions underlying the GenderUp approach, illustrating how specific workshop activities are expected to stimulate cognitive and social mechanisms that contribute to more socially inclusive and context-responsive scaling strategies.

GenderUp is designed as a five-stage conversational and reflective process intended to support innovation teams in critically examining scaling strategies from an inclusion and responsible scaling perspective. According to GenderUp documentation and interviews with a developer, the process begins with explicating the scaling ambition and current scaling strategy, where participants articulate current scaling goals, activities, partnerships, and assumptions underpinning the intervention (interview with developer; McGuire, Leeuwis, et al., 2024). According to McGuire et al. (2024a, 2024b) and the developer interviewed, the primary objective of this stage is to clarify the innovation being scaled, the goals associated with it, and the scaling strategy currently guiding the

Table 3
Elicited mechanisms of change underlying GenderUp

GenderUp stages	Activities (AC)	Mechanisms (M)	Expected outcomes (AO)
1. *Explicating the scaling ambition and current scaling strategy*	Guided discussion on pre-existing scaling strategies, followed by teams' exploration of questions on the online whiteboard	Through better understanding of what it takes to achieve goals and greater appreciation of the hidden assumptions underlying those activities	Leads to teams becoming more aware of their current/pre-existing scaling strategy (its components) and being able to articulate more coherently what they have been trying to achieve and how
2. *Exploring social diversity*	Sensitization of the team on gender and relevant diversity, followed by teams partaking in a survey on relevant social dimensions and then engaging with whiteboard questions on access and consequences for individuals and communities	Through the systematic exploration and better understanding of diversity dimensions	Leads to the team becoming more aware of the relevant dimensions of diversity— i.e., the segments that may not be able to access the innovation and/or may face negative effects
3. *Understanding intersectionality*	A training session on the intersectionality module and its relevance for agricultural innovations, followed by team engagement with whiteboard questions to untangle the intersectional puzzle of their innovation	Through a deeper appreciation of how vulnerabilities arise and become compounded within the scaling process	Leads to a better understanding of at-risk intersectional profiles who might face harm or exclusion and of the risks, harms, or negative consequences involved
4. *Mitigating unintended consequences*	Teams revisit all components of the current scaling strategy on the whiteboard and ask how each element could be modified to prevent likely negative consequences and/or leverage opportunities for positive effects	Through a deeper appreciation of the complexity surrounding the innovation, at-risk intersectional groups, and intended beneficiaries' socioeconomic circumstances	Leads to teams identifying and developing inclusive, opportunity-oriented, and risk-mitigating strategies
5. *Institutionalizing equity*	Teams explore the questions on the whiteboard, reflecting and documenting how strategies need to be translated toward adaptation into project management and organizational processes as well as resource allocations, structures, etc.	Through greater appreciation of the practical implications of embedding equity in projects in ways that ensure accountability and support tailored change	Leads to teams implementing equity-oriented scaling strategies

intervention.

This stage was developed in response to the observation that scaling processes in AR4D are often shaped by tacit assumptions influenced by institutional lock-ins, established innovation trajectories, donor priorities, and existing partnership arrangements (McGuire et al., 2022; McGuire, Leeuwis, et al., 2024; Sartas et al., 2020; Wigboldus et al., 2020). Consequently, as noted by McGuire, Leeuwis, et al. (2024): "*teams may make both explicit and implicit choices about critical scaling elements—including why they are pursued, for whom, by whom, and under what circumstances.*"

In this sense, stage 1 is intended to make such implicit assumptions visible and open to critical reflection. This was further emphasized by the developer interviewed, who explained that:

Scaling strategies are often shaped by assumptions that remain 'hidden' but are very important for the scaling process. Our aim for the first stage was to help teams better articulate their strategies, goals, and the assumptions through which they aim to achieve these goals […] moreover their existing scaling strategy serves as a starting point for making changes and also later serves as benchmark for assessing progress.

The elicited mechanism underlying this stage therefore suggests that structured reflection on existing strategies enhances participants' understanding of their own scaling pathways, operational assumptions, and implementation gaps, thereby enabling clearer articulation of scaling objectives and processes.

The second stage focuses on exploring diversity by guiding participants to identify the resources, capabilities, and conditions required to access and benefit from the innovation. According to the developer interviewed, participants are encouraged to reflect on how dimensions such as gender, age, education, wealth, disability, geography, and institutional affiliation shape differential access to these resources and opportunities. This is also reflected in the GenderUp workshop materials, which prompt teams to consider "*what social dimensions (e.g. wealth status, education, religion…) are most influential in determining access to the resources required for using the innovation*" *(McGuire, Leeuwis,* et al*., 2024)*.

The developer further explained that: "*The purpose is to move teams beyond simplistic framings such as men versus women and instead encourage reflection on how multiple social dimensions interact to shape innovation access and use*".

The elicited mechanism at this stage could therefore be summarized as through a systematic reflection on diversity; participants broaden their understanding of social differentiation beyond generalized categories and increase awareness of how access constraints vary across groups and contexts.

According to the developer interviewed, the purpose of stage 3 is to build on the exploration of diversity initiated in stage 2 by guiding innovation teams to identify at-risk intersectional profiles relevant to equitable access, use, and benefits of the innovation. As described in the GenderUp documentation, "*GenderUp guides teams through a process that helps identify at-risk intersectional profiles*" (McGuire, Leeuwis, et al., 2024).

McGuire, Leeuwis, et al. (2024) further explain that "*at-risk means that they are in danger of being negatively impacted by the innovation or the scaling process*".

Through this process, innovation teams identify individuals or groups who may be disproportionately affected by scaling activities and examine how overlapping social dimensions shape exposure to risks or opportunities. The purpose is not only to identify vulnerable groups, but also to understand "*how vulnerabilities overlap and reinforce each other within particular scaling environments*" *(McGuire, Leeuwis,* et al*., 2024)*.

The elicited mechanism underlying this stage is that reflecting on intersecting dimensions of disadvantage deepens participants' appreciation of how vulnerabilities are compounded within particular social and institutional contexts, thereby encouraging more nuanced understandings of inclusion and exclusion. Building on the identification of at-risk intersectional profiles, stage 4 focuses on revisiting and redesigning existing scaling strategies to mitigate potential negative consequences and strengthen inclusive outcomes. According to the developer interviewed, the objective of this stage is to co-develop revised scaling strategies that integrate relevant diversity dimensions, prioritize identified at-risk groups, and introduce mitigation measures that improve equitable access to innovation resources and opportunities. Our analysis of workshop guides and facilitation materials[3] revealed that discussions during this stage focus on identifying complementary innovations, adapting communication and training strategies, revising partnership arrangements, and introducing safeguards against exclusionary outcomes. Earlier GenderUp pilots documented similar responses, where innovation teams, after recognizing exclusionary dynamics, "*responded by incorporating mitigating activities and more user-specific complementary innovations into their strategy*" (McGuire, Leeuwis, et al., 2024).

The elicited mechanism underlying this stage is therefore that increased awareness of socially differentiated risks stimulates critical reflection on scaling trade-offs and encourages teams to redesign scaling strategies to better address exclusionary dynamics and unintended consequences. Finally, our analysis of the GenderUp process revealed that the fifth stage focuses on institutionalizing equity by embedding revised inclusive strategies into project planning, implementation, and governance arrangements. According to the developer interviewed, this stage is typically conducted some weeks after the initial workshop and is intended to support teams in integrating inclusive scaling considerations into operational systems, partnerships, communication plans, and implementation processes. The GenderUp documentation explicitly states that the objective of this stage is to "*streamline GenderUp discussions and new gender responsible scaling activities into the work plans and management tools of the scaling project*" (McGuire, Leeuwis, et al., 2024).

The elicited mechanism underlying this stage is that the team gains greater appreciation of the practical implications of embedding equity in projects through structured reflections, leading the team to institutionalize equity in the implementation of its project.

Collectively, these stages reflect GenderUp's broader theory of change: that structured dialogue, critical reflection, and iterative examination of diversity and intersectionality can reshape scaling strategies toward more responsible and socially inclusive outcomes. We find that the mechanisms mobilized in the GenderUp methodology, and as described in the previous sections, illustrate a strategy aimed at improving processes through structured dialogue and cumulative explication of underlying assumptions and potential blind spots generated by the scaling team's composition. It delivers on its aims of providing a methodology that is highly flexible to the context but that can also generate concrete guidance. Thus, it deviates from causal models that prescribe a recipe for generating particular scaling outcomes, and instead aligns with a process-oriented methodology for generating context-specific insights and recommendations. In the following section, we describe how this methodology is translated into practice and some of the potential challenges and opportunities in the realization of practical outcomes in this process-oriented approach.

3.2. Assessing GenderUp's mechanisms of change in the HQCP initiative

This section examines how the GenderUp approach supported the HQCP initiative's scaling team in developing a responsible scaling strategy. The analysis draws on data from observing the GenderUp workshop conducted with the HQCP scaling team, interviews with the scaling team after the workshop, and the project's scaling strategy

[3] Internal GenderUp workshop guides and facilitation materials were reviewed as part of the document analysis, but because these materials are unpublished and not publicly accessible, evidence from document analysis is primarily illustrated using published sources (McGuire, Leeuwis, et al., 2024) and developer interviews.

documents prepared both before and after the workshop. The analysis presented in this section illustrates the elicited AC–M–AO propositions presented in Table 3 to examine how GenderUp shaped reflection, decision-making, and scaling strategy development during the HQCP workshop.

3.2.1. Using HQCP to test stage 1's mechanism of change: Understanding scaling strategies better

In stage 1 of the GenderUp workshop, the HQCP scaling team mapped and discussed its existing scaling strategy using guided whiteboard exercises Workshop observations and comparison of pre- and post-workshop scaling documents showed that this process helped participants articulate previously implicit elements of the strategy, including operational pathways, partnerships, communication plans, and stakeholder roles. Figs. 1 and 2 present enhanced screenshots of the results from the GenderUp whiteboard—illustrating how existing strategies were elicited in this stage, including dimensions used for the elicitation. The images have been enhanced to ensure readability. In Fig. 1, it can be observed that the GenderUp whiteboard guided teams to define their innovation and the context in which they intend to scale this innovation. Fig. 2 shows that the whiteboard further guided teams to elicit planned training around the innovation, planned communication methods and channels, and the corresponding intended beneficiaries of the innovation and the various training needs.

Although most components such as innovation definition and scaling context description were mentioned in the original scaling strategy, the pathways for achieving them were often vague or implicit. Key operational elements, including training needs, communication plans, and partnerships needed to achieve the intended outcomes, were not explicitly defined in the original document but became clearly articulated by the end of stage 1.

One HQCP team member, a technology transfer officer at ILRI explained that stage 1 helped the team better understand what was required to achieve its intended goals, making previously implicit



Fig. 1. Enhanced screenshot of workshop results, stage 1 A: teams defined the innovation as part of existing scaling strategy articulation.



Fig. 2. Enhanced screenshot of workshop results, stage 1B: existing scaling strategy articulation.

assumptions and operational requirements more explicit: "*The first stage really made us think through what exactly was needed to achieve our goals and who needed to be involved. Some things we assumed were obvious had actually never been properly discussed*" (Interview with HQCP team member).

For instance, the pre-GenderUp scaling document contained only a brief reference to media engagement: "*Media monitoring: Track media coverage and social media engagement*". However, a comparison between the pre-workshop document and the GenderUp whiteboard result following stage 1 indicates that the team expanded this section to include specific plans such as collaborating with two local radio stations, producing YouTube videos, utilizing project websites and academic journals, and developing flyers for community outreach.

Similarly, discussions at stage 1 introduced additional scaling considerations: establishing market linkages and engagement with stakeholders, such as community leaders and government extension officers. Interviews with HQCP processors and poultry farmers confirmed that these stakeholders—who were not mentioned in the pre-GenderUp documentation—were critical to project success. By clarifying implicit assumptions and articulating detailed operational components, stage 1 enabled the HQCP team to transition from broad conceptual intentions to a concrete scaling strategy. This structured clarity established the foundation for inclusive reflection in subsequent stages. Arguably, stage 1 fostered critical reflections on scaling strategies and already induced change well before arriving at stage 5.

3.2.2. Using HQCP to test stage 2's mechanism of change: Understanding diversity dimensions better

During stage 2 activities, participants identified six resources considered necessary for effective use of the HQCP innovation: financial resources, technical skills and knowledge, market linkages, utilities, machinery and tools, and land. Using both the survey instrument and guided reflection on the online whiteboard, the team critically examined how different social dimensions such as wealth status, education, land ownership, age, gender, and disability intersect to shape access to these resources. This analysis set the stage for subsequent discussions on potential negative consequences for different user and non-user groups and for communities more broadly.

Our data from Oyo State further demonstrated that the resources identified through the GenderUp workshop were crucial to accessing the HQCP innovation, at least in some part of the country. For instance, establishing HQCP processing facilities requires machinery that can be prohibitively expensive, potentially excluding groups such as young women and youth with limited business experience who may need financial support. Similarly, access to land was recognized as an important factor nationally, with evidence suggesting that gender can constrain land ownership in some regions of Nigeria. Hence, women processors or HQCP entrepreneurs may face difficulties accessing land to establish HQCP processing factories. The importance of these outcomes was later confirmed by the agricultural extension director who elaborated that:

The early version of women only HQCP processing factories that we facilitated failed because of land issues. Initially everything was ok, until we realised that the land in those areas was mostly owned by men, even the one we set up the factory on. The women said it was theirs, but it belonged to their husbands. Now the men took the land after they realised the women were making more money.

Furthermore, our comparison of the pre- and post-GenderUp scaling

documents showed that many of these resources—except market linkages—had already been mentioned in the original strategy. However, although financial resources were frequently discussed, the original document failed to explain how social diversity might influence people's ability to access these resources. When gender was mentioned, it was generally framed in binary terms (men and women or young men and women), without deeper consideration of intersectional factors. Inception meeting observations and interviews with HQCP team members further suggested that such phrasing reflected donor reporting requirements rather than deliberate analytical intent. For example, we observed at the inception meeting that other consortium members reported similarly, with young women vs. young men. When queried, several members explained that the donor's emphasis on targeting such groups made such reporting mandatory.

One HQCP team member—country director of an international NGO—highlighted how GenderUp expanded his understanding of gender and diversity beyond binary categories:

I realised gender is not just men versus *women. Young women with lower levels of education, young men with no business experience, and persons with disabilities all have different vulnerabilities—and these definitely cannot be treated equally. That's not how I used to understand gender, certainly not before the GenderUp workshop*.

Stage 2 thus equipped the team with a nuanced understanding of diversity dimensions and their implications for scaling. It bridged the gap between gender awareness and intersectional analysis, setting the stage for the identification of specific at-risk profiles in stage 3.

3.2.3. Using HQCP to test stage 3's mechanism of change: Deeper appreciation of vulnerabilities

Building on the diversity dimensions identified stage 2 results, in stage 3 the scaling team focused on exploring how overlapping social characteristics could shape unequal access to, and differentiated effects from, the scaling of HQCP. Through guided GenderUp reflection questions and online whiteboard discussions, participants examined how dimensions such as gender, education, wealth status, disability, marital status, and geographic location intersected to produce varying forms of vulnerability.

Workshop observations and whiteboard outputs indicated that participants moved beyond broad social categories such as women and youth to construct more differentiated at-risk profiles representing social groups most likely to face exclusion or negative consequences in the absence of deliberate inclusive interventions. These included, for example, younger women with low levels of education, people without prior business experience, rural women living with disabilities, and processors located far from training or market centers. Participants repeatedly linked these profiles to specific constraints within the proposed scaling strategy.

For instance, the team anticipated that younger women with lower levels of education would face compounded disadvantages because of their inability to access information around HQCP (given the planned communication strategies), weaker financial standing, and restricted participation in training and partnership opportunities. Similarly, people with low levels of education and no prior business experience were seen as less able to navigate the technical and financial requirements that are part of the scaling strategy for accessing loans to support the establishment of HQCP businesses.

One participant explained during the workshop discussions:

Young women with lower levels of education may not even hear about some of these opportunities, and even when they do, they may not understand the loan processes or business requirements ….and in some areas, they may need their husband's approval to even participate in this project in the first place. Gender expert and HQCP team member.

The team also noted that women and men with limited financial resources or collateral cannot access the financial assistance (loans) planned as part of their scaling strategies and that young people living with disabilities may encounter additional barriers related to both physical accessibility (for example operating HQCP machines) and/or attending training organized as part of the scaling strategy. Residents located far from training locations, processing facilities, and markets were seen as being at risk of marginalization because of higher transportation costs, which might limit their ability to attend training (for those interested in training) and to access HQCP processing centers for others in such areas, hence limiting their access to the market. The team further highlighted the situation of married women who, despite formal access to resources, face decision-making constraints within household and community structures; the team anticipated that they might need permission from their husbands to participate in training or even establish an HQCP processing factory. Individuals without association membership or group affiliation—including women with disabilities living in rural communities—were understood to face the most significant compounding of barriers, as they lack both social capital and institutional recognition within the scaling framework.

Analysis of the pre-existing scaling strategy documents revealed that many of these differentiated vulnerability profiles were absent prior to the GenderUp workshop. Earlier documents generally framed target groups in binary categories such as women, men, or youth, without examining differences within those groups or how vulnerabilities might compound through intersecting social conditions.

The national program coordinator of the cassava producer association and an HQCP team member reflected during a follow-up interview:

Before this workshop, we mostly talked about women and youth generally. But now we are seeing that vulnerability is much more specific. Not every woman faces the same challenge and not every young person has the same opportunities.

Overall, stage 3 deepened the team's understanding of how vulnerabilities emerge and compound through multiple social intersections. This shift from a general gender focus toward a nuanced intersectional awareness established a critical foundation for stage 4, where the team began adapting their scaling strategies to mitigate risks and promote inclusive participation.

3.2.4. Using HQCP to test stage 4's mechanism of change: Deeper appreciation of scaling process complexity

At stage 4, participants engaged with a series of guided reflection questions on the GenderUp online whiteboard that prompted them to reassess bottlenecks, training plans, communication strategies, complementary innovations, and partnership needs within their current existing scaling strategy. The discussions focused on how previously identified vulnerabilities could translate into unequal outcomes if scaling strategies were not deliberately adapted.

Workshop observations and analysis of the whiteboard showed that participants increasingly linked social inclusion concerns to operational and institutional dimensions of scaling. Rather than discussing vulnerability only in abstract terms, participants began proposing practical changes to communication systems, training models, financing arrangements, and partnership structures. The team identified several existing critical bottlenecks that needed to be addressed to ensure equitable scaling outcomes. These bottlenecks included, among others, limited access to business planning and development training within specific sections of society (married women, people living far from areas of training, less educated people, etc.), weak market access by vulnerable groups, and inadequate technical capacity among some target users (less educated, young women). The team also noted that issues of accessibility—such as insufficient road networks and limited access to information—might further constrain participation, particularly for women, youth, and persons with disabilities. Our analysis revealed that almost all identified bottlenecks were based on, and in line with, previous stages, especially the stage 3 intersectional at-risk profiles. In response, the team proposed several adjustments to their existing training, capacity-building, and communication strategies and existing partnerships. They also proposed mitigation strategies to address some identified risks.

Analysis of workshop transcripts indicated that many of the proposed adaptations built directly on the intersectional at-risk profiles identified during stage 3. For example, in relation to training and capacity building, the team recommended the inclusion of business plan development modules for young men and women with low levels of education and no prior experience in business to enhance their ability to access loans, which are part of scaling strategies. Other proposed capacity building activities included training more fabricators to improve access to machinery, machine maintenance, and repairs. According to workshop observations, the team further proposed the establishment of incubation centers as a means to promote energy-efficient and environmentally friendly HQCP processing models while fostering innovation and collaboration among small-scale processors. These concerns were later echoed during field interviews with processors. One HQCP processor explained:

Up until now, we have had a septic hole here (pointing to a well-covered hole) that we direct the wastewater to, but we are working on a system to convert this waste to organic gas for cooking and for the community to fry gari with. We have some white people come over to access our proposal and they are very interested in this.

Revisions to communication strategies were another key outcome of this stage. The team recognized the need to make communication more inclusive and disability-friendly, while emphasizing the confidence and visibility of women participants. They also proposed the use of localized platforms—such as WhatsApp groups and community radio—to improve information flow, ensure consistent engagement, and strengthen collective learning among target users. These were proposed to help users (especially less educated groups) who might not be able to access communication strategies like YouTube videos, flyers, journal publications, and so on incorporated in existing scaling strategies.

To complement these measures, the team identified several mitigating strategies and complementary innovations, including the establishment of HQCP processor associations or platforms to enhance peer learning, facilitate information exchange, and strengthen negotiation power with buyers and suppliers. They also recommended exploring potential cooperative systems that would enable processors—especially those with disabilities or limited resources—to pool assets, share labor, and jointly access inputs and markets. Additionally, the facilitation of contractual agreements between HQCP processors, cassava peel suppliers, and aggregators was suggested to stabilize supply chains and enhance accountability. All of these were observed as being developed to address some unintended consequences anticipated in earlier stages.

Finally, the team outlined new partnerships required to operationalize the revised scaling strategy. These included financial institutions with youth- or gender-oriented financing schemes (e.g., Sterling Bank, Bank of Agriculture, Mastercard Foundation). The team further emphasized the need to engage business development agencies and extension systems to sustain capacity-building and long-term technical support.

At the end of stage 4, the HQCP scaling team demonstrated an improved understanding of how to integrate social inclusion into practical scaling strategy design. Through moving from abstract recognition of diversity-related challenges to concrete, actionable strategies aimed at addressing barriers at multiple levels—from access to training and capacity building, to communication and financing strategies, to organizational arrangements and institutional partnerships. Thus, the stage 4 outputs marked a significant step toward embedding equity, inclusion, and opportunity-driven approaches in the HQCP scaling strategy.

3.2.5. Using HQCP to test stage 5's mechanism of change: Greater appreciation of practical implication of equity in project

At the time of this assessment, the HQCP project implementation was only at its inception phase, thereby limiting opportunities to observe the full translation of GenderUp recommendations into practice. Nevertheless, we examined whether and how the project's planning and management documents reflected adjustments made following the GenderUp workshop. This involved a comparative analysis of the project's scaling strategy before and after the workshop to identify revisions informed by GenderUp recommendations. We also investigated the implications of existing partnerships for the use and implementation of GenderUp recommendations.

In addition, the study assessed the extent to which these documented changes—and the broader insights from the GenderUp process—were reflected in emerging implementation practices (thus current decisions being taken in the implementation process so far) and aligned with field realities in Oyo State, Nigeria. By integrating perspectives from stakeholders across the HQCP value chain, this stage provides an empirical assessment of how GenderUp-inspired adjustments resonate with the socioeconomic and institutional realities of scaling.

3.3. Field validation and further assessment of stage 5 implementation

The findings from the assessment of stage 5 implementation are presented under four thematic categories: boxed-in recommendations in strategy documents, institutional and partnership dynamics as enablers and constraints of implementation, the influence of stakeholder representation on workshop outputs, and the identification of barriers to scaling that remained outside the GenderUp process.

3.3.1. Boxed-in document but broader strategy remains largely unchanged

The comparative document analysis revealed that, although some revisions were made to the HQCP strategy document, these were largely confined to a new chapter summarizing the workshop's outcomes, leaving the broader scaling strategy largely unchanged. Thus, the results from the GenderUp workshop were boxed in the document as a chapter, without effecting the changes in the scaling strategy's operational framework.

Interestingly, our observations from the project's implementation inception meeting revealed that several GenderUp recommendations appeared in the HQCP team's presentation of their scaling strategy. These included commitments to improve inclusion—such as making provisions for persons with disabilities and engaging specific partner organizations in implementation. This raised important interpretive questions about intent and institutionalization: whether these inclusions reflected genuine efforts to embed GenderUp principles into practice, or whether they were selectively highlighted for presentation purposes because of the presence of consortium partners and donors. We also observed at the inception meeting that the HQCP team made some attempts to inculcate some of their recommendations in the consortium but were mostly brushed aside by lead consortium members.

3.3.2. Institutional and partnership dynamics as enablers/constraints of GenderUp implementation

Our analysis of interviews and inception meeting observations revealed that existing institutional partnerships played a dual role—functioning as both enablers of, and barriers to, the effective use and implementation of the adapted strategies resulting from the workshop.

A central finding was that donor requirements strongly influenced the HQCP scaling team's decision to participate in the GenderUp workshop. For many team members, integrating gender considerations was perceived not only as good practice, but also as a mandatory component of donor compliance. As one scaling team member explained: "*If I show you the document from the donor and the things they required us to do, you will know that the donor is big on the issue of gender, especially women and youth*".

The donor emphasis on gender mainstreaming motivated positively the HQCP team's engagement with GenderUp. However, the same donor-driven conditions also presented challenges for implementing the recommendations that emerged from the workshop. The consortium's broader project structure was designed with a dedicated gender component, which included explicit targets and deliverables related to

female and youth participation. Although this focus enhanced visibility and institutional commitment to gender issues, it also imposed quantitative performance pressures that sometimes conflicted with GenderUp's qualitative and reflexive ethos.

For instance, donor expectations regarding the number of beneficiaries to be reached (millions of women and youth) and the measurable outputs to be achieved led the consortium to make efficiency-driven decisions that inadvertently risked excluding vulnerable groups. For example, in efforts to reach a larger number of participants with limited resources, the project established eligibility criteria that favored farmers who already possessed key productive assets—criteria that potentially marginalized poorer or more resource-constrained groups. This reflects a broader tension between donor accountability frameworks, which emphasize scale and numbers, and the GenderUp approach, which prioritizes inclusivity and context sensitivity.

In addition to donor influence, we observed that the consortium's organizational composition and partnership structure and governance had the potential to shape further implementation of GenderUp outcomes. To begin with, the gender mainstreaming function—referred to as the Gender Equality and Social Inclusion (GESI) unit—was delegated to a specific consortium partner. Although this arrangement was intended to streamline gender integration, it inadvertently created compartmentalization, weakening both ownership of, and capacity to implement, the GenderUp outputs within the HQCP team. Notably, the partner responsible for GESI was not involved in the GenderUp workshop and only became aware of the process during the inception meeting, when the researchers (us) introduced it. GenderUp was conducted by the HQCP team. This lack of prior engagement meant that the partner had no sense of ownership or obligation toward the GenderUp recommendations.

Furthermore, the GESI partner had independently conducted its own gender assessment and had developed gender criteria and implementation guidelines for all consortium members. As these guidelines were designed primarily to meet donor quantitative targets, the GESI framework being advanced at consortium level did not integrate the insights, intersectional considerations, or context-specific mitigation strategies emerging from the GenderUp process.

This separation of functions also reflected underlying power dynamics within the consortium. Because the GESI partner controlled the funding allocation for gender-related activities, it effectively determined the gender actions that were prioritized and resourced. Consequently, the HQCP team's efforts to advocate for the adoption of GenderUp recommendations encountered resistance. On several occasions, the team's suggestions were subtly dismissed by the GESI lead, often justified on the basis that donor quantitative requirements made such adjustments 'unfeasible' within the project's design.

Furthermore, it became apparent that the consortium as a whole had not fully articulated several key elements of its scaling strategy at the inception meeting. For example, there was no clear plan for internal or external communication, nor agreement on what information should be included in outreach materials such as flyers or shared with project participants. Although general communication channels were mentioned, detailed responsibilities, messages, and feedback mechanisms remained undefined at the inception of the implementation process.

These dynamics illustrate that GenderUp outcomes were introduced into a complex institutional environment characterized by hierarchical relationships, competing mandates, and donor-driven conditionalities. Implementing the recommendations therefore required negotiation and alignment across multiple actors with differing priorities and authority. The findings reveal that, although the HQCP team demonstrated intent to integrate GenderUp principles, the absence of consortium-wide commitment and the compartmentalization of gender functions limited the adjustment of scaling strategies.

The findings underscore that institutionalizing GenderUp's inclusive principles requires not only project-level commitment, but also structural alignment across all partner organizations. Without collective ownership, the risk remains that equity-oriented practices may be compartmentalized within individual subprojects rather than embedded across the consortium's overall scaling strategy.

3.3.3. When representation shapes reality: Workshop reflections diverging from local gender and resource dynamics

Field interviews with farmers, HQCP processors, and feed millers revealed that, although the resources identified during the GenderUp workshop as essential for accessing the HQCP innovation—such as land, machinery, and technical skills—were broadly relevant in other part of Nigeria, their significance varied substantially within local contexts. In Oyo State, for instance, access to land, although generally recognized as important for HQCP production, was not perceived as a major constraint among the primary target groups of poultry farmers and entrepreneurs. Respondents explained that most poultry farmers already possessed land for their operations and that establishing HQCP processing units required only limited space. Consequently, they did not consider land access as a barrier to adoption.

Results from the GenderUp workshop emphasized that "*Gender plays a significant role in land ownership rights, with women often facing legal and cultural barriers that limit their ability to inherit or control land resource*" (Scaling document after-GenderUp, p. 24).

However, respondents in Oyo State largely disagreed with this generalization. They maintained that, within their local setting, access to land is determined primarily by financial capacity rather than gender.

In this area, all you need is your money and you will get land. As long as you have money, it does not matter if you are a man, woman, or young person—you will get land. We don't have any law or rule stopping anyone from getting land; maybe some other states in the north might have those cultures that stop women, but not here. HQCP processor.

This divergence illustrates how workshop participants' assumptions about gendered access to resources may not fully capture localized realities. Structural inequalities in land ownership persist elsewhere, defined by gendered tradition and culture, but such barriers appeared less salient in Oyo State.

A similar pattern emerged regarding access to machinery and technical skills. During the GenderUp workshop, participants identified the high cost of equipment and the presumed inability of women to operate machinery as significant barriers to scaling. However, respondents in Oyo State largely contested these claims. They explained that technological adaptation and local fabrication had substantially reduced equipment costs, enabling small-scale farmers and entrepreneurs to purchase customized or smaller models suited to their needs.

Working with donor organizations over the past 10 years since HQCP was introduced, we have developed smaller, more affordable versions of the machines, often using local materials. Farmers can now buy smaller or custom-built machines suited to their budgets, so cost is not an issue—maybe in places where they have not modified the machines. Fabricator.

Furthermore, respondents rejected the notion that women faced significant technical barriers to machine operation. Many emphasized that women are active participants in processing centers and increasingly manage equipment themselves. Recent technological improvements, such as the introduction of machines that start automatically using electricity, have further reduced the physical effort required for operation, effectively removing a barrier that once had gendered implications.

Taken together, these findings highlight that, although GenderUp effectively sensitized scaling teams to consider intersectional inequalities, the local realities in Oyo State reveal a more nuanced picture. However, arguably these results applied to GenderUp workshop participants because there was no local representation from Oyo State.

Moreover, our results point to a broader issue: even when contextual variation is recognized, some critical, context-specific barriers may still escape conversational frameworks such as GenderUp. In the HQCP case, one such overlooked constraint emerged not from gender or resource

access per se, but from the quality and trust dimensions embedded in local production systems. Section 3.3.4 explores how the challenge of *adulteration*—a factor largely absent from the GenderUp analysis—has become a defining barrier to HQCP scaling in Oyo State.

3.3.4. Capturing most, but missing key barriers to scaling: The case of adulteration

The GenderUp workshop identified several potential barriers to the effective scaling of the HQCP innovation in Nigeria, including: (i) the distance between HQCP processing factories and cassava peel sources—raising production costs; (ii) the distance between farmers and processing centers—limiting access and increasing transport costs; (iii) the limited number of processing centers and associated capacity constraints; and (iv) weak market linkages among key actors across the value chain. For each of these, the workshop generated practical solutions. For example, to address distance-related constraints, the team proposed two complementary measures: establishing mini processing units close to cassava peel sources (typically cassava processing clusters) to produce HQCP cake as an intermediate product; creating aggregation services for the efficient collection and transport of cassava peels. Similar solutions were outlined for the other barriers.

However, despite the relevance of these proposed measures, our interviews revealed that the workshop failed to recognize a critical barrier that fundamentally affects potential HQCP scaling in Oyo State: adulteration. Interviews with various stakeholders in Oyo State indicated that adulteration of HQCP feed material (intentional production of low quality HQCP with unwholesome cassava peel usually less than 24 h old) has emerged as the single most important factor undermining adoption, overshadowing other technical or logistical constraints. If this issue is not addressed, even full implementation of the GenderUp recommendations would not guarantee successful scaling.

According to several respondents, feed adulteration severely undermines farmers' confidence in HQCP. The Director of Feed Regulation in Nigeria explained:

I have gone round the field, and I can tell you the major reason why poultry farmers are not adopting HQCP is because many of the processors tamper with the feed. They sometimes mix old or rotten peels, making the feed harmful for livestock. How do you expect farmers to buy that? Even large feed operators have complained to me and said that was the reason they stopped using it.

Adulteration poses a direct risk to poultry production. Farmers reported that using substandard feed often resulted in reduced productivity—particularly among layers—and, in severe cases, high mortality rates. Consequently, most poultry entrepreneurs were unwilling to take the risk. Respondents emphasized that trust had become the defining factor in their purchasing decisions.

I stopped using HQCP after ILRI couldn't meet our demand. The issue wasn't availability—others were selling—but I only trusted ILRI's product. I never had problems using their feed, but when they couldn't supply, I got frustrated and stopped. Female poultry farmer, FGD.

HQCP processors confirmed that adulteration and trust deficits were major bottlenecks:

Come and check our warehouse. We have a lot of HQCP here, but the poultry farmers won't buy it. They say they only want the "original HQCP". Only the pig farmers are buying regularly. HQCP processor.

An agricultural researcher further explained that the impact of adulteration varied across livestock systems:

This happens because poultry is a high-risk business. Once one farmer experiences problems with a feed, others avoid it entirely. The processors are blamed for adulteration. But pigs have four stomach chambers, so they can tolerate it better—that's why pig farmers are using HQCP more than poultry farmers.

These findings suggest that, although GenderUp effectively addressed several structural and logistical barriers, it overlooked the trust and quality assurance dimensions critical for sustained adoption. Adulteration not only erodes confidence among users, but also reinforces perceptions that local processors cannot meet quality standards—a perception that institutional interventions must address. By failing to capture this key issue, the workshop's diagnostic process risked HQCP scalability in Oyo State. Unless quality assurance mechanisms and trust-building measures are explicitly integrated into the scaling strategy, the risk of adoption failure remains high, regardless of how well other barriers are addressed.

The prominence of adulteration as a scaling barrier underscores that technical solutions alone are insufficient for responsible scaling. Issues of product integrity, trust, and user confidence are fundamentally institutional in nature, shaped by governance arrangements, partnerships, and accountability systems. Addressing such challenges therefore depends on how project teams, donors, and implementing partners coordinate their roles and incentives to support transparent and equitable practices.

4. Discussion

This section interprets the findings in relation to the three research questions and wider debates on responsible scaling and gender-transformative approaches (GTA). Section 4.1 examines the mechanisms through which GenderUp is expected to foster reflexive and socially inclusive scaling decisions (RQ1), and identifies an important limitation concerning historical reflexivity. Sections 4.1–4.3 then consider how, why and under what contextual and institutional conditions these mechanisms were realised, enabled or constrained in the HQCP case (RQ2): respectively, the history of the innovation, the composition of the group involved in deliberation, and the partnership and donor arrangements within which recommendations had to be taken forward. Section 4.4 examines what the case reveals about GenderUp's contribution to operationalising gender-transformative thinking and draws together its implications for the uptake of recommendations and the design, facilitation and governance of dialogue-based scaling decision-support tools (SDSTs) (RQ3).

4.1. Mechanism of change working, but missing historical reflexivity

The findings reveal that GenderUp's underlying mechanisms of change demonstrate internal coherence and functional effectiveness. Each stage is linked to a plausible mechanism of enhanced understanding and deeper appreciation. Stage 1 fosters a better understanding of the requirements for achieving scaling objectives; stage 2 enhances better understanding of relevant dimensions of diversity; stage 3 deepens appreciation of intersectionality and vulnerability creation; stage 4 deepens appreciation of scaling process complexity; and stage 5 cultivates greater appreciation of practical implications of equity in scaling. Together, these stages operationalize core principles of responsible innovation and scaling, including anticipation, inclusivity, reflexive learning and responsiveness (Wigboldus and Brouwers, 2016; Wigboldus et al., 2020; Ewell et al., 2026). manifested through the adaptation of scaling strategies. The HQCP case provides evidence that a facilitated, staged dialogue can activate these mechanisms in practice. Following the workshop, the HQCP team moved beyond the binary categories of women and youth toward more differentiated at-risk profiles and revised elements of its strategy relating to communication, training, financing and partnerships. This resonates with the "responsibility check" on gender, age, inclusiveness and power equity described by Woltering et al. (2019), and with the role of participatory dialogue and critical reflection in GTA (Hillenbrand et al., 2024). However, these changes in understanding and planning do not, in themselves, demonstrate that all recommendations were implemented or that the resulting changes were transformative (see Section 4.4). In realist terms, when a facilitated, staged dialogue is conducted with a team that has an existing scaling strategy (context), structured prompts can trigger reflection and a deeper appreciation of diversity and risk (mechanism), contributing to revisions to elements of that strategy

(outcome). The following sections examine the conditions and boundaries of this configuration.

The findings also indicate that GenderUp's processes lack historical reflexivity; understood here as scaling actors' capacity to systematically engage with an innovation's past trajectory, including earlier successes, failures, and unintended consequences, and to integrate such learning into adaptive and context-responsive scaling decisions. The workshop enabled participants to identify relevant barriers to scaling, such as the distance to processing centres and weak market linkages, but it did not capture a constraint that had hindered HQCP adoption in Oyo State for nearly a decade: feed adulteration, a problem of quality and trust rather than of gender or resource access. The responsible scaling literature underscores that continuous reflexive learning and monitoring are essential in complex adaptive systems such as AR4D interventions (Sartas et al., 2020; Schut et al., 2020; Wigboldus et al., 2020). SDSTs, however, have been critiqued for their predominant reliance on forward-looking causal reasoning embedded in their mechanisms, often neglecting backward learning derived from previous innovation trajectories and scaling experiences (Damtew et al., 2023). Our case confirms this critique and refines it. Scaling scholarship has argued that scaling efforts neglect the relationships, history and incentives, that shape whether an innovation can take hold (Woltering et al., 2019), and that every context has its own history and particularities (Wigboldus and Brouwers, 2016). The HQCP case shows that a tool can succeed at the first reflexive task, anticipating how a strategy may affect different groups, and still miss the second, learning from an innovation's past, because the two draw on different knowledge.

This raises a broader question about the mechanisms of dialogue-based SDSTs: do they systematically prompt users to learn from an innovation's previous trajectory, particularly when the innovation has already been developed, piloted or used for some time? The question becomes even more salient when looked at it from the broader multi-level perspective on system change; constraints such as feed adulteration maybe a reflection of established practices, relationships and institutional arrangements within a (value chain) system in which the innovation is embeded, and such practices and rules can stabilize arrangements and constrain change (Leeuwis et al., 2021). Moreover innovations may differ in how far they follow dominant arrangements or depart from them (Gaitán-Cremaschi et al., 2019). This suggests anticipating consequences for different social groups needs to be paired with an analysis of the configuration of the system into which an innovation is scaled, including its history of trust, quality assurance and governance. Embedding mechanisms that explicitly incorporate backward learning, through structured reflection on past innovation histories and scaling trajectories could enhance the tool's ability to support truly adaptive and context-responsive responsible scaling. Also a combination with tools such as Scaling Readiness (Sartas., 2020) may also help users assess the innovation's maturity and the wider conditions that enable or constrain its scaling.

4.2. Representation and contextualization: Whose presence and whose knowledge shapes the development of responsible scaling strategies?

The HQCP case connects to the wider emphasis on inclusion in responsible scaling and innovation: who gets to speak, whose knowledge counts and who decides. Recent work on responsible scaling adds legitimacy and epistemic justice to the principles discussed above (Ewell et al., 2026; Altshul, 2025). This shifts attention from what scaling decisions should consider to how they are made and whose knowledge informs them. Two questions are central: who makes decisions, and how, which concerns governance and procedural fairness; and whose evidence counts, which concerns the recognition and integration of different forms of knowledge. The same questions run through the systems perspective, in which governing system change involves deciding who makes decisions, whose knowledge counts and how power shapes the choices available (Leeuwis et al., 2021; Leeuwis et al., 2026). They bear directly on the composition of a group taking part in a dialogue and on how far its perspectives reflect the experiences of those affected by scaling.

Contextual validity and embeddedness in innovation development and scaling strategies are core tenets of responsible scaling scholarship, which emphasizes the need for decisions and strategies to reflect socially differentiated realities (McGuire, Al-Zu'bi, et al., 2024; McGuire et al., 2022; McGuire, Leeuwis, et al., 2024). A related principle is the need to base decisions on contextually valid knowledge and information; an important principle of responsible scaling requires not only that scaling strategies reflect local realities, but also that diversity exists within the decision-making processes that shape these strategies (Steinke et al., 2022). Diversity is not merely a procedural issue, but also a substantive determinant of the kinds of knowledge and assumptions that inform strategy design. However, the findings from this study reveal that such socially differentiated realities cannot be meaningfully captured without the active representation of those most directly affected by the innovation and its associated scaling processes.

The HQCP case demonstrates that, although GenderUp's underlying mechanisms of change are effective in enhancing critical reflection on dimensions of social diversity, resource access and intersectionality, the degree to which these reflections align with the lived realities of intended beneficiaries depends largely on whose perspectives inform them. In other words, who is in the room during a GenderUp workshop matters profoundly for effective contextualization.

In the HQCP workshop, participants were primarily national-level experts and institutional representatives with experience in system-level scaling. Consequently, their discussions and resulting strategies reflected a broad, nation-level understanding of social diversity and resource access; effectively a national average. In contrast, local respondents in Oyo State acknowledged that, although such challenges might exist elsewhere in Nigeria, they did not align with Oyo's specific local conditions. For instance, gendered constraints to land ownership, identified by the workshop participants as a barrier, were widely dismissed by local actors who emphasized that land access in their communities depended primarily on financial capacity rather than gender norms.

This divergence underscores a critical insight: although GenderUp's mechanisms of change posit that mechanisms of critical reflection are activated when diverse participants engage with structured prompts, leading to inclusive scaling strategies, the full realization of this mechanism is contingent on participant composition. When the room lacks locally embedded voices such as smallholder poultry farmers, women cassava processors, and community-based fabricators that are crucial to HQCP scaling: the reflective process tends to operate on abstract understandings of diversity derived from those in the room, rather than on the lived experiences of those who will ultimately be affected by the innovation. This is consistent with what Wigboldus and Brouwers (2016) describe as ecological and composition fallacies among common scaling misconceptions: what holds at one level or for one group cannot be assumed to hold at another level or for everyone. McGuire, Leeuwis, et al. (2024) and Rietveld et al. (2022) similarly note that, in dialogue-based settings meant to surface intersectional social diversity, the absence of potential end users can shape what the dialogue produces. The implication is not that national expertise should be excluded, but that it should be combined with locally grounded knowledge and treated as an assumption to be tested rather than a definitive description of local reality. Research on gender and food systems likewise shows that gender relations and access to productive resources vary across and within contexts, including through their interaction with other dimensions of social difference (McGuire et al., 2022; McGuire, Leeuwis, et al., 2024; Njuki et al., 2023). Studies in Ghana point the same way for access to productive resources and to information (Ankrah et al., 2020; Ngissah, 2024).

In realist terms, when a dialogue is convened mainly with national-level experts (context), structured prompts activate reflection on

diversity and access (mechanism), but this draws on a national average and yields strategy elements that did not fully match Oyo State realities (outcome). Local representation, or a step that tests workshop insights with local actors, is the condition under which the same mechanism could work on local knowledge.. The HQCP case thus shows that contextual validity in responsible scaling is a matter of epistemic diversity as well as methodological design. The presence and knowledge of participants shape deliberation and determine whether reflection processes, such as those facilitated through GenderUp, lead to context-sensitive and socially inclusive scaling strategies.

Inclusion is also linked to debates about categories and intersectionality. The workshop's movement beyond the binary categories of women and youth is a meaningful methodological gain. Intersectional approaches draw attention to how gender intersects with other dimensions of social position to shape people's opportunities and constraints, yet translating this into programme design and implementation remains challenging (McDougall et al., 2021; Rietveld et al., 2022; Hillenbrand et al., 2024). GenderUp gave the team a structured opportunity to discuss differentiated risks and potential disadvantages. However, the consortium continued to use broad categories such as women and youth in some planning and reporting processes, which team members associated with donor requirements. Developing intersectional insights in a workshop therefore does not necessarily change the categories used in institutional planning, funding and accountability. Section 4.3 considers the significance of this gap for uptake.

Building on these insights, future iterations of GenderUp and comparable dialogue-based SDSTs should include clear guidance on participant selection, to balance representation across institutional levels and social groups. Because not every relevant actor can join a single workshop, an added stage of cross-checking and information gathering, in which workshop insights are validated and refined through localized data collection, would strengthen the link between reflective dialogue and on-the-ground realities. Such a step would also connect facilitated reflection more closely to the locally grounded and reflexive practice advocated in gender-transformative research (Rietveld et al., 2022).

4.3. How existing partnerships, institutional arrangements, and donor requirements shape the use and implementation of SDST recommendations

The scholarship on responsible scaling highlights the critical role of institutional governance, incentive structures, and partnership configurations in shaping how scaling interventions are designed and implemented (McGuire et al., 2022; McGuire, Leeuwis, et al., 2024; Nchanji et al., 2023; Sánchez Rodríguez et al., 2021; Wigboldus et al., 2020). The HQCP case demonstrates how these institutional and donor dynamics can both enable and constrain the effective use of SDSTs such as GenderUp.

Donor mandates first enabled attention to gender. The scaling team's decision to take part in the GenderUp workshop was strongly influenced by donor requirements on women and youth, and the team treated gender integration as part of compliance as well as good practice Hillenbrand et al. (2024) similarly report that donor support can create organizational mandates and momentum for gender-transformative work. The same donor conditions then constrained implementation in two ways.

The first constraint is the donor's performance metrics. Requirements to reach millions of women and youth to be reached may have created pressure to maximize scale rather than inclusivity. This quantitative framing encouraged efficiency-driven approaches that risked sidelining vulnerable groups—contradicting GenderUp's focus on equity and context-specific adaptation. Consequently, although GenderUp promoted project-level critical reflection and adaptation, its broader influence was potentially constrained by donor accountability requirements that prioritized numerical targets.

These findings reflect well-documented tensions between accountability for numbers and the aims of responsible and gender-transformative scaling. Woltering et al. (2019) argue that donor emphasis on the number of end users reached can lead scaling teams to prioritize visible, measurable results over the systemic changes needed for lasting impact. The HQCP case highlights the same tension between demonstrating numerical reach and responding to the differentiated needs that GenderUp brought into view. It also raises a broader question about what claims of women's economic participation in development projects signify. Gender scholarship cautions that women's economic participation can be promoted instrumentally, without challenging the norms and power relations that shape women's opportunities or ensuring a more equitable distribution of benefits (McDougall et al., 2021; Hillenbrand et al., 2024). In the HQCP initiative, some eligibility criteria favored farmers who already possessed key productive assets, raising concern that the initiative could reach women and young people without reaching those facing the greatest constraints. The available evidence does not establish that quantitative reporting requirements led to exclusion at the time of the research. The possibility that numerical targets and eligibility criteria could work against substantive inclusion nonetheless remains an important risk.

The second constraint is how the consortium organized gender expertise. Having both a designated gender specialist and an attempt to mainstream gender as an integral part of all work can create tensions in partnerships. Hillenbrand et al. (2024) report that gender is often treated as the responsibility of “the gender person”, that gender expertise may be under-resourced and devalued, and that donor support can give gender advisers more influence over budgeting, although some respondents also described a shift toward shared responsibility. McDougall et al. (2021) similarly argue that organizations must embody gender equality in their own practices if they are to act as agents of change, while noting persistent resistance to institutionalizing gender within development organizations.

In the HQCP consortium, the GESI function was assigned to one partner that controlled gender-related funding. GenderUp was organized for the HQCP scaling team, and the GESI partner did not take part in the workshop as they were not invited. Its guidelines were shaped by donor targets, and suggestions developed by the HQCP team were dismissed as unfeasible. This case reinforces existing accounts and refines them in two ways. First, it locates the challenge not simply in the presence or absence of gender expertise, but in how authority and resources are distributed between those who develop socially inclusive strategies and those who control gender-related funding. Second, it shows how this division can constrain the implementation of SDST recommendations even when the participating team is committed to inclusive scaling. Wigboldus et al. (2023) identify a related dynamic in scaling partnerships, where the political economy of incentives, interests and power relations unrelated to the innovation constrained its institutionalization.

In realist terms, when gender mandates and funding sit with a donor-driven partner outside the team that took part in the dialogue (context), the recommendations that emerge from reflection lack collective ownership and compete with numerical targets (mechanism), and the strategy changes remain confined to a chapter of the strategy document rather than reaching its operational framework (outcome). Institutional arrangements and donor frameworks are therefore part of the conditions under which any SDST operates, and they need to be addressed explicitly if socially inclusive principles are to go beyond workshop-level commitment (see Section 5).

4.4. Strengthening gender-transformative thinking in scaling: What the HQCP case shows and adds

We positioned GenderUp as a way to operationalize gender-transformative thinking in specific scaling settings by drawing attention to differentiated innovation packages and the potential for addressing gender norms and structures as part of these. In the HQCP case, GenderUp helped the team move beyond broad categories of

women and youth toward more differentiated at-risk profiles. The revised strategy included measures to improve access to skills, finance, markets and information, such as business-planning modules, training for fabricators, localized and disability-friendly communication, producer associations and partnerships with financial institutions.

Viewed through the continuum from accommodative to transformative approaches (McDougall et al., 2021; Hillenbrand et al., 2024), these measures were mainly accommodative. Although the workshop identified concerns such as women's land rights and ability to operate machinery, the revised package did not include measures directly addressing discriminatory norms or the power relations underpinning them. Nor did it address the structural constraints discussed in Section 4.3, including donor targets and control over gender-related funding. As implementation was still at an early stage, subsequent changes cannot be ruled out. However, the available evidence does not demonstrate that the package had begun to transform these underlying norms and structures. GenderUp therefore supported more differentiated and inclusive strategy discussions, but translating this reflection into transformative action requires further design and institutional support.

The gender transformative potential of GenderUp could be enhanced in several ways; First, Stage 4 could more explicitly prompt teams to consider interventions addressing relevant norms and structural constraints, and who would implement them. Second, assumptions about discriminatory norms should be tested with local actors before mitigation measures are designed (Section 4.2). Third, donors and partners controlling gender-related funding should be involved so that proposed actions can be considered alongside the authority and resources needed to implement them (Section 4.3). These adjustments could help connect GenderUp's diagnostic contribution to calls for gender-transformative research to address the normative and structural constraints that sustain inequality (Rietveld et al., 2022).

5. Conclusion and recommendations

We studied how GenderUp, a conversation-based SDST, contributed to the development and deployment of responsible and socially inclusive scaling strategies within the AR4D context. Although our findings may be context-specific to the case study, they provide insightful lessons and opportunities to discuss broader themes in responsible scaling scholarship. We highlight the role of SDSTs in developing inclusive scaling strategies within the AR4D context. We also demonstrate that the development and implementation of socially inclusive scaling strategies are not determined by the design of an SDSTs alone, but also shaped by representation among participants engaged with the tool, historical reflexivity, donor requirements and institutional partnerships, and their governance structures.

Our findings indicate that GenderUp embodies coherent and robust mechanisms of change that effectively foster critical reflection and adaptive strategy design. The tool helps innovation teams identify relevant dimensions of diversity, analyze intersectionality, and revise scaling strategies to minimize harm and enhance inclusion. However, the realization of these outcomes is highly contingent on contextual and institutional dynamics.

Three factors emerged as especially influential. First, representation—who participates in the workshop—determines the contextual validity of outcomes, as the perspectives present in the room shape how diversity and inclusion are understood. Second, historical reflexivity is limited: although GenderUp stimulates forward-looking reflection, it can overlook persistent structural barriers, such as market adulteration in the case of HQCP, that historically constrain scaling. Third, institutional arrangements—including donor requirements and partnership structures—both enable and constrain GenderUp's influence, highlighting tensions between tool intent and systemic realities.

At the start of this paper we positioned responsible scaling and GenderUp as attempts to operationalize systemic and gender transformative thinking in AR4D scaling initiatives, and overcome ingrained tendencies of focusing on scaling single solutions and overlooking inequitable consequences of scaling for different intersectional groups or geographical communities. Overall, we can conclude from our realist evaluation that SDSTs such as GenderUp can indeed activate mechanisms that support AR4D innovation teams to think and talk more systemically and responsibly about their scaling initiatives and translate this toward the development of coherent action plans. However, such action plans may be based partially on invalid information and interpretations. Moreover, the implementation of the scaling strategies included in these plans can easily be constrained by the kinds of institutional arrangements that have tended to affect the enactment of systemic innovation approaches more broadly. Such constraints are insufficiently anticipated in the program theory assumed in the GenderUp design.

These findings suggest several refinements that could make GenderUp more effective in practice. More broadly, they point to design and governance questions for any dialogue-based SDST that seeks to support responsible scaling. Future iterations of GenderUp, and comparable dialogue-based SDSTs, should include explicit guidance on participant selection to ensure balanced representation across institutional levels and social groups. Given that not all relevant actors can be included in a single workshop, adding a cross-checking and information gathering stage to validate and contextualize workshop insights through localized engagement could enhance practical alignment. Embedding mechanisms for backward learning, drawing on past innovation trajectories, would further strengthen adaptive decision-making. Finally, effective implementation requires early engagement with consortium leadership and donors to foster shared ownership, resource alignment, and institutional commitment to responsible scaling. Future research should test these propositions across other cases and SDSTs, follow implementation beyond the inception phase to see whether tool-supported strategies change gender norms and structures over time, and examine how local actors can be involved in dialogue-based tools without losing the system-level view that national and institutional participants bring.

CRediT authorship contribution statement

Ebenezer Ngissah: Writing – review & editing, Writing – original draft, Visualization, Methodology, Investigation, Formal analysis, Data curation, Conceptualization. **Esther Kihoro:** Writing – review & editing, Writing – original draft, Supervision, Formal analysis, Data curation, Conceptualization. **Cees Leeuwis:** Writing – review & editing, Writing – original draft, Supervision, Formal analysis, Conceptualization. **Katharine Legun:** Writing – review & editing, Supervision, Conceptualization. **Mareike Smolka:** Writing – review & editing. **Comfort Freeman:** Writing – review & editing, Conceptualization. **Erin McGuire:** Writing – review & editing. **Adebola Bunmi Adebayo:** Data curation. **Hanna Ewell:** Writing – review & editing. **Jasada Ijudai:** Conceptualization.

Ethical approval

The researchers wish to state that the research conducted in this study received ethical approval from the International Livestock Research Institute's Ethics Committee. All procedures performed in this study were in accordance with the ethical standards of the institution. Moreover, all study participants were asked for informed consent, with a consent form.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgement

This work was carried out with support from the CGIAR Scaling for Impact Program. We would like to thank all funders who supported this research through their contributions to the CGIAR Trust Fund (www.cgiar.org/funders).

Data availability

Data will be made available on request.

References

Altshul, H.J., 2025. Framework Analysis Report: Responsible Innovation and Scaling in CGIAR Instruments.

Ankrah, D.A., Freeman, C.Y., Afful, A., 2020. Gendered access to productive resources – evidence from small holder farmers in Awutu Senya West District of Ghana. Sci. Afr. 10, e00604. https://doi.org/10.1016/j.sciaf.2020.e00604.

Barrett, C.B., Benton, T., Fanzo, J., Herrero, M., Nelson, R.J., Bageant, E., Wood, S., 2022. Socio-technical innovation bundles for Agri-food systems transformation. In: Socio-Technical Innovation Bundles for Agri-Food Systems Transformation. Springer International Publishing, Cham, pp. 1–20.

Cornwall, A., 2003. Whose voices? Whose choices? Reflections on gender and participatory development. World Dev. 31 (8), 1325–1342.

Damtew, E., Douthwaite, B., Schut, M., Sartas, M., Leeuwis, C., 2023. Improving scaling performance in research for development: learning from a realist evaluation of the scaling readiness approach. Eur. J. Dev. Res., 0123456789 https://doi.org/10.1057/s41287-023-00586-w.

Darnhofer, I., 2015. Socio-technical transitions in farming: Key concepts. In: Transition Pathways towards Sustainability in Agriculture: Case Studies from Europe. CABI, Wallingford UK, pp. 17–31. https://scholar.sun.ac.za/.

Douthwaite, B., Hoffecker, E., 2017. Towards a complexity-aware theory of change for participatory research programs working within agricultural innovation systems. Agric. Syst. 155, 88–102.

Douthwaite, B., Mayne, J., McDougall, C., Paz-Ybarnegaray, R., 2017a. Evaluating complex interventions: a theory-driven realist-informed approach. Evaluation 23 (3), 294–311.

Douthwaite, B., Mur, R., Audouin, S., Wopereis, M., Hellin, J., Saley Moussa, A., Bouyer, J., 2017b. Agricultural research for development to intervene effectively in complex systems and the implications for research organizations.

Ekboir, J.M., 2003. Research and technology policies in innovation systems: zero tillage in Brazil. Res. Policy 32, 573–586.

Esterhuyse, H.W., 2019. Understanding the societal impact of research through productive interactions and realist theory-based evaluation: select cases of agricultural research in South Africa. https://scholar.sun.ac.za.

Ewell, H., Smolka, M., Leeuwis, C., McGuire, E., Ngissah, E., 2026. Too late for responsibility? Dialogic prompts for responsible decision-making across agricultural research for development.

Funnell, S.C., Rogers, P.J., 2011. Purposeful Program Theory: Effective Use of Theories of Change and Logic Models. John Wiley & Sons.

Gaitán-Cremaschi, D., Klerkx, L., Duncan, J., Trienekens, J.H., Huenchuleo, C., Dogliotti, S., Contesse, M.E., Rossing, W.A.H., 2019. Characterizing diversity of food systems in view of sustainability transitions. A review. Agron. Sustain. Dev. 39, 1. https://doi.org/10.1007/s13593-018-0550-2.

Gebreyes, M., Mekonnen, K., Thorne, P., Derseh, M., Adie, A., Mulema, A., Kemal, S.A., Tamene, L., Amede, T., Haileslassie, A., Gebrekirstos, A., Mupangwa, W.T., Ebrahim, M., Alene, T., Asfaw, A., Dubale, W., Yasabu, S., 2021. Overcoming constraints of scaling: critical and empirical perspectives on agricultural innovation scaling. PLoS One *16* (5 may), 1–19. https://doi.org/10.1371/journal.pone.0251958.

Gilmore, B., McAuliffe, E., Power, J., Vallières, F., 2019. Data analysis and synthesis within a realist evaluation: toward more transparent methodological approaches. Int. J. Qual. Methods 18. https://doi.org/10.1177/1609406919859754;JOURNAL:JOURNAL:IJQA;CTYPE:STRING:JOURNAL.

Grin, J., Rotmans, J., Schot, J., Geels, F., Loorbach, D., 2010. Routledge Studies in Sustainability Transitions. Routledge, New York.

Habermann, B., Vogl, C.R., Mekonnen, K., Bekele, K., Felt, U., 2021. Farmers and scientists in AR4D: looking at a watershed management project through an STS lens. NJAS: Impact in Agricultural and Life Sciences 93 (1), 126–151. https://doi.org/10.1080/27685241.2021.1993095.

Hall, A., Clark, N., 2010. What do complex adaptive systems look like and what are the implications for innovation policy? J. Int. Dev. 22, 308–324.

Hall, A., Sulaiman, V.R., Clark, N., Yoganand, B., 2003. From measuring impact to learning institutional lessons: an innovation systems perspective on improving the management of international agricultural research. Agric. Syst. 78 (2), 213–241.

Hillenbrand, E., Obi, C., McDougal, C., Rossignoli, C.M., 2024. Scaling-out gender transformative approaches in agriculture research for development. Gend. Technol. Dev. 28 (1), 50–77. https://doi.org/10.1080/09718524.2023.2272348.

Hussain, F., Brown, S., 2018. Developing a comparative framework of ICT4D initiatives in the global south: a critical realist approach. Electron. J. Inf. Syst. Dev. Ctries. 84 (6), e12055.

Kaboub, F., 2004. Realistic EvaluationRay Pawson and Nick Tilley, sage, London, 1997, 256 pages. Soc. Sci. J. 41 (1), 153–154. https://doi.org/10.1016/s0362-3319(03)00108-3.

Klerkx, L., Van Mierlo, B., Leeuwis, C., 2012. Evolution of systems approaches to agricultural innovation: concepts, analysis and interventions. *Farming systems research into the 21st century: the new dynamic*, pp. 457–483.

Kruijssen, F., Dhamankar, M., Van Schagen, B., Posthumus, H., 2021. Applying a theory of change based approach to livestock research for development: learnings from the cgiar research program on livestock Agri-food systems. https://www.cgiar.org/funders/.

Leeuwis, C., Boogaard, B.K., Atta-Krah, K., 2021. How food systems change (or not): governance implications for system transformation processes. Food Secur. 13 (4).

Leeuwis, C., Schut, M., Klerkx, L., 2017. Systems research in the CGIAR as an arena of struggle: Competing discourses on the embedding of research in development. In: Agronomy for Development. Routledge, pp. 59–78.

Leeuwis, C., Smolka, M., Ewell, H., McGuire, E., Kihoro, E., Neudert, P., Ngissah, E., Mugambi, S., Schut, M., Klerkx, L., 2026. Moving downstream in responsible innovation: seven building blocks for responsible scaling strategies. J. Responsible Innov. https://doi.org/10.1080/23299460.2026.2739111.

Letty, B.A., Hart, T.G.B., Murugan, S., Naidoo, T., Rai, S., Thiam, D., Zake, J., Sershen, & Annor-Fremong, I., 2022. A Critical Comparative Review of Evidence on Alternative Instruments for Supporting Research and Innovation in Agri-Food Systems in the Global South (Frontiers in Sustainable Food Systems).

Lugalla, I.M., 2008. Gender in the participatory agricultural development and empowerment project 2003–2008 a case study of iringa district tanzania. Institute of Social Studies file:///C:/Users/Ngiss001/Downloads/Lugalla%20WGD%202007-08%20(1).pdf.

MacArthur, J., Carrard, N., Davila, F., Grant, M., Megaw, T., Willetts, J., Winterford, K., 2022. Gender-transformative approaches in international development: a brief history and five uniting principles. Women's Stud. Int. Forum 95, 102635. Pergamon.

Maluka, S., Kamuzora, P., SanSebastián, M., Byskov, J., Ndawi, B., Olsen, Ø.E., Hurtig, A. K., 2011. Implementing accountability for reasonableness framework at district level in Tanzania: a realist evaluation. Implement. Sci. 6 (1), 11.

Maru, Y., Sparrow, A., Stirzaker, R., Davies, J., 2018. Integrated agricultural research for development (IAR4D) from a theory of change perspective. Agric. Syst. 165, 310–320.

McCown, R.L., 2002. Changing systems for supporting farmers' decisions: problems, paradigms, and prospects. Agric. Syst. 74 (1), 179–220.

McDougall, C., Badstue, L., Mulema, A., Fischer, G., Najjar, D., Pyburn, R., Elias, M., Joshi, D., Vos, A., 2021. Toward structural change: Gender transformative approaches. In: Pyburn, R., van Eerdewijk, A. (Eds.), Advancing Gender Equality through Agricultural and Environmental Research: Past, Present, and Future. International Food Policy Research Institute, pp. 365–401.

McGuire, E., Al-Zu'bi, M., Boa-Alvarado, M., Luu, T.T.G., Sylvester, J.M., Leñero, E.M.V., 2024a. Equity principles: using social theory for more effective social transformation in agricultural research for development. Agr. Syst. 218. https://doi.org/10.1016/j.agsy.2024.103999.

McGuire, E., Rietveld, A.M., Crump, A., Leeuwis, C., 2022. Anticipating gender impacts in scaling innovations for agriculture: insights from the literature. World Development Perspectives 25. https://doi.org/10.1016/j.wdp.2021.100386.

McGuire, Erin., Leeuwis, Cees., Rietveld, A.M., Teeken, B., 2024b. Anticipating social differentiation and unintended consequences in scaling initiatives using GenderUp, a method to support responsible scaling. Agr. Syst. 215. https://doi.org/10.1016/j.agsy.2024.103866.

Mosse, D., 1994. Authority, gender and knowledge: theoretical reflections on the practice of participatory rural appraisal. Dev. Chang. 25 (3), 497–526.

Nchanji, E., Puskur, R., Cosmas, L., Ketema, D., Gartaula, H., Mukhopadhyay, P., Mapedza, E., Johnstone, K., Barrett, S., 2023. Bundling Socio-Technical Innovations to Empower Women as Partners and Drivers of Climate Change Solutions. . (Proceedings of the Stakeholder Consultation Workshop).

Ngissah, E., 2024. The nexus between information sources, gender and adoption of fall armyworm management practices in southern Ghana. International Journal of Agriculture Environment and Food Sciences 8 (2), 401–411. https://doi.org/10.31015/jaefs.2024.2.15.

Ngissah, E., Kihoro, E., Leeuwis, C., Legun, K., Cieslik, K., Smolka, M., Abegaz, D., Schut, M., Ewell, E., 2026. Enhancing responsible scaling in AR4D in developing countries: an analysis of decisions and decision support tools across the innovation development trajectory working paper) .

Ngissah, E., Leeuwis, C., Cieslik, K., Freeman, C., 2025. Uberisation of mechanisation: exploring the features and establishment of matchmaker digital tractor-hire platforms in Ghana through the lens of actor network theory. *NJAS: impact in agricultural and* Life Sci. 97 (1). https://doi.org/10.1080/27685241.2025.2531122.

Njuki, J., Eissler, S., Malapit, H., Meinzen-Dick, R., Bryan, E., Quisumbing, A., 2023. A review of evidence on gender equality, women's empowerment, and food systems. In: Science and Innovations for Food Systems Transformation. Springer, pp. 165–189.

Ørngreen, R., Levinsen, K.T., 2017. Workshops as a research methodology. Electr. J. E Learning 15 (1), 70–81, 569.

Parra, J.D., 2023. Realist evaluation and its role in the stages of explanatory research based on critical realism. J. Crit. Realism 22 (5), 859–881.

Pawson, R., 2000. Middle-range realism. European Journal of Sociology/Archives 40, 283–325. https://doi.org/10.1017/S0003975600007050.

Pawson, R., Tilley, N., 2004. Realist Evaluation.

Ponelis, Shana.R., 2015. Using interpretive qualitative case studies for exploratory research in doctoral studies: a case of information systems research in small and

medium enterprises. International Journal of Doctoral Studies 10, 535–550. http://ijds.org/Volume10/IJDSv10p535-550Ponelis0624.pdf.
Rietveld, A., Hellin, J., Fisher, E., Gartaula, H., Kramer, B., Farnworth, C.R., Teeken, B., Lopez, D.E., Mujawamariya, G., Bailey, A., Choudhury, A., 2022. In: CGIAR GENDER Impact Platform (Ed.), A Community of Practice for Gender-Transformative Research Methodologies (CGIAR GENDER Impact Platform Working Paper #007).
Rose, D.C., Sutherland, W.J., Parker, C., Lobley, M., Winter, M., Morris, C., Dicks, L.V., 2016. Decision support tools for agriculture: towards effective design and delivery. Agr. Syst. 149, 165–174.
Sánchez Rodríguez, A.M., MacLachlan, M., Brus, A., 2021. The coordinates of scaling: facilitating inclusive innovation. Syst. Res. Behav. Sci. 38 (6), 833–850. https://doi.org/10.1002/sres.2740.
Sartas, M., Schut, M., Proietti, C., Thiele, G., Leeuwis, C., 2020. Scaling readiness: science and practice of an approach to enhance impact of research for development. Agr. Syst. 183 (May), 102874. https://doi.org/10.1016/j.agsy.2020.102874.
Schut, M., Leeuwis, C., Thiele, G., 2020. Science of scaling: understanding and guiding the scaling of innovation for societal outcomes. Agr. Syst. 184 (April), 102908. https://doi.org/10.1016/j.agsy.2020.102908.
Spielman, D.J., Birner, R., 2008. How Innovative Is your Agriculture?: Using Innovation Indicators and Benchmarks to Strengthen National Agricultural Innovation Systems. World Bank, Washington, DC, USA.
Steinke, J., Ortiz-Crespo, B., van Etten, J., Müller, A., 2022. Participatory design of digital innovation in agricultural research-for-development: insights from practice. Agr. Syst. 195. https://doi.org/10.1016/j.agsy.2021.103313.
Wigboldus, S., Brouwers, J., 2016. Using a Theory of Scaling to Guide Decision Making: Towards a Structured Approach to Support Responsible Scaling of Innovations in the Context of Agrifood Systems. Centre for Development Innovation, Wageningen University & Research.
Wigboldus, S., Klerkx, L., Leeuwis, C., 2020. Making scale work for sustainable development: A framework for responsible scaling of agricultural innovations. In: Science, Technology, and Innovation for Sustainable Development Goals. Oxford University Press, pp. 518–544. https://doi.org/10.1093/oso/9780190949501.003.0025.
Wigboldus, S., Klerkx, L., Leeuwis, C., Schut, M., Muilerman, S., Jochemsen, H., 2016. Systemic perspectives on scaling agricultural innovations. A review: systemic perspectives on scaling agricultural innovations. Agron. Sustain. Dev. 36 (3), 46.
Wigboldus, S., McEwan, M.A., van Schagen, B., Okike, I., van Mourik, T.A., Rietveld, A., Amole, T., Asfaw, F., Hundayehu, M.C., Iradukunda, F., Kulakow, P., Namanda, S., Suleman, I., Wimba, B.R., 2023. Understanding capacities to scale innovations for sustainable development: a learning journey of scaling partnerships in three parts of Africa. Environ. Dev. Sustain. 25, 8197–8231. https://doi.org/10.1007/s10668-022-02394-4.
Woltering, L., 2024. Stuck on Scale: Rethinking Scaling and Systems Change for Agricultural Research for Development for Food Systems Transformation (Doctoral dissertation, Wageningen University and Research).
Woltering, L., Fehlenberg, K., Gerard, B., Ubels, J., Cooley, L., 2019. Scaling – from "reaching many" to sustainable systems change at scale: a critical shift in mindset. Agr. Syst. 176, 102652. https://doi.org/10.1016/j.agsy.2019.102652.
Woltering, L., Valencia Leñero, E.M., Boa-Alvarado, M., Van Loon, J., Ubels, J., Leeuwis, C., 2024. Supporting a systems approach to scaling for all; insights from using the scaling scan tool. Agr. Syst. 217. https://doi.org/10.1016/j.agsy.2024.103927.