



Independent
Science for
Development
Council



ISDC Review of 12 Initiative Proposals

February 2022

Credit: 2015CIAT/ManonKoningstein

Contents

Executive Summary	1
Section 1: Essential Details of Proposal Review	1
Reviewer Composition and Diversity	1
QoR4D Criteria	2
Scoring of Criteria	5
Section 2: High-level Synthesis of Proposals	5
Criteria Across Proposals that Scored as Needing Improvements	6
Conformity to Eschborn Principles	8
Comparative Advantage	8
Observations Related to Companion Document	8
Individual Proposal Reporting	9
1. AgriLAC Resiliente: Resilient Agrifood Innovation Systems Driving Food Security, Inclusive Growth, and Reduced Out-Migration in Latin America and the Caribbean (LAC)	10
2. Foresight and Metrics to Accelerate Food, Land, and Water Systems Transformation.....	17
3. From Fragility to Resilience in Central and West Asia and North Africa (F2R-CWANA)	23
4. Fruit and Vegetables for Sustainable Healthy Diets (FRESH)	33
5. Harnessing Digital Technologies for Timely Decision-Making across Food, Water, and Land Systems.....	41
6. HER+: Harnessing Gender and Social Equality for Resilience in Agrifood Systems	52
7. NATURE+: Nature-positive Solutions for Shifting Agrifood Systems to More Resilient	61
8. Resilient Aquatic Food Systems for Healthy People and Planet	70
9. Resilient Cities Through Sustainable Urban and Peri-urban Agrifood Systems.....	75
10. Securing the food systems of Asian Mega-Deltas for Climate and Livelihood Resilience	83
11. Sustainable Intensification of Mixed Farming Systems (SI-MFS)	89
12. Transforming AgriFood Systems in West and Central Africa (TAFS-WCA)	98
Appendix A: Eschborn Principles	106
Appendix B: Companion Document Review Responses	107

Executive Summary

The external review of 12 Initiatives is an essential part of good governance and quality assurance, delivering benefits for the CGIAR researchers, leadership, and System Council. The main benefit is an assurance that the best possible science is conducted to deliver the intended development outcomes. The reviews presented in this report intend to inform funders' investments in One CGIAR research, to ensure they are appropriately targeted with high chances for success. The backbone of the review criteria stems from the Quality of Research for Development in the CGIAR Context (Qo4RD [2020])¹ and the Eschborn Principles (Appendix A). Qo4RD is a framework that facilitates CGIAR System-wide agreement on the nature and assessment of the quality of science.

To operationalize the Qo4RD framework for the Initiative assessments, ISDC embarked on a codesign process with One CGIAR scientists that resulted in 17 criteria using the four elements of Qo4RD: relevance, scientific credibility, legitimacy, and effectiveness. The Qo4RD criteria also aligned with the Eschborn Principles—a set of codesigned principles developed by system funders and other stakeholders and endorsed by System Council in April 2020. The criteria were framed to ensure proposals presented context understanding, anticipated needs and opportunities of end-users, and built partnerships and activities. The Executive Summary is divided into two sections.

- Section 1 presents essential details necessary to understand the review process
- Section 2 provides a high-level synthesis of the 12 proposals

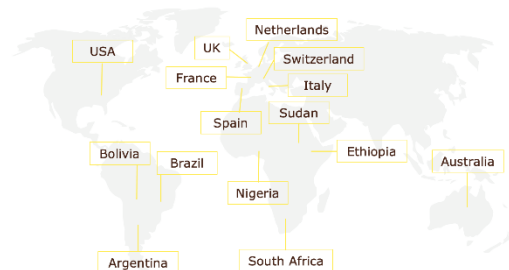
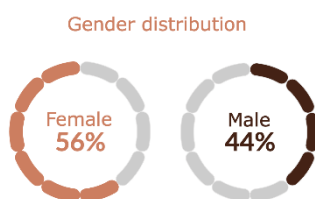
Section 1: Essential Details of Proposal Review

Each Initiative was reviewed by an independent and anonymous review team comprised of three external subject matter experts (SMEs), led by an ISDC member, and supported by the CGIAR Advisory Services Shared Secretariat (CAS Secretariat). The CAS Secretariat identified SMEs through a competitive roster enrollment that contains more than 100 social and biophysical scientists representing more than 25 countries. The CAS Secretariat matched SMEs to proposals based on their expertise to each Initiative review team, with one serving as a coordinator who aggregated and built a consensus among the team, working closely with the ISDC member proposal lead. The ISDC member lead vetted the matched SMEs.

The reviews of the first 19 proposals in 2021 and the reviews in 2022 followed the same 17 Qo4RD criteria. Approximately half of the external SMEs were reviewers in 2021 while the rest were new to the process. Their inclusion in the reviews depended on their availability (December 2021–February 2022) and expertise.

Reviewer Composition and Diversity

The names of all SMEs who served as reviewers for all 32 proposals (including Accelerated Crop Improvement through Precision Genetic Technologies to be reviewed in May 2022)² are listed on the CAS Secretariat [website](#). The information in this report provides analytics on the diversity of the reviewers for this wave of reviews. Each team had a minimum of one social scientist. The composition was 56% female and 44% male, located across 15 countries. Diversity among the review teams was essential because of the cross-cutting goal of the Initiatives and five Impact Areas. The diversity of the reviewers explains, in part, the variance among the Qo4RD scores of each review that can be found in the proposal reporting.



¹ See Qo4RD reference materials at <https://cas.cgiar.org/isdc/Qo4RD>.

² The three external reviewers for Accelerated Crop Improvement through Precision Genetic Technologies have been contracted and are included in the aggregate listing of SMEs.

QoR4D Criteria

Table 1 on the following pages depicts the 17 QoR4D criteria along with each Eschborn Principle and where the criterion should be presented in proposals. Bolded words represent primary QoR4D element.

Table 1. Criteria for Proposal Assessment and Mapped QoR4D Elements, Eschborn Principles, and Related Proposal Sections

Criteria	QoR4D Elements	Eschborn Principles ³	Proposal Section
1. Clearly defined research problem that addresses Impact Areas, is a high priority in the targeted geographies, is well aligned to shared, multi-funder priorities, and is well informed by previous research findings and evaluations	Relevance, Effectiveness	4, 6	Challenge statement 2.1, Learning from prior evaluations and Impact Assessments 2.3, Impact statements 5
2. Evidence that the Initiative is demand driven through codesign with key stakeholders and partners (Investment Advisory Groups, governments, private sector, funders) and research collaborators within and outside CGIAR ⁴	Relevance, Effectiveness	4, 5, 6, 11	Participatory design process 2.6, Challenge statement 2.1, Work Package ToCs 3.2
3. Research questions, objectives, outputs, and outcomes are aligned to the research problem, and are measurable with defined deliverables	Relevance, Effectiveness	4, 7, 10	Work Package ToCs 3.2, Measurable three-year (End of Initiative) outcomes 2.2, Priority-setting 2.4, Management plan 7.1
4. Overall Theory of Change with intended outputs, outcomes, and impacts at scale clearly described. Assumptions are documented, causal linkages are clear, especially the role of partners in driving impact	Effectiveness, Relevance	3, 7, 10	Full Initiative ToC 3.1
4a. Individual work package Theory of Change Work package 1 Work package 2 Work package 3 Work package 4	Effectiveness, Relevance	3, 7, 10	Work Package ToCs 3.2
5. Research methodology and methods (and supporting activities) are fit-for-purpose, feasible, and assumptions and risks are clearly stated	Credibility, Relevance, Effectiveness	2, 5	Work Package ToCs 3.2, Priority-setting 2.4
6. Analysis of trade-offs and synergies across the CGIAR Impact Areas; ex-ante assessment of project benefits provides logical rationale for scaling of impacts	Effectiveness, Credibility	4,6	Projection of benefits 2.7, Result framework 6.1, Impact statements 5
7. Evidence that the Initiative will likely lead to impact at scale through integrated systems approaches that drive innovation in research and partnerships, including linking to and leveraging of other Initiatives within and outside CGIAR	Effectiveness, Credibility, Relevance	5, 6, 9, 11	Projection of benefits 2.7, Work Package research plans and ToCs 3.2
8. Ethics, including equitable partnerships, information disclosure, biases, and potential conflicts of interest are considered; proposal defines how formal research ethics approvals will be sought/granted ⁵	Legitimacy, Credibility	11	Policy compliance and oversight 8

³ See Appendix A for Eschborn Principles

⁴ The types, range, and roles of partners need to be fully explained. For example, partners involved in research implementation may be different to those partners needed for delivery of outcomes and scaling of impacts and they will have different roles in codesign and codelivery. How these partners have been included in the Initiative design process needs to be described with evidence of their support.

⁵ Proposal do not include individual Initiative ethic statements but robust all-CGIAR policies and mechanisms section. Initiatives will confirm alignment with CGIAR Research Ethics Policy. This was a CGIAR decision during proposal development.

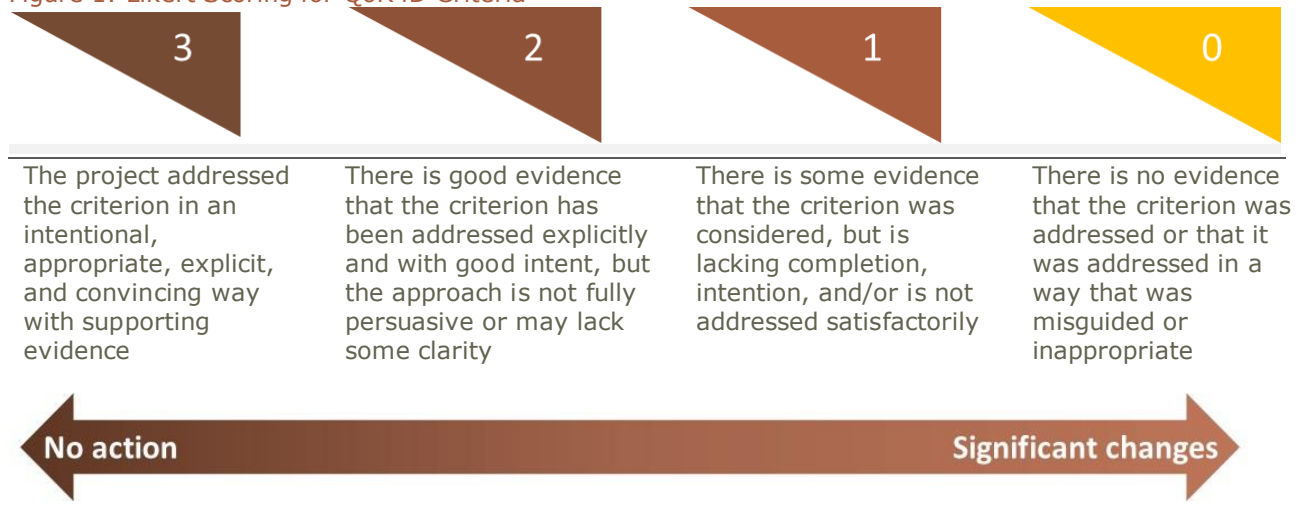
Criteria	QoR4D Elements	Eschborn Principles ³	Proposal Section
9. Research design and proposed implementation demonstrates gender and social inclusion that can be tracked in outcomes	Legitimacy, Effectiveness	2	Gender equality, youth & social inclusion 5.3, Projection of benefits 2.7
10. A risk framework that details main project risks and mitigation actions across areas of science, funding, operations, partnerships, ethics, and environment	Credibility, Legitimacy, Relevance	9	Risk assessment 7.3
11. CGIAR capacity and its comparative advantage and appropriateness to lead the work is justified. This includes the skills, diversity and multi-/trans-disciplinarity of the research team and approaches to meeting gender and diversity targets	Relevance, Legitimacy, Effectiveness	2, 5, 6	Comparative advantage 2.5, Initiative team 9.1, Gender, diversity and inclusion in the workplace 9.2
12. Capacity building within project teams, partners, and stakeholders captured in capacity development plan. This can include development of early career researchers and partner staff, support/empowerment for under-represented stakeholders, and building partner networks	Credibility, Legitimacy	2, 6	Capacity development 9.3
13. Project management mechanisms and (if applicable) additional scientific oversight and governance measures effectively and efficiently support the Initiative objectives ⁶	Legitimacy, Credibility	7, 11	Management plan and Risk assessment 7, Research governance 8.1
14. Justified and transparent costing explicitly linked to expected Research for Development results	Legitimacy, Effectiveness	8	Financial Resources 10
15. Anticipated research outputs (knowledge, technical, or institutional advances, specific technologies or products, policy analyses) are described and knowledge/gaps they will fill are evident. Protocols for open-data and open-access compliance are evident in plan (including budget)	Credibility, Effectiveness	4, 9	Work Package research plans and ToCs 3.2, Open and FAIR data assets 8.2
16. Monitoring, evaluation & learning (MEL) plan for the Initiative is clearly defined, with flexibility to adapt. MEL plan supports effective management and learning, including baseline data collection, and evaluative and review processes corresponding to stage-gates and course-correction decisions. MEL occurs during the life of Initiative and is used proactively to reflect on and adapt the Theory of Change, where appropriate	Credibility, Effectiveness, Legitimacy	4, 7, 10	MELIA plan 6.2, Planned MELIA studies and activities 6.3, Measurable three-year (End of Initiative) outcomes 2.2
17. Well-defined plan for Initiative-level evaluation and impact assessment based on expected end-of-Initiative outcomes and impact. Links between the impact assessment plan and indicators in the Theory of Change are clear	Effectiveness, Relevance	3, 4, 10	MELIA plan 6.2, Planned MELIA studies and activities 6.3, Full Initiative ToC 3.1, Work Package ToCs 3.2 Projection of benefits 2.7

⁶ Each proposal had standard text on CGIAR research governance arrangements already agree for section 8.1 This was a decision during proposal development.

Scoring of Criteria

Figure 1 below describes the scoring for the 17 QoR4D criteria.

Figure 1: Likert Scoring for QoR4D Criteria



Section 2: High-level Synthesis of Proposals

Initiative proposals are unique, and each proposal had its strengths and weaknesses, which are described qualitatively in proposal reports. This section presents relevant aggregated quantitative data based on the QoR4D criteria and feedback from reviewers for 12 proposals. ISDC acknowledges the difficulties of developing proposals under the current One CGIAR circumstances and despite the positive aspects of the Initiatives, the reviews highlight areas of improvement.

Similar to the first 19 proposals reviewed, ISDC found that many Initiative proposals were lacking in solid scientific justifications outlining why the research is needed. The proposals continued to have a strong development and impact emphasis, and with tight word limits, attentiveness to some of the underpinning best practice in presenting scientific research appears to have been sacrificed. ISDC again urges the Initiative Design Teams of **all proposals** to provide a much better balance between the science and development before implementation. **Knowledge gaps that inhibit further development should be articulated, followed by the research questions and their underlying hypotheses.**

Along with the balance of science and development in the Initiatives, is the concern that many proposals present overall vague and poorly defined research questions. As one review commented, "Research questions are very broad and could be more explicit." Another review stated that under the header of Research Questions, only broad descriptions of what is proposed was presented.

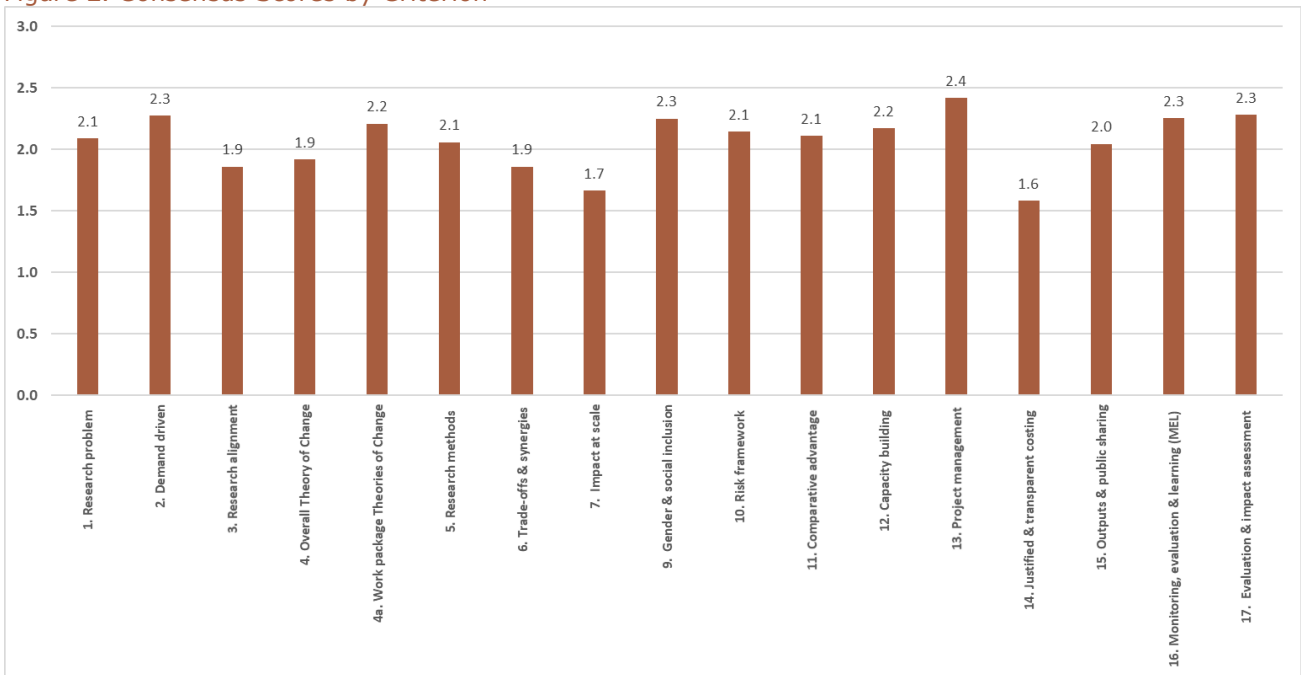
Further in this report, observations of the Companion Document are made despite the Companion Document only including the first 19 Initiatives reviewed. A missed opportunity exists to amalgamate some or parts of Initiatives that address very similar issues (e.g., Nature+, Transformational Agroecology, and SI-MFS). ISDC also acknowledges the difficulty of broadening the 2030 Research and Innovation Strategy beyond the traditional expertise of CGIAR. Reviewers asked if CGIAR has the social science capacity, which relates to the System's comparative advantage.

Figure 2 below shows the average consensus score for each criterion across the 12 Initiatives. Criterion 8, that focused on ethics, was not scored because Initiative Design Teams were instructed to use standard language across proposals. While this directive helped to address some of the procedural aspects of compliance with ethical guidelines and requirements, it is not sufficient to judge the broader aspects of legitimacy of the research process. At the aggregate level, the criteria that received scores less than "2" include the following.

- research alignment (criterion 3)
- overall ToC (criterion 4)
- trade-offs and synergies (criterion 6)
- impact at scale (criterion 7)
- justified and transparent costing (criterion 14)

Overall, this round of proposals received five aggregate scores less than “2.” For the initial 19 reviews, only criterion 14 (justified and transparent costing) received a score less than “2” at the aggregate level. A score of two was described as, “There is good evidence that the criterion has been addressed explicitly and with good intent, but the approach is not fully persuasive or may lack some clarity.”

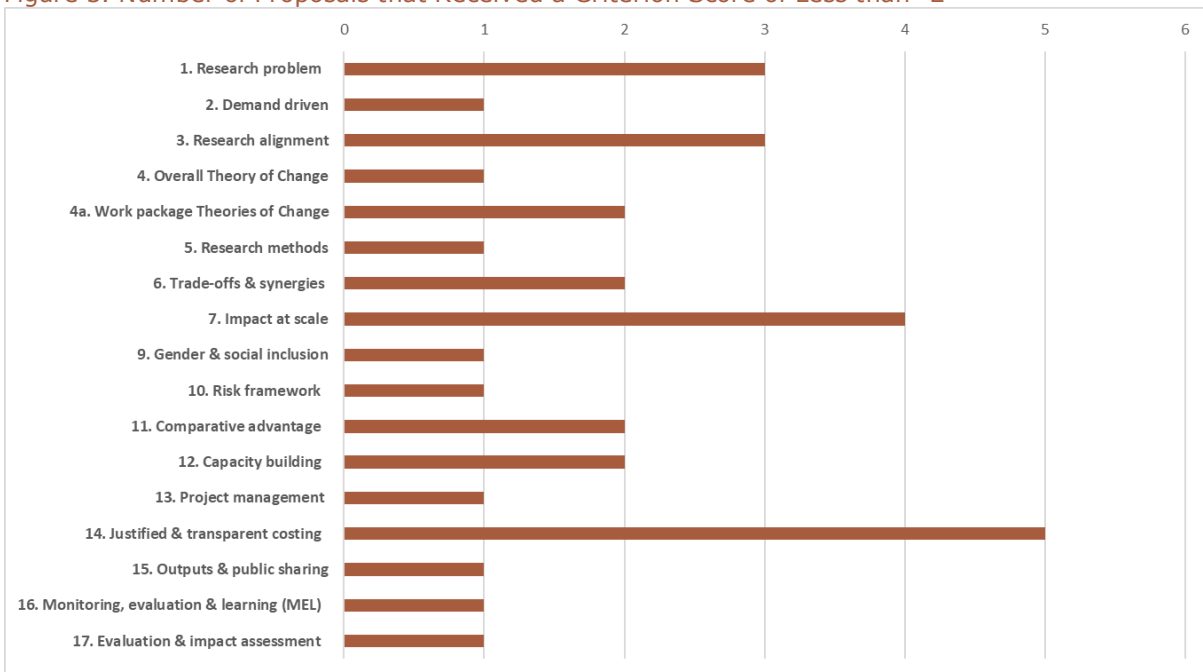
Figure 2. Consensus Scores by Criterion



Criteria Across Proposals that Scored as Needing Improvements

Figure 3 shows the number of proposals with a consensus score of below “2.” A score of “1” was defined as, “There is some evidence that the criterion was considered, but is lacking completion, intention, and/or is not addressed satisfactorily.” Many of the consensus scores were not round numbers (integers) because the decision on how to report the consensus was up to the review teams; some teams decided to use the average of their scores as representing consensus. Each proposal summary found in this report includes individual reviewer scores, as well as the team’s consensus score to show the variance among reviewers and the consensus scores if the mathematical average was not used.

Figure 3. Number of Proposals that Received a Criterion Score of Less than “2”



To further refine areas for improvement, Tables 2, 3, and 4 below highlight criteria where at least five (Table 2), four (Table 3), and two (Table 4) Initiatives scored less than “2.”

Five of the 12 proposals (42%) scored the criterion related to budgets lower than a “2.” The main cause for the low score was the lack of information and insufficient granularity due to budget template limitations. Reviewers expressed concerns that the costing lacked clarity; at a minimum, budgets should detail salaries, operating, and capital investment costs. The proposal budgets only presented line items by Work Package and country, organized by year. Expectations of co-investments from partners were also absent. Like the first 19 proposals reviewed, this criterion received the lowest scores overall citing the same rationale. One comment was, “Justification for the budget is not presented. This makes it difficult to evaluate if there is an appropriate balance of financial resources across the Work Packages, and whether those funds are indeed sufficient/realistic for the proposed activities.”⁷

Table 2. Criteria with Eight Proposals Scoring Less than “2”

Criteria	QoR4D Elements	Eschborn Principles
14. Justified and transparent costing explicitly linked to expected Research for Development results	Legitimacy, Effectiveness	8

Four proposal reviews (33%) cited the impact at scale (criterion 7, Table 3) as needing improvements. One review asked, “The indicators of outcomes and impact seem unrealistic. How can this Initiative in such a short period of time (2025) achieve the outcomes/impact that the CRPs failed to achieve in 10 years? What are the substantial changes in approach and methods brought by this Initiative to substantiate such claims?”

The feasibility of CGIAR-led research was a challenge highlighted in the recent CAS Secretariat Independent Evaluation CRP synthesis. “The CRPs have produced high-quality and relevant research products, but there is a disconnect between the time to impact and the lifespan of a project or CRP. From one side, the long-term and complex nature of the research-to-impact pathway make it difficult to obtain the commitment of partners required to enable uptake into use. From the other side, donor expectations put pressure on the CRPs to articulate short- to medium-term development outcomes that are unrealistic for CGIAR research by itself and put weight on the CGIAR Centers to focus on more adaptive research, for which they may not have a comparative advantage, at the expense of longer-term exploratory research” (p. 18).⁸

Impact at scale also requires successful partnerships. One review reflected, “How the partnerships will be set up and which local research institutions and universities that will be involved should be explained in greater depth in the proposal.” Four of the 19 first reviewed proposals also received a score of less than “2” for impact at scale. This equates to **26% of the 31 proposals** reviewed as having impact at scale as needing improvement.

Table 3. Criteria with Five Proposals Scoring Less than “2”

Criteria	QoR4D Elements	Eschborn Principles
7. Evidence that the Initiative will likely lead to impact at scale through integrated systems approaches that drive innovation in research and partnerships, including linking to and leveraging of other Initiatives within and outside CGIAR	Effectiveness, Credibility, Relevance	5, 6, 9, 11

Table 4 below provides details on the two additional criteria that scored less than “2,” which occurred in three proposals (25%). Having a score of below “2” is a grave indication for criterion 1 on a “clearly defined research problem” since as one review stated, “A poor problem definition leads to unjustified research questions and to a weak Theory of Change.” Only one proposal of the first round of 19 reviewed received a score below a “2.”

Commentary on research alignment (i.e., criterion 3) also is provided under the heading of “strengths and weaknesses of the reviews” for each of the Initiatives. One review noted, “There is a mismatch between the research problem, which suggests a focus on system-level interactions and outcomes to address a broad set of cross-sectoral challenges (including poverty reduction, improved nutrition, and increased equity) and the Work Packages, which appear to substantially narrow the focus of the Initiative to climate change impacts and the agriculture sector.” The first 19 proposals reviewed had four proposals

⁷ ISDC was aware after the first round of reviews that, similar to the ethics criterion, due to the nature of the template and instructions, criterion 14 could have been noted as not applicable. However, to maintain consistency in the process between the first and second rounds the criterion was maintained.

⁸ CAS Secretariat (CGIAR Advisory Services Shared Secretariat). (2021). Synthesis of Learning from a Decade of CGIAR Research Programs. Rome: CAS Secretariat Evaluation Function.

(21%) that scored less than “2” for the research alignment criterion, resulting in **23% (7) of the 31 Initiative proposals** reviewed scoring less than “2.”

Table 4. Criteria with Four Proposals Scoring Less than “2”

Criteria	QoR4D Elements	Eschborn Principles
1. Clearly defined research problem that addresses Impact Areas, is a high priority in the targeted geographies, is well aligned to shared, multi-funder priorities, and is well informed by previous research findings and evaluations	Relevance, Effectiveness	4, 6
3. Research questions, objectives, outputs, and outcomes are aligned to the research problem, and are measurable with defined deliverables	Relevance, Effectiveness	4, 7, 10

Conformity to Eschborn Principles

The CGIAR 2030 shared agenda is to be supported by pooled funding. Since the QoR4D criteria aligned with the Eschborn Principles (Appendix A), the scoring echoes how well those Principles are addressed. The lowest scores were associated with Eschborn Principles 8, 6, and 4 (Table 5).

Table 5: Eschborn Principles Associated with Low QoR4D Scores

Eschborn Principle and Proposal Concern
<p>#4. A clear problem statement, rigorous priority-setting, purpose-driven solutions and a focused set of metrics for success.</p> <p>Concern: Excessive vagueness and lack of rigorous scientific problem definition, characterize some Initiatives. Furthermore, there might be some discrepancies between research problem and outcomes that address a broad set of cross-sectoral challenges. In some cases, there is a need to specifically define problems per country/region, with identifiable knowledge gaps, and consequent/justified research needs. Overall, there is little discussion of metrics to be developed and incorporated into the Initiatives.</p>
<p>#6. Apply operational and geographic focus in areas of recognized CGIAR competencies and achieve impact by working strategically with partners that have complementary competencies, at all stages of research-for-development.</p> <p>Concern: Broadening partnerships to achieve impacts is recommended. The codesign of activities with a wider network of partners, particularly NGOs, private sector, and governments, is critical for scaling. Involving local organizations should be expanded at all stages of the research.</p>
<p>#8. Realistic and transparent costing explicitly linked to expected results.</p> <p>Concern: Most reviewers consider that budget material is insufficiently detailed to assess if costing is linked to expected results. That said, Initiative proposals follow a required template on financial resources, i.e., yearly budget breakdown by activity and by geographical area. Such a budget breakdown does not allow rigorous linkage between budget and expected results. Lack of systematic granularity in the distribution of critical funds (e.g., personnel, capital equipment, capacity building, partnership development, and investments into innovations) is coupled with the absence of narrative justifying the expenditures.</p>

Comparative Advantage

The need for a better understanding and articulation of CGIAR’s comparative advantage is evident from the vastly different interpretation of what constitutes comparative advantage across the Initiative proposals. Initiatives mostly interpreted comparative advantage in terms of input variables such as reputation and geographic location of researchers rather than any comparative advantage in delivering outcomes. Across the 12 proposals reviewed, the comparative advantage criterion (11) received a score of 2.1, compared to 2.4 for the first 19 proposals. During this review, several teams questioned if CGIAR has the in-house social science expertise essential to successfully implement the Initiatives. In contrast, the lack of horizontal research “with” and not “for” stakeholders was evident in some proposals.

Observations Related to Companion Document

During the first 19 reviews, System Council asked ISDC to review the Companion Document to the 2022–2024 CGIAR Investment Prospectus. The consensus report for each proposal included a response to the question: “Does the Initiative align with the cohesion of the portfolio as described in the Companion

Document?” Although the Companion Document was not updated to include the entire portfolio of 32 Initiatives, Appendix B provides responses to this question for the 12 Initiatives.

Unfortunately, with the Companion Document only including the first 19 Initiative, it does not include SI-MFS that has strong linkages with other Initiatives such as HER+ and EiA. There is no mention of links with seed systems or plant breeding Initiatives (SeEdQual and ABI) or with climate mitigation strategies (MITIGATE+). Such links are just as important as links with agronomy and pest management for sustainable intensification in farming systems.

Coherence among the Initiatives continues to be a concern that is highlighted across the reviews. Some Initiatives identify divergent drivers and propose contradictory solutions particularly in agronomy and farming systems where approaches proposed in Transformational Agroecology and Nature+ appear to be at odds with Excellence in Agronomy and some duplication of activities with Genetic Innovation Initiatives is apparent. In addition, Initiatives read as if they were developed independently, which may lead to a significant risk of contradictions, duplications, overlaps, and gaps. A key aspect of this is that partnerships are not detailed. Without having a solid understanding of the partnerships, assessment of who does what is difficult.

Although the merging of Initiatives was proposed during previous System Council meetings, only one merged before the external ISDC reviews. Merging parts of Initiatives could address some concerns of duplication and contradictory solutions.

A special concern arises regarding Regional Integrated Initiatives, which in general are not cohesive, resulting in missed opportunities for mutual learnings and duplication of efforts. CGIAR needs to give careful thought to whether Regional Integrated Initiatives are intended as a different local of research than the Global Thematic Initiatives or if, instead, Regional Integrated Initiatives are to plan mainly an integrative, translational, and partnership-strengthening role. The Regional Integrated Initiative proposals, as written, are very much the former, resulting in gaps, duplication of efforts, and potentially counterproductive intra-CGIAR competition between global and regional Initiatives. The Initiatives will probably generate regional and international public goods (i.e., beyond their respective geographical scope) that have not been mapped across the Initiatives. Furthermore, there is no strategic approach to institutionalize capacity strengthening at any level, within and across Initiatives. This will require serious attention, commitment, and resources, especially to help strengthen technical capacity for adaptive research, monitoring, impact evaluation, foresight and trade-off analyses with local and regional partners. CGIAR may be able to leverage advanced research institution partners in the Global North to assist; it does not need to shoulder those responsibilities alone.

Individual Proposal Reporting

All proposal review reports are presented in the following section. ISDC developed a consensus template for review teams to complete in coordination with an ISDC member. The template included a mix of qualitative commentary (e.g., review summary and actional recommendation(s) and three strengths and weaknesses) and quantitative consensus QoR4D scores. To provide additional information, the CAS Secretariat developed Figures to highlight QoR4D individual reviewer score variance and the resulting consensus score for each criterion. The proposals are presented in alphabetical order. The review reports received light, technical editing for understanding and clarity.

1. AgriLAC Resiliente: Resilient Agrifood Innovation Systems Driving Food Security, Inclusive Growth, and Reduced Out-Migration in Latin America and the Caribbean (LAC)

Review Summary and Actionable Recommendation

The AgriLAC Resiliente proposal focuses, mainly, on the use of digitally enabled agro-advisory services to sustainably intensify production, improve nutrition and appropriately anticipate and manage climate risks in the local and regional agrifood systems (AFS). The Work Packages address the linkages between resilient AFS, environmental sustainability, and the socioeconomic situation of people in the region, particularly vulnerable groups such as women and youth who are more likely to migrate out of the region because of the lack of opportunity in agriculture. The proposal clearly explains how the outcomes will contribute to meet the 2030 UNFSS, SDG, and COP26 targets by increasing the climate resilience of vulnerable farming communities, expanding employment opportunities, and reducing the pressure of out-migration in the region. These targets will be achieved by promoting nutrition-sensitive socio-ecological-technological (SET) innovations, empowering agrifood system actors through a digital ecosystem of agro-advisory services, integrating low-emission strategies, setting up InnovaHubs for learning and knowledge management and scaling up these interventions to shape national and regional policies.

Although MELIA activities are discussed in terms of complexity involved in all phases of the proposal, they need to carefully be adapted to the realities confronted during implementation. An overall weakness relates to the challenges with credibly discerning the impacts of an Initiative like this. Parallel Initiatives will seek to achieve similar and related goals, making isolation and identification of Initiative Work Packages' impacts particularly challenging. A much sharper articulation of how scaling is expected to work and what the drivers or enablers—and their related challenges or barriers are—is needed urgently. These could vary quite substantially across the five Work Packages and target countries.

Overall Strengths of Proposal

The main strength of the proposal is the holistic- and evidence-based approach to the main challenges of agriculture in an important part of the LAC region. It addresses how vulnerable the food systems in the LAC region are and how aspects such as climate change are impacting and widening the economic and social disparities in the region. Additionally, the Work Packages proposed are coherent and cohesive and tackle the main challenges with pragmatic solutions which are argued to be scalable and are likely to have a broad impact in the region.

All five Work Packages have a strong socioeconomic component and emphasize the need to address gender equality, youth, and social inclusion. The proposal places much emphasis on mechanisms in which nutrition-sensitive SET innovations can be introduced or codesigned with stakeholders and beneficiaries. This includes local nutrition practices that have been demonstrated to work and that can be adapted to the local context and adopted for broader impact through the implementation and use of InnovaHubs. The proposed Work Packages dissect the vulnerabilities of the food systems in the LAC region and focus on scalable solutions to curb out-migration, through youth and social inclusion and concrete actions to reduce gender inequalities.

AgriLAC Resiliente proposes mixed quantitative and qualitative methods and socio-economic research as equal pillars to agronomic and genetic research and innovations (i.e., climate-smart, drought-tolerant crops) to achieve a resilient agrifood systems and a profound system transformation. Past lessons of implemented projects are taken as learning experiences and the proposal clearly states how inclusive, participatory approaches in research and the planning process can harness better results. Socio-economic aspects and social inequalities are at the core of the vulnerabilities of the food systems in the LAC region and the proposal stresses the need for collaboration, multi-sectorial involvement, and an inclusive, user-centered approach.

Overall Weaknesses of Proposal

The Initiative timeline is ambitious, given the tight timeframe for the implementation and the scale it seeks to reach in two cycles. AgriLAC Resiliente expects to help seven countries to design and deploy low-emission, resource efficient pathways to support AFS by increasing the climate resilience of the rural poor. However, achieving the outcomes and upscaling from four countries (Guatemala, Honduras, El Salvador, and Nicaragua) to Colombia, Peru, and Mexico is an undertaking that may not be feasible. Taking into consideration the differences in ecosystems and socio-economic situations, as well as the root causes of out-migration, a focus on the Central American region is recommended. The four Central American countries share common challenges and are geographically and culturally very close. However, the socio-economic-environmental-political complexities of Colombia, Peru, and Mexico must be more carefully considered, despite the long-lasting experience the CGIAR international research centers have in these three countries. A sharper articulation of how scaling is expected to work and of what the drivers or enablers and nature of the challenges or barriers that scaling of the work under

each of the WPs are likely to encounter—which could vary quite substantially across the five Work Packages—is needed.
The proposal mentions collaboration and partnerships and the importance of transdisciplinary research as the basis of the planned activities, with a particular focus on the inclusion of social sciences and mixed methods to strengthen agronomic research. External partnerships will be critical to increase CGIAR effectiveness. A core issue will be strengthening agrifood system resilience through the inclusion of knowledge generated and applied in collaboration with local research institutions. How the partnerships will be set up and which local research institutions and universities will be involved should be explained in greater depth in the proposal. The collaboration and engagement with local institutions and stakeholders will be central to the success of the initiative and should be considered as a fundamental pillar of AgriLac Resiliente.
The LAC region is fraught with political instability, which affects government engagement and stakeholder involvement. Depending on the interests and priorities of individual governments, AgriLAC Resiliente may face the risk of not achieving the proposed agenda. A contingency plan or further explanation of how to mitigate this risk would be helpful. A comprehensive engagement of local stakeholders can reduce risks.

Areas of Divergence among Review Team and ISDC Resolution

Not applicable.

Does the Initiative Align with the Cohesion of the Portfolio as Described in the Companion Document?

The outputs, outcomes, and impacts of the Initiative are well-aligned with the CGIAR Impact Areas and the SDGs. The proposal states that the work will be conducted in collaboration with the regional and country offices. This alignment to the CGIAR Performance and Results Management Framework will allow for a coherent and cohesive integration of the portfolio. The proposal demonstrates cohesion through synergies with local partners. The different elements of the CGIAR Performance and Results Management Framework have been taken into consideration and are explicitly mentioned in the proposal. Regional and activity integration, co-creation with local partners, and country offices are an important part of the proposal and addressed in the Work Packages.

The reviewers scored each of the 17 QoR4D criterion individually. The review teams then built a consensus score. In the majority of reviews, the consensus score was the same as the mathematical average of the three individual reviewer scores. The criteria that received between a 0 or 1 are presented with a rationale below. Although not required, in order to keep review reports brief and readable, some teams opted to record in their reports their consensus-built rationale for scores above 1. Following the criteria is a figure presenting the individual reviewer scores and consensus. Please refer to p. 5 for the Likert scoring definitions.

Criteria	Proposal Sections	QoR4D Elements	Consensus Score
1. Clearly defined research problem that addresses Impact Areas, is a high priority in the targeted geographies, is well aligned to shared, multi-funder priorities, and is well informed by previous research findings and evaluations	Challenge statement 2.1, Learning from prior evaluations and Impact Assessments 2.3, Impact statements 5	Relevance, Effectiveness	3
The problem statement and objectives are strong and addresses the main challenges and drivers that will shape sustainable AFS in the next decade. Evidence of lessons learned within and outside the CGIAR exist to support research questions and the need for the One CGIAR to move forward on the huge challenge of transforming AFS in the target countries, selected through an evidence-based priority setting process. AgriLAC Resiliente is an opportunity to “ensure that the expertise, research evidence base and results—hereto dispersed across various CGIAR Centers and AR4D partners are united”—an important element in the CGIAR reform.			
2. Evidence that the Initiative is demand driven through codesign with key stakeholders and partners (Investment Advisory Groups, governments, private sector, funders) and research collaborators within and outside CGIAR	Participatory design process 2.6, Challenge statement 2.1, Work Package ToCs 3.2	Relevance, Effectiveness	3

Criteria	Proposal Sections	QoR4D Elements	Consensus Score
<p>The proposed agenda has merit and the documentation of consultations to demonstrate codesign and evidence of the Initiative being demand driven are comprehensive and satisfactory. There is ample evidence (see Annex 2) pointing to a consultation process involving the academy, public and private sector representatives, and the civil society in general. Codesign implies to give voice to weak stakeholders to resolve imbalances among them, small farmers and, eventually their associations, those that will ultimately be impacted by the Initiative.</p>			
<p>3. Research questions, objectives, outputs, and outcomes are aligned to the research problem, and are measurable with defined deliverables</p>	<p>Work Package ToCs 3.2, Measurable three-year (End of Initiative) outcomes 2.2, Priority-setting 2.4, Management plan 7.1</p>	<p>Relevance, Effectiveness</p>	<p>2</p>
<p>The Theory of Change is broad and relies on, at times, very strong assumptions that are expressed in generic language. Given the scale of the effort and the space constraints, this is perhaps inevitable. At the same time, it makes it very difficult for an informed reader to discern what these Initiatives might look like, how they will be implemented on the ground, and how this will lead to lasting impact. The texts under the research question headings in each of the sub-sections in section 5 are not always research questions, as in 5.1, where they instead are broad descriptions of what is proposed.</p> <p>The Work Packages include the main research questions, and these are aligned with the expected outcomes and deliverables. The path of obtaining the intended results and outcomes, in all the geographies and scales considered, is expressed in general terms as it is difficult to anticipate how the AFS transition process will evolve across space and time. The management plan is designed to capture the details during the implementation phases and provide evidence for eventual adjustments.</p>			
<p>4. Overall Theory of Change with intended outputs, outcomes, and impacts at scale clearly described. Assumptions are documented, causal linkages are clear, especially the role of partners in driving impact</p>	<p>Full Initiative ToC 3.1</p>	<p>Effectiveness, Relevance</p>	<p>2</p>
<p>The estimates of impacts and the numbers of individuals who will be reached by and benefit from this Initiative are obtained using adoption numbers from empirical studies, information on farmers with similar characteristics, overall population figures and other relevant, secondary data. A weakness here is the treatment of scaling as seemingly frictionless. The constraints identified are difficult to overcome without the generation of intangibles at country level (good will, commitment, and strategic engagement). While the endorsement and support letters provide important evidence of commitment to and support of scaling by key government and other local stakeholders and partners, evidence on adoption often rely on comparatively small samples and type of efforts and commitment (e.g., NGOs vs bureaucrats) that it may be challenging to replicate at scale. It is important to say that trying is not a total failure in this case, but the initiative must develop a very strong communication strategy from the very beginning to align expectations and attract engagement in the first phase selected 4 countries: these concerns need to be addressed in the revised version.</p> <p>The pillars of the proposal Theory of Change are very much associated with two strategic goals: one, to help AFS in LAC to transition to low-emission pathways; and two, enable smallholders' access to diversified sources of income, to discourage out-migration. The five Work Packages are designed to achieve the intended outcomes but depend on strong assumptions that may not materialize. In this regard, it is very important to give credit to others than the CGIAR constituencies themselves and acknowledge contributions from smallholders' work. The proposed dialogues will help and should be at the very heart of the activities to be conducted in a continuing and genuine state-of-art codesign process.</p>			
<p>4.a Individual work package Theories of Change (score individually)</p> <p>Work package 1 Work package 2 Work package 3 Work package 4 Work package 5</p>	<p>Work Package ToCs 3.2</p>	<p>Effectiveness, Relevance</p>	<p>WP1: 3 WP2: 2 WP3: 3 WP4: 3 WP5: 2</p>
<p>Work Packages are all well designed, innovative, connected and very much interdependent.</p> <p>Work Package 1 aims at establishing a scalable model (End of Initiative 1) of nutritional-sensitive and socio-ecological-technological "best bets" linked to the transition to agriculture production systems aligned to the five</p>			

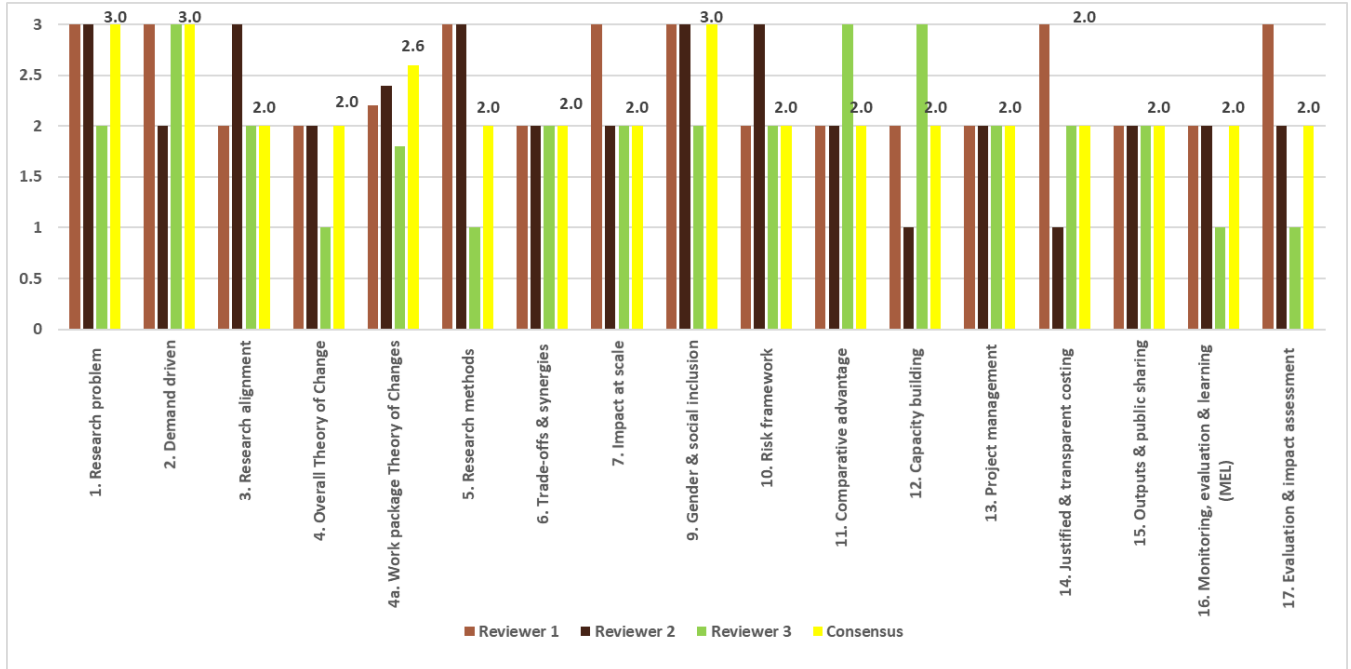
Criteria	Proposal Sections	QoR4D Elements	Consensus Score
<p>Impact Areas, a package that will bring attention to the need to introduce or develop diverse and holistic approaches to agricultural development.</p> <p>Work Package 2 needs to develop, along with the advisory service creation (agro-climatic prediction models), financial tools tailored to the specific needs of the different value chain and AFS stakeholders. These financial tools should include agricultural insurance schemes. The digital ecosystem proposed in the WP has a huge potential and other service-providers and stakeholders can be involved. Work Package 2 seems to be among the most complex and challenging in its pathway to End-of-Initiative-outcome 2. The targets for 2024 seem too ambitious. The refinement of platform prerequisites and its subsequent development could be just good enough and an important legacy for continuing of the Initiative. Tasks attributed to the CGIAR to co-create the basis for the digital decision-making environment to become operational (data hubs, infrastructure, data streams, tools, and services etc.) are huge considering the proposed scales. A task such as prediction models needs large inputs and experimentation at the local level. Thus, partnerships with universities or NARES are essential as are common, comparable protocols.</p> <p>Work Package 3 on synergies and trade-offs among low-emission and sustainable development has a cultural component in a somewhat conservative and traditional AFS in the four first-phase countries, a fact that could hamper change at the scales considered. The revised version should explicitly address this, possibly giving rise to new research questions. The tone is quite technocratic.</p> <p>Work Package 4 is key for the success of the Initiative by facilitating the coordination of knowledge channels and maximize synergies among farmers, farmer's associations, and field researchers, in addition to the links with mainly Work Package 1 and 2. The idea of establishing CoP is commendable, despite of the challenges involved in keeping them properly running.</p> <p>For Work Package 4 and 5, synergies can be explored with global initiatives, CGIAR Initiatives, but it's important to also capitalize on the knowledge and expertise of local research institutions and universities. The action points on how to include these stakeholders in the InnoHubs need to be more clearly explained. Work Package 5 complements the efforts of the other four Work Packages. The Initiative must incorporate improved language to make it clear that results are not CGIAR-only results, so that the science, evidence, and tools informing AFS policies and incentives are genuinely codesigned and developed, with credits going to those who deserve to be credited. Authentic attribution will be essential for CGIAR's future success.</p>			
<p>5. Research methodology and methods (and supporting activities) are fit-for-purpose, feasible, and assumptions and risks are clearly stated</p>	<p>Work Package ToCs 3.2, Priority-setting 2.4, Innovation Packages and Scaling Readiness Plan 4.1</p>	<p>Credibility, Relevance, Effectiveness</p>	<p>2</p>
<p>The research questions are adequate for the nature and objectives of the Initiative, and responsive to contexts identified through the priority setting process. The assumptions are very strong and were transparently stated in the proposal: a similar comment applies to the risks that assumptions may not hold (materialize).</p> <p>However, the description of methods is broad, and it is hard to discern the quality of the design of survey instruments, the analysis that will feed into the tailoring of interventions to address contextual challenges related to gender and youth and the granularity and fine-tuning of implementation efforts. As mentioned above, the lack of systematic engagement with scaling challenges is an overall weakness of the proposal. Another overall weakness relates to the challenge of credibly discerning the impacts of a program of this kind which has ambitions at scale but where there is a high likelihood that parallel initiatives will attempt to achieve similar and related goals: this makes the isolation and identification of impacts of program initiatives particularly challenging and more challenging than for many other initiatives. The narrative for impact assessment plans in section 6.2 is very short and too general to be of value: a similar comment applies to table 6.2.3. which is vague and falls short of expectations of methodological rigor and clarity for a proposal of this size.</p>			
<p>6. Analysis of trade-offs and synergies across the CGIAR Impact Areas; ex-ante assessment of project benefits provides logical rationale for scaling of impacts</p>	<p>Projection of benefits 2.7, Result framework 6.1, Impact statements 5, Innovation Packages and Scaling Readiness Plan 4.1</p>	<p>Effectiveness, Credibility</p>	<p>2</p>

Criteria	Proposal Sections	QoR4D Elements	Consensus Score
<p>The synergies across the CGIAR Impact Areas and the rationale for scaling impacts have been explored. However, scaling of impacts and how stakeholders will be involved could be further clarified. The prospects for scaling of reach and impacts are much less straightforward than the impression one gets from reading the proposal.</p> <p>The data and information collected and the derived calculations to the estimated projections seem to be appropriate (Annex 4). The trade-offs and synergies among CGIAR Impact Areas were adequately addressed in the proposal. The Initiative is very much dependent on data and information management tools to be developed or adapted according to needs. A positive point is the intended plan to document and bring learnings from previous and ongoing experiences that have led to both successful and failed scaling processes. Yet, the Initiative fails to provide clear evidence on how these will be efficiently communicated.</p>			
<p>7. Evidence that the Initiative will likely lead to impact at scale through integrated systems approaches that drive innovation in research and partnerships, including linking to and leveraging of other Initiatives within and outside CGIAR</p>	<p>Projection of benefits 2.7, Work Package research plans and ToCs 3.2</p>	<p>Effectiveness, Credibility, Relevance,</p>	<p>2</p>
<p>There is evidence that the projected benefits are adequate but must be very carefully monitored in order to detect the need for corrective action and propose revisions since the scale of the project is large. An effective communication strategy among all partners and stakeholders is needed as a supporting tool to disseminate the results obtained.</p> <p>It is quite possible that CGIAR teams work with an underlying understanding of how scaling is to be achieved based on the in-house experience with taking advantage of existing partnerships and working with relevant stakeholders who are and will be involved in the implementation and scaling efforts in the countries of interest.</p>			
<p>9. Research design and proposed implementation demonstrates gender and social inclusion that can be tracked in outcomes</p>	<p>Gender equality, youth & social inclusion 5.3, Projection of benefits 2.7</p>	<p>Legitimacy, Effectiveness</p>	<p>3</p>
<p>All the aspects of this criterion seem to be adequately addressed by the proposal and efforts will be undertaken to deal with social and cultural aspects that hamper gender and youth inclusion in co-creating the transition process.</p> <p>While gender and social inclusion feature, including in the brief emphasis of the prominent role of women in LAC agriculture, this often feels like a mandatory add-on: the text would benefit from a sharper articulation of familiarity with the nature of the challenges that may be encountered in the different work packages. While the capacity of CGIAR research staff to address these issues well is not in doubt, the proposal could be improved by including a few brief examples.</p>			
<p>10. A risk framework that details main project risks and mitigation actions across areas of science, funding, operations, partnerships, ethics, and environment</p>	<p>Risk assessment 7.3</p>	<p>Credibility, Legitimacy, Relevance</p>	<p>2</p>
<p>The risk scores appear reasonable and point to the considerable challenges confronting Initiatives with transformative ambitions. At the same time, and again referring to particularly the scaling arguments from above, the nature and granularity of challenges are neither well-articulated, nor, it seems, sufficiently well understood. While the formative and piloting phase (2022-24) is a clear strength, the program could significantly reduce its overall risk exposure through better and more granular Theory of Change preparations for each work package and a much clearer articulation of the intervention design challenges, on the one hand, and enablers of and hurdles to both implementation and scaling success, on the other. One option would be to present each Work Package Theory of Change as a draft that will be made more granular and regularly updated through the first phase (2022-24) piloting and learning.</p> <p>Additionally, some contextual aspects which could pose risks for the implementation of the agenda are missing, such as the political instability which Nicaragua, Honduras, and Guatemala might face during project implementation. The evaluators suggest adding this to the risk matrix.</p>			
<p>11. CGIAR capacity and its comparative advantage and appropriateness to lead the work is justified. This includes the skills, diversity and multi-/trans-disciplinarity of the research team and approaches to meeting gender and diversity targets</p>	<p>Comparative advantage 2.5, Initiative team 9.1, Gender, diversity and inclusion in the workplace 9.2</p>	<p>Relevance, Legitimacy, Effectiveness</p>	<p>2</p>

Criteria	Proposal Sections	QoR4D Elements	Consensus Score
<p>The composition and the competence and breadth and depth of experience of the team members and of the CGIAR are good. The proposal transparently states that the team composition is unlikely to meet CGIAR's gender target of a minimum of 40% women in professional roles (currently estimated to be at 35%) and will not be comprised of individuals from diverse backgrounds. The Initiative proposes to consciously consider diversity beyond gender when recruiting and during other planned activities. Some activities to add female participants on STEM project-related areas of expertise can be considered, for example.</p>			
<p>12. Capacity building within project teams, partners, and stakeholders captured in capacity development plan. This can include development of early career researchers and partner staff, support/empowerment for under-represented stakeholders, and building partner networks</p>	<p>Capacity development 9.3</p>	<p>Credibility, Legitimacy</p>	<p>2</p>
<p>Capacity building is mentioned, but not in enough detail. Early-career researchers, partners, staff included in Initiative team, but with no further detail on transdisciplinary collaboration or how research and implementation teams can work together. This is presented in very general terms, making it very difficult to assess. The training topics are generally well identified, but some section deserves further elaboration not only by means of providing more details on the content but also on institutional responsibilities, duration of courses, methods, and tools to be used, among other aspects.</p> <p>One additional objective could be to think of the program as an opportunity to enhance CGIAR learning and capacity to manage large scale transformative initiatives. What is the strategy for learning from the successes and failures of a transformative initiative of this kind? Given the high risks, some failures are inevitable: acknowledging and ensuring that these provide as important lessons as successes is vital given the very thin overall evidence on the drivers of and hurdles to transformative success.</p>			
<p>13. Project management mechanisms and (if applicable) additional scientific oversight and governance measures effectively and efficiently support the Initiative objectives</p>	<p>Management plan and Risk assessment 7, Research governance 8.1</p>	<p>Legitimacy, Credibility</p>	<p>2</p>
<p>The research governance statement is presented, but further information or link to management plan or risk assessment should be included. The transformative agenda is welcome, but also very ambitious: it seeks to tackle considerable challenges within a relatively short time window. Given the multi-country, high-risk exposure nature of the Initiative and the fact that many of the planned activities may encounter, effective management will be challenging. It is good to see that a scaling coordinator is in place: a few more details about the experience of management team members with leading high-risk, transformative initiatives would be helpful. What is not clear how these structures will relate to the CGIAR governance, a point that could be made explicit.</p>			
<p>14. Justified and transparent costing explicitly linked to expected Research for Development results</p>	<p>Financial Resources 10</p>	<p>Legitimacy, Effectiveness</p>	<p>2</p>
<p>This criterion cannot be properly assessed based on the provided information, despite sounding reasonable for the nature and scope of the initiative. Must be discussed in much more detail before a final decision is made.</p>			
<p>15. Anticipated research outputs (knowledge, technical, or institutional advances, specific technologies or products, policy analyses) are described and knowledge/gaps they will fill are evident. Protocols for open-data and open-access compliance are evident in plan (including budget)</p>	<p>Work Package research plans and ToCs 3.2, Open and FAIR data assets 8.2</p>	<p>Credibility, Effectiveness</p>	<p>2</p>
<p>Protocols for open-data and -access compliance and budget were not included. However, links to Open and Fair Data Assets Policy were added. Work Packages outputs are highly dependent on strong assumptions that might not materialize. The proposal has presented ways and means to deal with this critical aspect. The development of a communication strategy to share concerns over time in this regard would help and seems to be lacking.</p>			
<p>16. Monitoring, evaluation & learning (MEL) plan for the Initiative is clearly defined, with flexibility to adapt. MEL plan supports effective management and learning, including baseline data collection, and evaluative and review processes corresponding to stage-gates and course-correction decisions. MEL occurs during the life of Initiative and is used proactively to reflect on and adapt the Theory of Change, where appropriate</p>	<p>MELIA plan 6.2, Planned MELIA studies and activities 6.3, Measurable three-year (End of Initiative) outcomes 2.2</p>	<p>Credibility, Effectiveness, Legitimacy</p>	<p>2</p>

Criteria	Proposal Sections	QoR4D Elements	Consensus Score
<p>MEL activities, despite the degree of complexity involved in all phases of the proposal, are clearly defined but will inevitably require adaptation as implementation proceeds. Reiterating from above, an overall weakness relates to the challenges with credibly discerning the impacts of a program of this kind which has ambitions at scale and where the likelihood that parallel initiatives will seek to achieve similar and related goals makes the isolation and identification of impacts particularly challenging. The narrative for impact assessment plans in section 6.2 is short and too general: a similar comment applies to the MELIA plan in table 6.2.3 which falls well of short of the expectations for a proposal with a budget of this size.</p>			
<p>17. Well-defined plan for Initiative-level evaluation and impact assessment based on expected end-of-Initiative outcomes and impact. Links between the impact assessment plan and indicators in the Theory of Change are clear</p>	<p>MELIA plan 6.2, Planned MELIA studies and activities 6.3, Full Initiative ToC 3.1, Work Package ToCs 3.2, Projection of benefits 2.7</p>	<p>Effectiveness, Relevance</p>	<p>2</p>
<p>Given the budget, the ambitions, and the fact that the initial phase (2022-24) will involve extensive learning that will likely result in quite extensive intervention design changes, the further fact that 6.2 and 6.3 as well as the Theories of Change are vague, it is difficult to see how this criterion can be satisfactorily met. Outcomes would be more effectively monitored if they include the components of policy implementation such as derived plans, programs, and incentives, linked, where appropriate, to the five Impact Areas, opening a better structure to measure outcomes and impacts.</p>			

The Figure below represents original scoring from each reviewer and consensus scoring for each criterion. *The consensus score across reviewers may not reflect the mathematical average.* For purposes of the Figure, the QoR4D criteria have been shortened. Please refer to p. 3 for full criteria definitions. Please note that criterion 8 on ethics was not scored.



*No consensus score was greater than 0.5 variance from the mathematical average.

2. Foresight and Metrics to Accelerate Food, Land, and Water Systems Transformation

Review Summary and Actionable Recommendation

This proposal builds on longstanding and highly successful modeling efforts in CGIAR. The research challenge defined by the Foresight and Metrics Initiative is central to successful food, land, and water systems transformation. Addressing today's interlinked challenges will require greatly expanded analyses and increased capacity to understand system-level interactions and outcomes across multiple sectors and scales, and to convert these analyses into practical guidance to inform policy options.

However, the priorities and work packages proposed by the Foresight and Metrics Initiative suggest that its approach may hew too closely to current work with IMPACT and RIAPA models. The risk is that FORESIGHT will remain overly focused on the agriculture sector and fail to deliver the more systemic analysis of trade-offs and synergies across multiple sectors that is needed.

More clarity is needed regarding: a) metrics (will new ones be developed? Will existing ones be tested or incorporated in models in new ways?); b) the current baselines and expected targets for outputs and outcomes—in general, MELIA plans need rethinking once there is greater clarity about the focus and outcomes of FORESIGHT; c) relationships with other CGIAR research and how results of CGIAR research and CGIAR innovations will be incorporated into models; d) specification of risks, and better definition of strategies for managing identified risks; and e) what kinds of policies and investments may be modeled for consideration by global, regional, and national decision makers.

Finally, the list of personnel to be recruited appear overly oriented towards the biophysical and technological components of the project compared to social science, including political economy, and expertise in partnership creation, maintenance, and communication. Given the fundamental importance to the Initiative of successful transfer, adaptation and adoption of the tools developed in the Initiative, the Initiative will need to ensure that recruited staff have the skills and experience to accomplish this.

Overall Strengths of Proposal

1. Research problem/challenge statement. There is widespread agreement that the food, land, and water systems transformation required to address today's interlinked challenges will require greatly expanded analyses and increased capacity at global, regional, and national levels. Expanded analyses and capacity are needed to understand system-level interactions and outcomes across multiple sectors and scales, and to convert these analyses into practical guidance to inform policy options.

2. Management plan. The management plan of the Initiative is well organized with both a scientific leadership team and a technical advisory committee, plus specialists for managing financial and other activities.

3. Comparative advantage. The reviewers strongly concur that the CGIAR is well-placed to lead this work and acknowledge the team's extensive track record of excellent work. However, given the recommendations of prior evaluations (2.3), the list of personnel to be recruited appears overly focused on biophysical and technological expertise compared to social scientists, including political economists, health/nutrition expertise, and specialists in partnership creation, maintenance, and communication.

Overall Weaknesses of Proposal

1. Lack of clarity and inconsistency of outputs and outcomes. Different sections focus on different outcomes in terms of science, geographic coverage, and desired measurable impacts. The focus of 2.2 is the six countries; 3.2 includes global actors; 6.1 has some specific indicators not mentioned elsewhere. In 6.1, the current reference values and expected targets in 2025 are never indicated for outputs and outcomes. The MELIA plans need rethinking once there is greater clarity about the focus and outcomes of the Foresight and Metrics Initiative.

2. Mismatch between the research problem, priority-setting, and work packages. The research problem suggests a focus on system-level interactions and outcomes to address a broad set of cross-sectoral challenges. The work packages substantially narrow the focus to climate change impacts and the agriculture sector. While the IMPACT and RIAPA models are well-respected, it is not clear what further development may be required to expand from their current agriculture sector focus to the broader cross-sectoral analyses of synergies and trade-offs implied by the research question.

3. Lack of articulated integration with the rest of the CGIAR. For example, there is no discussion of how agronomic or veterinary expertise will inform the Foresight and Metrics Initiative exercise. Which specific CGIAR Initiatives or innovations will be considered under Work Package 1, for example?

4. Inadequate consideration of risks and mitigation strategies. 7.3 refers almost solely to the risks inherent in working with the six target countries and does not consider risks at regional and global levels. Additional risks are not considered, such as the failure to develop/adapt foresight models and other analytical tools to identify cross-sectoral synergies and impacts—beyond agriculture—to encompass natural resources, climate change effects, health/nutrition, poverty, and equity. There is insufficient attention to mitigation strategies for risks related to poor data, and to address capability and capacity constraints.

Areas of Divergence among Review Team and ISDC Resolution

Not applicable.

Does the Initiative Align with the Cohesion of the Portfolio as Described in the Companion Document?

One of the great challenges posed by the UNFSS was the need to de-silo the agriculture sector, and enable informed decision making about food systems transformation, through analyzing synergies and trade-offs inherent in alternative policy and investment pathways across food, land, water and health sectors. The potential contributions of foresight modeling (and related capacity strengthening at regional and national levels) to meet these challenges would seem to be immense. Yet this proposal falls disappointingly short, especially given the CGIAR's considerable foresight expertise and track record to date.

It is not clear that the vision and planning for the Initiative stretches more than incrementally beyond the usual agriculture sector focus. If it cannot adapt to meet the broader, cross-sector needs—and identify more specifically how existing models will change to do so—it is unclear how the Foresight and Metrics Initiative can serve the linchpin role anticipated for it in One CGIAR and that is so badly needed at country, regional, and global levels. In addition, although the Foresight and Metrics Initiative mentions planned linkages with several other initiatives throughout the proposal, it is not at all clear how these will be effected/coordinated. And, as mentioned elsewhere in this review, there is little/no discussion of how CGIAR expertise will be used to inform the Foresight and Metrics Initiative model development and analyses.

The reviewers scored each of the 17 QoR4D criterion individually. The review teams then built a consensus score. In the majority of reviews, the consensus score was the same as the mathematical average of the three individual reviewer scores. The criteria that received between a 0 or 1 are presented with a rationale below. Although not required, in order to keep review reports brief and readable, some teams opted to record in their reports their consensus-built rationale for scores above 1. Following the criteria is a figure presenting the individual reviewer scores and consensus. Please refer to p. 5 for the Likert scoring definitions.

Criteria that Scored an Average of 1 or 0	Proposal Section	QoR4D Elements	Consensus Score
3. Research questions, objectives, outputs, and outcomes are aligned to the research problem, and are measurable with defined deliverables	Work Package ToCs 3.2, Measurable three-year (End of Initiative) outcomes 2.2, Priority-setting 2.4, Management plan 7.1	Relevance, Effectiveness	1.3
<p>There is a mismatch between the research problem, which suggests a focus on system-level interactions and outcomes to address a broad set of cross-sectoral challenges (including poverty reduction, improved nutrition, and increased equity) and the Work Packages, which appear to substantially narrow the focus of the Initiative to climate change impacts and the agriculture sector.</p> <p>The Foresight and Metrics Initiative will also rely heavily on IMPACT and RIAPA models. It is not clear what further development may be required to expand from their current agriculture sector focus to the broader cross-sectoral analyses of synergies and tradeoffs implied by the research question, to better incorporate health/nutrition and natural resource synergies/trade-offs/options. For example, health and nutrition do not appear as a core feature of the global modeling framework.</p> <p>Further, in 3.1. and 2.7 the main indicator suggested for nutrition, health, and food security is number of people meeting minimum dietary energy requirements, which does not address more recent diet quality and affordability priorities. The global modeling frameworks and descriptions discuss linking to livestock models, but fish and aquatic systems are not mentioned.</p>			

<p>There are some inconsistencies in the different outcomes and measures of outcomes between the sections 2.2 and 3.2, which reflects some basic confusion over partners/ audiences for the research in this proposal. Outcomes in 2.2 focus on "six countries" (presumably those identified at the bottom of p. 6), but numerous other statements in the proposal allude to other actors/investors in global food and agriculture, and as well as the four technical partner countries. Which global and regional partners will be most important for engagement?</p>			
<p>5. Research methodology and methods (and supporting activities) are fit-for-purpose, feasible, and assumptions and risks are clearly stated</p>	<p>Work Package ToCs 3.2, Priority-setting 2.4, Innovation Packages and Scaling Readiness Plan 4.1</p>	<p>Credibility, Relevance, Effectiveness</p>	<p>1.7</p>
<p>The actual methods to be used in strengthening the Foresight and Metrics Initiative analysis are not described in the proposal, but rather the processes for making them more widely available and useful. In particular, the proposal is very vague regarding which innovations or policies might be analyzed and how that might provide useful information. Furthermore, given that metrics is half of the proposal title, there is remarkably little discussion of metrics to be developed/ incorporated into the Foresight and Metrics Initiative.</p>			
<p>6. Analysis of trade-offs and synergies across the CGIAR Impact Areas; ex-ante assessment of project benefits provides logical rationale for scaling of impacts</p>	<p>Projection of benefits 2.7, Result framework 6.1, Impact statements 5, Innovation Packages and Scaling Readiness Plan 4.1</p>	<p>Effectiveness, Credibility</p>	<p>1.3</p>
<p>Sections 2.7 and 5 don't align very well. Section 2.7 uses the projected meta-benefits from the entire CGIAR system and assumes that 3% conservatively accrue to this effort. While better decision making by donors and country policymakers will undoubtedly contribute to better outcomes from CGIAR innovations, this approach seems simplistic.</p> <p>Section 5 does not provide specific metrics other than that policymakers in six countries will have improved capacity to analyze questions associated with the five Impact Areas. One way to assess scale of impact might be to look at these six countries in terms of their relative global importance within these five Impact Areas. The misalignment of sections 2.7 and 5 reflects a broader inconsistency in the proposal in terms of audiences, partners, and impact.</p> <p>Is the project primarily for global audiences, including the rest of the CGIAR? And the work with specific countries an addition to this existing core effort? Or both (as seems to be indicated by the budget)? If both, then ways to measure impact need to be articulated for both levels of activity.</p>			
<p>7. Evidence that the Initiative will likely lead to impact at scale through integrated systems approaches that drive innovation in research and partnerships, including linking to and leveraging of other Initiatives within and outside CGIAR</p>	<p>Projection of benefits 2.7, Work Package research plans and ToCs 3.2</p>	<p>Effectiveness, Credibility, Relevance,</p>	<p>1.3</p>
<p>The Initiative plans to link to and leverage several other Initiatives. However, a major weakness of the proposal is the lack of specificity about these linkages and, more generally, the lack of articulated integration with the rest of the CGIAR. For example, there is no discussion of how agronomic or veterinary expertise will inform the Foresight and Metrics Initiative exercise. Which specific CGIAR Initiatives or innovations will be considered under Work Package 1, for example?</p> <p>Will the Foresight and Metrics Initiative and cooperating initiatives operate from a shared database at country- and regional-level to reduce duplication of effort? Will each Initiative have its own set of contacts with national ministries or regional entities, or will there be a single CGIAR relationship coordinator for key country and other external partners? How will work priorities be coordinated among cooperating Initiatives?</p> <p>The distinction between focus countries and the four countries selected as technical partners (China, India, Indonesia, South Africa) is not clear. What role is anticipated for the technical partners in scaling foresight approaches?</p>			
<p>9. Research design and proposed implementation demonstrates gender and social inclusion that can be tracked in outcomes</p>	<p>Gender equality, youth & social inclusion 5.3, Projection of benefits 2.7</p>	<p>Legitimacy, Effectiveness</p>	<p>1</p>

The proxy indicator for the group—gender equality, youth, and social inclusion—is number women in new jobs. This is simplistic and inadequate to track the impact of the initiative/foresight planning on youth and other socially disadvantaged groups.

The ways in which the Foresight and Metrics Initiative efforts will address gender equality, youth, and social inclusion are not well specified. Section 5.3 discusses the development of metrics without providing some description of how those will be developed. More importantly, how the Foresight and Metrics Initiative model(s) will provide insights into the determinants of these metrics is not discussed. Presumably new kinds of modeling efforts will be needed to adequately answer questions about gender, youth etc., and it is not clear where and how that will occur.

Civil society (women, youth organizations, representatives of underserved minorities) and private sector representatives are not included in the list of partners who will be consulted 'using participatory and quantitative foresight approaches to analyze how gender equality, youth and social inclusion may evolve....'

10. A risk framework that details main project risks and mitigation actions across areas of science, funding, operations, partnerships, ethics, and environment	Risk assessment 7.3	Credibility, Legitimacy, Relevance	1.7
---	------------------------	---	-----

This is a complex proposal, and the list of risks and mitigation strategies provided here is inadequate. The statement of risks in 7.3 refers almost solely to the risks inherent in working with the six target countries and does not consider any risks for the regional/global impact of the Foresight and Metrics Initiative, or additional, serious risks such as:

The research challenge that the Initiative proposes to address relies on developing/revising foresight models and other analytical tools to identify cross-sectoral synergies and impacts—beyond agriculture—to encompass natural resources, climate change effects, health/nutrition, poverty, and equity. This would appear to require not only a straightforward application of current techniques but some rethinking/revision/coordination to meet these newly identified needs. Is there not some risk of failure here that should be acknowledged by the Initiative?

Regarding the second risk identified (low uptake of scientific innovation of the initiative (Work Package 2)—is it really the case that this could only be jeopardized by the misalignment of the initiative with local policy cycles? As noted above, it is puzzling that the initiative views foresight/modeling as a technical exercise that is the domain of government policy experts and research institutions. To ensure that country models best reflect constraints and meet priority needs at country level, wider engagement of partners—including civil society and private sector—is needed. Not to include them represents a risk that the developed model may be inappropriate or unacceptable to key stakeholder groups. Perhaps this will be the domain of a partner CG initiative—in that case this should be stated clearly.

Regarding the third risk, dependency on legacy data. This risk and its proposed mitigation both seem to be extremely understated. We assume that availability of credible, current data will represent a serious constraint in most countries and regions. While this Initiative cannot tackle the problem alone, if foresight modeling is seen as a critical activity for regions and countries, it may help to energize effort to improve and maintain data collection and access.

Regarding the capability and capacity constraints (fourth and fifth risks), the mitigation plans here seem inadequate. The role of the technical partner countries is not clear. Is it the case that they may provide an ongoing trainer-of-trainers and help desk function after the project ends? Are there other regional technical partners that possess foresight expertise and could be strengthened to provide an ongoing resource, especially in non-focus countries?

12. Capacity building within project teams, partners, and stakeholders captured in capacity development plan. This can include development of early career researchers and partner staff, support/empowerment for under-represented stakeholders, and building partner networks	Capacity development 9.3	Credibility, Legitimacy	1.3
---	--------------------------	-----------------------------------	-----

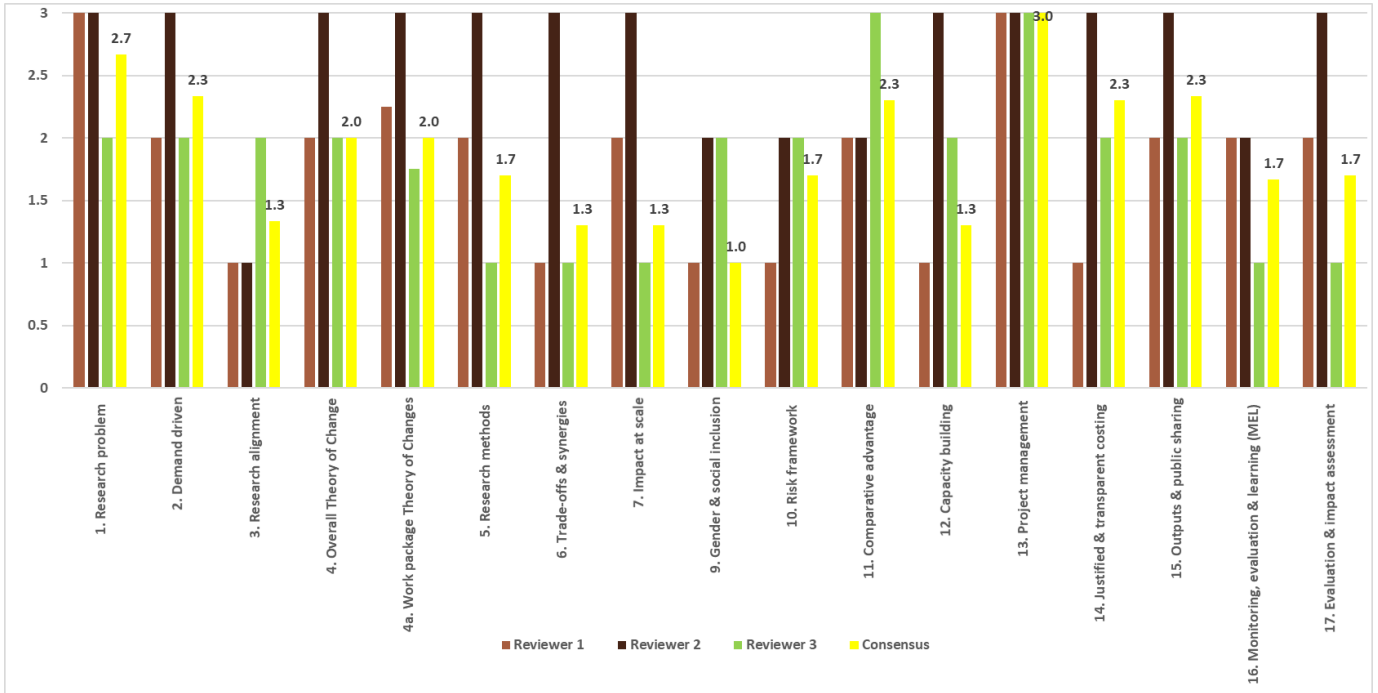
Capacity building outlined for internal CG staff is the focus of 9.3. The capacity development activities described in Work Package 4 for technical partners and focus countries focus on in-person and online training programs and the development of relevant training materials and case studies. These are necessary but fail to convey a dynamic of progressive transfer of leadership and responsibility and resources for foresight planning to national and regional teams.

The plan for strengthening the capacity of junior partners in the focus countries is not well-developed. Work Package 4 (section 3.2) states that needs will be identified, and different kinds of training will be used. Such customization will entail a considerable amount of effort and the capacity of the research team to carry out such customized, intensive efforts is not clear.

Work Package 4's section 6 refers to number of Master and PhD degrees, but there is no clear allocation of funds to support such training. PhD degrees would obviously occur beyond the scope of a three-year project.			
16. Monitoring, evaluation & learning (MEL) plan for the Initiative is clearly defined, with flexibility to adapt. MEL plan supports effective management and learning, including baseline data collection, and evaluative and review processes corresponding to stage-gates and course-correction decisions. MEL occurs during the life of Initiative and is used proactively to reflect on and adapt the Theory of Change, where appropriate	MELIA plan 6.2, Planned MELIA studies and activities 6.3, Measurable three-year (End of Initiative) outcomes 2.2	Credibility, Effectiveness, Legitimacy	1.7
<p>The MELIA plans need rethinking once there is greater clarity about the focus and outcomes of the Foresight and Metrics Initiative. Several proposed activities in the Sections 6.2 and 6.3 are quite unclear and would imply significant new effort. For example, Section 6.2. refers to development of a "problem tree analysis" that will somehow allow identification of cause and effect, as well performance indicators. This seems quite vague, and such analysis might be better used to identify risks rather than performance. Another example in section 6.3 is the use of a "model-based evaluation" to understand how policies are informed by "our innovations." It is unclear what this means? How could the Foresight and Metrics Initiative be used to identify how it informs policies? As these two examples show, there are a wide variety of approaches mentioned, and this lack of focus seems to indicate a solid MELIA plan is lacking.</p> <p>In the tables reported in section 6.1, the current reference values and expected targets in 2025 are never indicated for outputs and outcomes. Having these values established would be very useful to understand whether the target values are achievable in the execution phase, and to provide a basis for evaluating the impact of the initiative in the assessment stage.</p> <p>Considerable weight is given to uptake of the Initiative's innovations by partners. But the majority of effort and outputs focus on the technical work of foresight models and tools, and development of learning materials, with relatively less attention to the process of engaging partners and stakeholders to participate in model/scenario development, and for technical partners to gain capacity and assume responsibility for foresight analyses.</p> <p>Ongoing use of foresight tools depends on capacity but also on agency of partners/stakeholders. What are benchmarks that define different levels of capacity improvement, and, similarly, different levels of perceived agency/ownership of foresight tools and processes, by different stakeholders?</p>			
17. Well-defined plan for Initiative-level evaluation and impact assessment based on expected end-of-Initiative outcomes and impact. Links between the impact assessment plan and indicators in the Theory of Change are clear	MELIA plan 6.2, Planned MELIA studies and activities 6.3, Full Initiative ToC 3.1, Work Package ToCs 3.2, Projection of benefits 2.7	Effectiveness, Relevance	1.7
See comments under 16.			

The review is continued on next page.

The Figure below represents original scoring from each reviewer and consensus scoring for each criterion. *The consensus score across reviewers may not reflect the mathematical average.* For purposes of the Figure, the QoR4D criteria have been shortened. Please refer to p. 3 for full criteria definitions. Please note that criterion 8 on ethics was not scored.



*Two consensus scores were greater than 0.5 variance from the mathematical average. Please refer to criteria 7 and 12 above for rationale of these consensus scores (both with a variance of 0.7 from the mathematical average).

3. From Fragility to Resilience in Central and West Asia and North Africa (F2R-CWANA)

Review Summary and Actionable Recommendation

This Initiative addresses a very complex, relevant set of challenges in a broadly convincing way. Many of the review criteria have been addressed explicitly and with good intent, but overall, the approach is not fully persuasive and, as indicated below, significant elements lack clarity. Central and West Asia and North Africa (CWANA) obviously is a region in which prospects for profound changes in climate (which already are apparent) will almost certainly exacerbate existing problems of food insecurity, conflict, and water scarcity; tackling these challenges is a tall order. Despite the pandemic, the Initiative Design Team engaged in an impressive amount of consultation during the process to create this proposal, which capitalizes on many of the CGIAR's established strengths within the region and also globally. However, the Initiative may be overly ambitious given the limited timeframe and resources combined with very real risks of disruption by events outside researchers' control, which they acknowledge.

The Initiative combines elements of two CGIAR Action Areas: Resilient Agrifood Systems and Genetic Innovation. Linking these could have advantages, but integration and articulation necessary to achieve those synergies are not always sufficiently clear. The absence of a clear, concrete agrifood systems framing and rationale in terms of "integrated systems approaches" (or even citations to relevant scientific publications) raises some fundamental concerns regarding understanding and capacity to implement a "Resilient Agrifood Systems" Initiative. Moreover, lack of a systems framework raises questions about the strategic nature of this work and potential for generating regional and international public goods. **Although necessary detail for full evaluation is lacking, this appears to be very much a "business as usual" framing for the Genetic Innovation aspects of the Initiative and lessons from past experience do not seem to be incorporated.** A very different and fresher approach would seem appropriate for a new research-for-development strategy. The review team therefore wonders if the right capabilities, particularly regarding social sciences, are in place to address the key challenges of this region. There is also a missed opportunity to engage with Agricultural Research Institutes based in similar climatic regions via genuine partnerships, including sharing of real resources.

Finally, one wonders how this proposal might be different (and more innovative) if the "humanitarian-development-peace nexus" approach (mentioned in the risk assessment) were embraced as an Initiative design challenge, not merely as mitigating funding uncertainty? A stronger and more explicit focus on rural poverty reduction rather than just technological solutions to productivity might have assisted in developing the missing integrated systems approach. A further concern is that the Initiative is not written using a concise and convincing style. It is full of generalities, jargon, and excessively long sentences that impede comprehension. This does not make for compelling reading, particularly for non-experts who will make major funding decisions based on this proposal. This constitutes a reputational risk for CGIAR. Hence, the ISDC through this review, recommends the three specific weaknesses outlined below are addressed through a substantive rewrite.

Overall Strengths of Proposal

Evidence on extent of consultations by the Initiative in validating issues, strategies, and activities (see criterion 2 below)

Project management (see criterion 13 below)

Monitoring, evaluation, and learning (MEL) plan (see criterion 16 below)

Overall Weaknesses of Proposal

Characterization of research problem (see criterion 1 below for more detail). While the challenges mentioned are compelling, there is a lack of hard evidence and scientific documentation of the scope of these problems and the corresponding prospects for the Initiative to address them at significant scale. The language is often vague and full of empty jargon. While there is recognition of the major issues impeding rural development in this region, including political instability and institutional rigidity, the research strategy does not seem to be attuned to these realities. The concerns raised in the summary above regarding the poor articulation and presentation of arguments is particularly evident for this criterion. All the buzzwords are there, but it is impossible to understand the research imperative. **To address this issue will require a reframing of the proposal as part of a substantive rewrite.**

Analysis of trade-offs across Impact Areas (see criterion 6 below for more detail). Neither discussion nor analysis of trade-offs could be found. Similarly, credible *ex-ante* assessment necessary to understand regional scope for impacts is absent. Hence, prioritization as a basis for strategic choices is impossible. It is therefore difficult to know which Impact Areas will be addressed and how impediments to development would be overcome.

For instance, what trade-offs might be involved between environmental health and biodiversity versus poverty reduction, livelihoods, and gender equality?

Overall theory of change (see criterion 4 for more detail), especially lack of clarity on causal linkages and roles of partners in delivering outcomes and impacts and superficial articulation of underlying assumptions. While text in Section 3.1 lists Work Packages, activities, and aspirational outcomes and impact, it lacks both logical causal connections and critical presentation of assumptions and risks; it also is vague on partners' roles in driving outcomes and impacts and on the role of ARI's and global science in advancing knowledge and action. The underlying "theory" would seem to be a technology transfer model, with research activities aimed at filling unspecified knowledge gaps and without serious attention to the full range of factors underlying yield gaps and barriers to adoption. **Little insight is provided into how people will be helped by these interventions, why these innovations have not happened already, and if there are barriers, what can researchers do to surmount them?**

Areas of Divergence among Review Team and ISDC Resolution

Criterion 7. Evidence for impact at scale through integrated systems approaches. Split scores 3, 3, and 1. Rationale for the score of 1 by the SME coordinator: Absence of a clear, concrete agrifood systems framing and rationale in terms of "integrated systems approach" (or even citations to relevant scientific publications) raises fundamental questions regarding understanding and capacity to implement a "Resilient Agrifood Systems" Initiative. Moreover, lack of a systems framework raises questions about the strategic nature of this work and potential for developing International Public Goods. However, there seems to be good articulation between Work Package 4 and NEXUS Gains— indeed one wonders why this Work Package 4 is in a separate proposal rather than incorporated within NEXUS Gains? Work Package 5 is innovative (for CGIAR), but even here the pursuit of innovations across different elements of the food system is not the same as an agri-food systems approach. The Initiative does not address the fact that benefits don't always scale. For instance, while higher yields might be regarded as a policy and public good success at the macro level, they often provided little or no benefits for farmer (micro level). This highlights the importance of foresight and trade-off analyses and attention to rural poverty reduction, which are aspects that are largely absent from this Initiative.

The ISDC review team lead agrees with the consensus score of 1 for this criterion.

Does the Initiative Align with the Cohesion of the Portfolio as Described in the Companion Document?

At one level, F2R CWANA's Work Package-level theory of change diagrams are quite effective in depicting how elements of this Initiative interact with other CGIAR Initiatives. On the other hand, like so much of this proposal, the text is not really clear about how these are articulated with the broader CGIAR portfolio. More important than all those arrows on the diagrams, and as emphasized in the commentary on specific review criteria, the absence of a coherent theory of change for this Initiative. This raises questions about synergistic links with the entire One CGIAR portfolio. One can imagine how some of the relevant Genetic Innovation Initiatives might contribute germplasm to F2R CWANA (though this only is touched upon in the proposal and without specifics), but it is difficult to see the potential for flows in the opposite direction or even potential for the generation of regional or international public goods. One exception here might be the interactions described with NEXUS Gains, though that seems confined to Work Package 4. The Companion Document stresses (on p. 12) the "Impact Area Platforms" in generating "impact-oriented cohesion" (which would be great!), but the very weak treatment of trade-offs/synergies across Impact Areas and absence of ex-ante assessment of scaling of benefits to impacts (discussed under criterion 6) gives neither a logical rationale nor a basis for confidence.

The reviewers scored each of the 17 QoR4D criterion individually. The review teams then built a consensus score. In the majority of reviews, the consensus score was the same as the mathematical average of the three individual reviewer scores. The criteria that received between a 0 or 1 are presented with a rationale below. Although not required, in order to keep review reports brief and readable, some teams opted to record in their reports their consensus-built rationale for scores above 1. Following the criteria is a figure presenting the individual reviewer scores and consensus. Please refer to p. 5 for the Likert scoring definitions.

Criteria	Proposal Sections	QoR4D Elements	Consensus Score
1. Clearly defined research problem that addresses Impact Areas, is a high priority in the targeted geographies, is well aligned to shared, multi-funder priorities, and is well informed by previous research findings and evaluations	Challenge statement 2.1, Learning from prior evaluations and Impact Assessments 2.3,	Relevance, Effectiveness	1

Criteria	Proposal Sections	QoR4D Elements	Consensus Score
	Impact statements 5		
<p>There is no doubt that climate change will present considerable problems for agriculture in CWANA. However, the challenge statement cites no supporting contextual evidence apart from the IPCC report and does not elaborate on the nature of past and expected future changes in the climate. Similarly, it is undeniable that climate instability and other drivers of change in CWANA seem poised to exacerbate food import dependence, national food insecurity, and water supply conflicts, or that massive out-migration (especially by young males) leaves many women in charge of agriculture and other rural activities. And yet, the daunting challenges underpinning F2R-CWANA are neither clearly defined nor supported by quantitative evidence or sufficient citations. This is a huge and diverse region (really a number of regions): What portion of its economic (income, employment), social (poverty, gender, youth, refugees), food security, water scarcity, and agrobiodiversity challenges could be addressed by the F2R CWANA Initiative? A simple table of indicators for the region and target countries would be very helpful in assessing this proposal. Furthermore, while it is true that field crop production is important in agriculture across CWANA, and of course closely related to livestock production, the fact that fruit and vegetable production is ignored needs some explanation. The "End of Initiative Outcomes" also are clear, but under "technology" only genetic innovations are addressed, with no mention of mechanical, IT, or other technologies. Overall, there is a bias towards addressing on-farm problems and issues, while the Initiative title, problem statement, and Work Package theories of change highlight the agrifood system as a whole. Furthermore, the impact statements are difficult to assess because almost none of them have three-year goals associated with them. While needs for capacity building to address myriad challenges are clear, discussions of capacity building activities could be improved.</p> <p>This proposal does not appear to reflect lessons from past AR4D in the region. Section 2.3 ("Learning from prior evaluations ...") comes across as superficial and selective. Where is the candid discussion and reflection on practical lessons from earlier efforts, not just Grain Legumes and Dryland Cereals (GLDC), but also the mixed record of prior Consortium Research Program (CRP) predecessors: Dryland Systems, Dryland Cereals, and Grain Legumes, <u>particularly the importance of sound priority setting, including demand analysis, and integration of breeding (Work Package 2) with systems agronomy (Work Package 3)</u>? Nonetheless, we do sympathize with scientists facing the daunting risks and challenges of conducting AR4D research in CWANA. It is intriguing to think of a possible approach to project design that fully embraces an operational, systems science framing of the "humanitarian-development-peace nexus approach," (which is only mentioned in passing on p. 71 as a mitigating aspect of fundraising risk). This may be a missed opportunity for a truly innovative framing for a CGIAR initiative spanning CWANA that could have great relevance for this region and for other regions going forward in an increasingly uncertain/disrupted world. And, without considering the implications of this "nexus approach," one worries about the risks entailed with a largely conventional approach to AR4D in this region.</p>			
<p>2. Evidence that the Initiative is demand driven through codesign with key stakeholders and partners (Investment Advisory Groups, governments, private sector, funders) and research collaborators within and outside CGIAR</p>	<p>Participatory design process 2.6, Challenge statement 2.1, Work Package ToCs 3.2</p>	<p>Relevance, Effectiveness</p>	<p>3</p>
<p>There is strong evidence of consultation, including reports for almost all the countries involved, as well as regional surveys. The proposal also documents a creative approach to participatory design, despite challenges posed by the COVID-19 pandemic, engaging an impressive combination of stakeholders from different entities and sectors. However, within the CGIAR, commitments of various centers could be clearer. Finally, and most fundamentally, who among these high-level partners is in a position to speak to the AR4D priorities of the target beneficiary groups? Do rural women and youth want these innovations? Were these groups consulted directly in the design of this proposal? How? If not, will their input be sought early in the process? How?</p>			
<p>3. Research questions, objectives, outputs, and outcomes are aligned to the research problem, and are measurable with defined deliverables</p>	<p>Work Package ToCs 3.2, Measurable three-year (End of Initiative) outcomes 2.2, Priority-setting 2.4, Management plan 7.1</p>	<p>Relevance, Effectiveness</p>	<p>2</p>

Criteria	Proposal Sections	QoR4D Elements	Consensus Score
<p>The three-year outcomes could be clarified and may not be measurable in their current form. However, within the Work Packages, the outcome definitions are better and will be easier to quantify; deliverables are well-defined. Characterization of problems appears to be included as a first step in many Work Packages and is fundamental to Work Package 1. Overall, however, scientific grounding for the research questions is not clear nor are sufficient scientific citations provided to document and justify the approaches proposed. The narrative relies heavily on (undefined) jargon, e.g., “sustainable intensification, conservation agriculture, scale appropriate mechanization, nexus governance, inclusive entrepreneurship, national innovation platforms,” to name but a few. A large literature is relevant to each of these, but that prior science is not critically assessed nor are the most important publications cited. Many of these terms—for example “sustainable intensification” versus “conservation agriculture”—need to be critically assessed regarding their practical implications, strengths, and weakness (which are contested in the literature and amongst research and development agencies on the ground) as concepts underpinning this Initiative. Even on the most practical level, there are problems of clarity. What are the “Best Bet Genetic Innovations”? What crops and livestock are of central interest? What farming systems and fisheries? Have these already been selected? If so, how? If not, how will they be prioritized? It is not even clear whether the proposal focuses on rainfed, irrigated, or both types of systems and in which countries? Specifically, in section 4.1, Innovation Packages and Scaling Readiness Plan (p. 35): “The F2R-CWANA Initiative will evaluate and scale a range of proven and well characterized set of Core Innovations generated through CRPs and bilateral projects.” The annexes on “harvesting Golden Eggs” and “Core Innovations” (linked from p. 35) mainly list general research methods and data sources; when specific crops and livestock/fish are mentioned, their relevance to CWANA is not always clear or, in cases where the studies of Core Innovations focus on CWANA, the information provided in the table is too general to evaluate (viz., “new food and feed crops, scale appropriate mechanization platform”).</p>			
<p>4. Overall Theory of Change with intended outputs, outcomes, and impacts at scale clearly described. Assumptions are documented, causal linkages are clear, especially the role of partners in driving impact</p>	<p>Full Initiative ToC 3.1</p>	<p>Effectiveness, Relevance</p>	<p>1</p>
<p>Rationale: The graphic for the full Theory of Change effectively clarifies how the different elements of the Initiative fit together. Arguments and explanations regarding intended outputs, outcomes, and impacts at scale are convincing, but what is the overarching “<u>Theory of Change</u>?” Text in Section 3.1 lists Work Packages, activities, and aspirational outcomes and impact, but it lacks both logical causal connections and critical presentation of assumptions and risks; it also is vague on partners’ roles in driving outcomes and impacts. The underlying “theory” would seem to be a technology transfer model, with research activities aimed at filling unspecified knowledge gaps and without serious attention to the full range of factors underlying yield gaps and barriers to adoption. Little insight is provided into how people will be helped by these interventions, why these innovations have not happened already, and if there are barriers, what can researchers do to surmount them?</p>			
<p>4.a Individual Work Package Theory of Change (score individually)</p> <p>Work Package 1 Work Package 2 Work Package 3 Work Package 4 Work Package 5</p>	<p>Work Package ToCs 3.2</p>	<p>Effectiveness, Relevance</p>	<p>WP1: 2 WP2: 2 WP3: 2 WP4: 3 WP5: 3</p>
<p>Work Package 1: Is this Work Package primarily a channel for extension of innovations from the other Work Packages? The Work Package 1 Theory of Change highlights inflows of “innovations to be scaled” from the other Work Packages, but certainly results from WP1 also are relevant and important for the other Work Packages? This seems to be related to RQ1.02 and RQ1.04, but the methods do not seem appropriate to answering these questions or producing key outputs. Specifically, what is the mechanism for feedback from “national innovation platforms” and Work Package 1 research results into research activities in the other Work Packages? It would be helpful to have more practical clarity on the operational meaning of “national innovation platform,” who will be involved, and why would they choose to participate? If there are policy differences and powerful competing interests vested in the status quo, how will institutional and policy issues that are barriers to adoption be addressed? Which agro-ecologies will have physical “living labs?” How are those agro-ecologies prioritized? What will be their scope, coverage, and representation of important domains within the CWANA region? The absence of any mention of the following issues is an omission that pervades the entire report: insurance, food storage capacity, and agricultural labor productivity (i.e., worker safety in extremely hot environments has been shown to be helped by simple interventions, such as having shade to rest in). Many of these are systemic issues that could be addressed in Work Package 1</p>			

Criteria	Proposal Sections	QoR4D Elements	Consensus Score
<p>Work Package 2: Three big unaddressed questions here are (1) whether or not these innovations are profitable (i.e., does the additional financial benefit from these improved varieties exceed the cost of using them in terms of labor and inputs?), (2) are relative risks and profitability attractive compared to alternatives, including the existing cultivars? and (3) what are the barriers to adoption? <u>These should be research questions, not just assumptions in which these questions are (at best) alluded to indirectly.</u> We also were not able to find sufficient specifics on WP2 genetic innovation priorities. Although “Best Bet Genetic Innovations” are mentioned, these are unspecified. What crops/livestock/fish will be included or are priority candidates? How will those be selected? By whom and using what criteria for prioritization and for balancing tradeoffs across attributes? How will the needs of women and youth be identified and factored into breeding priorities? It is indicated that women and other stakeholders will participate in on-station and on-farm assessment and validation, but who sets the primary priorities for crops and systems? How will agrobiodiversity conservation (RQ 2.03) be incorporated into breeding (RQ 2.01)? In that same vein, how will WP 3 (farming systems/agronomy) insights shape activities in WP 2 (breeding) and how will breeding programs draw on WP 4 and WP 5 to address water and climate issues in order to build resilience through adaptation of breeding priorities and even changing crop, livestock and fisheries product priorities? All these questions remain unanswered, which highlights the general weakness of this initiative: excessive vagueness and a lack of rigorous, scientific problem definition. Although there may be some two-way flow between WP 2 and both WP1 and WP5, apparently there are no flows with either WP 3 or WP 4, and this articulation and integration would seem crucial to a successful breeding program aimed at resilience and adaptation.</p> <p>Work Package 3: In parallel with Work Package 2, three fundamental questions are unaddressed in Work Package 3: (1) whether or not these agronomic innovations are profitable (i.e., does the additional benefit from these innovations and practices exceed the cost of using them in terms of labor and inputs?), (2) are relative risks and profitability attractive compared to alternatives, including the existing farming/aquaculture systems?, and (3) what are the barriers to adoption of these innovations and practices? It is not even clear what systems and innovations will be the focus in each country. Specifically, what are the “principal farming systems” mentioned in the proposal? Moreover, while the proposal mentions that “the region’s insufficient knowledge and service delivery infrastructure, capacity, and enabling policies are key bottlenecks,” apart from RQ3.04, research is focused elsewhere (on agronomic and genetic components of yield gaps) and even RQ 3.04 is vague on how those infrastructure, institutional, and policy issues will be addressed. Overall, as already emphasized, the six assumptions listed (e.g., “farmers ... see the value in this conservation and climate-smart agriculture approach ... over other alternatives,” “strategies and business models are workable and profitable”) really are research questions to be assessed <i>ex ante</i> and fed into breeding and farming systems agronomy programs. Finally, how will Work Package 3 interact with Work Package 2? The Theory of Change indicates “release of Global Genetic Innovations” flowing from Work Package 2 into Work Package 3, but how do agronomic insights from WP3 flow into Work Package 2 and shape genetic improvement strategies? More fundamentally, why are the Work Package 2 and 3 activities separated across work packages rather than being more closely integrated?</p> <p>Work Package 4: This is the clearest and most compelling set of research questions among these WPs, though the scientific methods are simply listed and not spelled out with sufficient specificity. There seems to be good integration with WP4 across WP 1, 3, and 5 (as well as NEXUS Gains and other CGIAR initiatives), but notably not to WP 2, which is a major concern. It is hard to assess which partner(s) will provide the distinctive skill sets required for this WP and disappointing that there is no capacity development impact pathway in this WP, given the importance of challenges arising from the water-climate-food nexus in CWANA. Furthermore, many issues related to water involve multiple stakeholders with competing interests regarding this increasingly scarce resource: expertise in conflict resolution, political science, public administration, and water law would seem to be essential, but are not mentioned in the proposal.</p> <p>Work Package 5: “Food value chains” (FVCs) are a central concept of this WP (and the Initiative); thus, a clear operational definition of what is meant by FVC (and the scope of that central concept) is required either here or in the overall problem statement, along with supporting references to the relevant literature. WP 5 holds significant potential for innovative action research, including experimentation with models for public/private investment opportunities, perhaps including development of various insurance products to mitigate risks for various FVC actors and entities. Some of the research questions seem to be very context-specific; how will answering these questions advance knowledge more generally and provide regional and international public goods? Presumably, non-CGIAR partners play key roles in activities for RQ 5.03 and RQ 5.04 (which don’t read like research questions), but the roles and mechanisms and levels of funding of these essential partners - and especially “selected incubator/accelerator programs” - are not sufficiently clear within the overall design of the Initiative.</p>			

Criteria	Proposal Sections	QoR4D Elements	Consensus Score
5. Research methodology and methods (and supporting activities) are fit-for-purpose, feasible, and assumptions and risks are clearly stated	Work Package ToCs 3.2, Priority-setting 2.4, Innovation Packages and Scaling Readiness Plan 4.1	Credibility, Relevance, Effectiveness	2
<p>The methods laid out in the Work Packages are directly related to the research questions proposed. The stated methodologies are clear and reasonable, with the broad exception (note above in 4.a regarding Work Package 2 and Work Package 3): if one does not understand whether or not these varieties, crops, technologies, and practices are profitable at an acceptable level of risk, then scope for adoption and impact is questionable. Profitability/adoptability is sometimes assumed, rather than being something that should be tested. Crucially, higher yields do <u>not</u> always translate to net positive profits at acceptable levels of risk for farmers (under increasingly risky conditions). More generally, key “assumptions” are superficial (too often taking the form “innovations will be adopted”) and beg fundamental questions of farmer engagement, prioritization, fit, appropriateness, and responsiveness to farmers’ needs and capabilities, especially for women and youth. Purposes of some research methods can be somewhat vague (e.g., under Q4.05 “Socio-economic data analysis”—To what end? Amongst whom? Where obtained? What type?). Finally, as mentioned above in Criterion 3, scientific grounding is not clear nor are sufficient scientific citations provided to document and justify the methods and approaches proposed. Regarding the components that relate to Resilient Agrifood Systems—the proposal lacks a coherent, scientific grounding in systems science and the nature and challenges of resilience within an agri-food systems/agricultural development context in middle income countries. Important terms are not defined operationally within the relevant scientific frameworks and concepts.</p>			
6. Analysis of trade-offs and synergies across the CGIAR Impact Areas; ex-ante assessment of project benefits provides logical rationale for scaling of impacts	Projection of benefits 2.7, Result framework 6.1, Impact statements 5, Innovation Packages and Scaling Readiness Plan 4.1	Effectiveness, Credibility	1
<p>The results framework is highly effective in clarifying what is going to be done and how it will be measured. The proposal seems cognizant of the broad set of complementary efforts across the CGIAR. However, the projections of benefits seem quite optimistic and there is <u>no</u> apparent effort to analyze trade-offs across CGIAR Impact Areas. From section 4.1, we infer the intention is to take existing “Core Innovations” that have been generated by other projects, use WP1 to assess readiness to adopt under all five Work Packages, then develop implementation strategies for scaling up in collaboration with partners. The question remains whether the end results will have significant impacts at the level of CGIAR impact areas. The impact statements (Section 5) are vague, lack supporting empirical evidence, and do not provide critical assessment of existing published research. The absence of scientific citations is striking. Taking each of the five CGIAR Impact Areas in turn:</p> <p><u>Nutrition, health, food security</u>: impossible to assess based on the sketchy analysis and lack of evidence in this proposal. One does not even learn what nutrition-dense crops (or relevant farming systems) will be included in the research. We found no mention of horticultural crops, which in addition to nutritional values, also could provide particularly important income opportunities for women and youth and innovation opportunities for Work Package 5.</p> <p><u>Poverty reduction, livelihoods, and jobs</u>: arguments are not convincing. Indeed, the track record in agricultural development as an engine of poverty reduction is not encouraging for many of these countries, most of which are quite far along with structural transformation. This is recognized in the proposal on p. 37: “Studies have confirmed high poverty and unemployment levels regionally,” [it would have been helpful to see these data or other indicators on these issues] “but few case studies demonstrated that agricultural innovations can contribute to improve livelihood security, raise incomes and reduce poverty in efficient and inclusive ways.”</p> <p><u>Gender equality, youth, and social inclusion</u>: entrenched challenges are acknowledged generally, but historic, structural barriers are not addressed in the proposal for insights on what really is feasible within the timeframe of the Initiative.</p> <p><u>Climate adaptation and mitigation</u>: strangely absent in Work Package 1; simplistic in Work Package 2 and 3; strongest in Work Package 4; and implicit in Work Package 5 and could be presented in much more compelling ways in each, for example by placing geographic and material scope for impact within the context of sub-regions in this large, diverse region.</p>			

Criteria	Proposal Sections	QoR4D Elements	Consensus Score
<u>Environmental health and biodiversity</u> : biodiversity impacts are neither clearly specified nor soundly articulated, particularly as these may involve tradeoffs with other impact areas, such as food security and poverty reduction.			
7. Evidence that the Initiative will likely lead to impact at scale through integrated systems approaches that drive innovation in research and partnerships, including linking to and leveraging of other Initiatives within and outside CGIAR	Projection of benefits 2.7, Work Package research plans and ToCs 3.2	Effectiveness, Credibility, Relevance,	1
No consensus among reviewers, with two scoring this criterion as a 3 and one as a 1. See "Areas of Divergence among Review Team" above for rationale for score of 1. This final score is supported by ISDC.			
9. Research design and proposed implementation demonstrates gender and social inclusion that can be tracked in outcomes	Gender equality, youth & social inclusion 5.3, Projection of benefits 2.7	Legitimacy, Effectiveness	2
Rhetorically, this proposal places great emphasis on women and youth as intended beneficiaries/adopters, populations that would be very challenging almost anywhere, and especially so in this region. While the intentions are good, commitment of necessary resources to address these huge challenges is not clear in either section 9.3 (Capacity Development) nor in section 5.3 (Gender equality, youth, and social inclusion); both sections are vague and lack a vision for innovative approaches that actually might enable tangible progress in a foreseeable timeframe. Such tangible progress is necessary for success of all Work Packages and deserves greater attention in project design. One wonders what progress on gender and social inclusion really can be made in the timeframe for End of Initiative outcomes? At a very basic level, it is not even clear how women, youth, and other marginalized groups (including refugees) have been (or will be) consulted directly regarding their needs, capacities, and constraints. We recognize that the pandemic and other challenges make consultations difficult, but key questions remain: how can we know these groups want the innovations envisioned in this Initiative and that those technologies and practices will be feasible for them to adopt?			
10. A risk framework that details main project risks and mitigation actions across areas of science, funding, operations, partnerships, ethics, and environment	Risk assessment 7.3	Credibility, Legitimacy, Relevance	2
No doubt, this is a very high-risk region regarding a range of threats (political, military, economic, socio-cultural, climatic, and environmental). The identification and prioritization of risks in Section 7.3 seems appropriate and comprehensive in addressing <u>risks to the project</u> . Particular kudos for candor about the risks of conflict in the region. However, the "substitute countries" include Afghanistan, Syria, and Yemen, which present similar (or currently worse?) security risks than Sudan. The inevitably subjective rankings of likelihoods and impacts (and consequently the risk scores) seem rather conservative (too low) and mitigation measures (especially in the last row on tensions between public and private sector) seem superficial. Although it is reported that "unintended consequences of technologies/innovations" were considered, there is no discussion of these <u>risks from project activities</u> .			
11. CGIAR capacity and its comparative advantage and appropriateness to lead the work is justified. This includes the skills, diversity and multi-/trans-disciplinarity of the research team and approaches to meeting gender and diversity targets	Comparative advantage 2.5, Initiative team 9.1, Gender, diversity and inclusion in the workplace 9.2	Relevance, Legitimacy, Effectiveness	2
Conflict and policy fragmentation are emphasized as two of the forces to reckon with in what the proposal refers to as the "perfect storm" facing CWANA. Who on the core team has the training to understand these issues fully? While clearly indispensable to Work Package 1 and 4 (and arguably necessary for all Work Packages), political science, public administration, legal, mediation, and conflict resolution expertise does not seem to be included in section 5.1 (or elsewhere). There is well-established CGIAR capacity across this region related to Work Package 2 and 3; however, the risk here is continuation of business as usual (see comments above in point 1 regarding lack of apparent response to lessons from previous experience in CGIAR plant breeding and farming systems work in the region), particularly if it proves impossible to engage sufficient numbers of women professionals and extension agents (as acknowledged by the team). Work Package 4 (and perhaps also Work Package 1) can benefit from links via IWMI leadership in both this proposal and NEXUS Gains. Work Package 5 is NOT an area in which CGIAR has demonstrated comparative advantage, but key partners (e.g., MENA WEF Hub, Berytech) do appear to have expertise and it is encouraging that Berytech leadership is represented in the core IDT group.			

Criteria	Proposal Sections	QoR4D Elements	Consensus Score
However, there is not sufficient detail in the budget to assess how/whether resources will be shared beyond CGIAR partners, who dominate most aspects of this proposal.			
12. Capacity building within project teams, partners, and stakeholders captured in capacity development plan. This can include development of early career researchers and partner staff, support/empowerment for under-represented stakeholders, and building partner networks	Capacity development 9.3	Credibility, Legitimacy	2
Capacity building is mentioned frequently but is not developed sufficiently anywhere in the proposal. Specifics in the proposal are largely focused on internal capacity building for team members; it could be helpful to have greater clarity on mentoring and on whether capacity building should include technical as well as managerial skills. Plans for capacity building for stakeholders (especially women, youth, and marginalized groups including refugees) are vague and it is not clear sufficient resources (skilled staff and budgets) are included in plans. In CWANA, capacity development for these stakeholder/beneficiary groups requires a major and innovative effort. Building capacities for these groups to participate in, contribute to, and benefit from this initiative will be essential to realizing End of Initiative outcomes and CGIAR Action Area impacts. On the other hand, there is laudable candor about difficulty in attaining gender CGIAR goals in project leadership (Section 9.2, p. 75) and it is worth noting that a woman from the region is designated as co-lead. This candid, practical approach needs to be extended to design of capacity building for stakeholders more broadly.			
13. Project management mechanisms and (if applicable) additional scientific oversight and governance measures effectively and efficiently support the Initiative objectives	Management plan and Risk assessment 7, Research governance 8.1	Legitimacy, Credibility	3
Thorough and complete, including plans for a project steering committee comprising stakeholders, which is essential to the legitimacy of the initiative and to an adaptive approach to project management.			
14. Justified and transparent costing explicitly linked to expected Research for Development results	Financial Resources 10	Legitimacy, Effectiveness	3
Overall, the balance in budgeting across Work Packages seems appropriate. However, an impediment to our review is that there is no detail on budgeting for work by specific partners nor information on funding commitments for specific areas of expertise/skill sets. Also, it seems strange that all the budget is allocated for activities within specific countries: where is the funding (which should be significant) for cross-country synthesis and sharing of regional and international public goods?			
15. Anticipated research outputs (knowledge, technical, or institutional advances, specific technologies or products, policy analyses) are described and knowledge/gaps they will fill are evident. Protocols for open-data and open-access compliance are evident in plan (including budget)	Work Package research plans and ToCs 3.2, Open and FAIR data assets 8.2	Credibility, Effectiveness	3
These are well-described and appropriate. Protocols for open data and open access seem thorough and align with CGIAR practices, which are improving. Overall, however, the flavor of the Initiative is closer to a development project rather than an AR4D program and, apart from elements of Work Package 4, it is not always clear how these activities will advance knowledge or produce international public goods.			
16. Monitoring, evaluation & learning (MEL) plan for the Initiative is clearly defined, with flexibility to adapt. MEL plan supports effective management and learning, including baseline data collection, and evaluative and review processes corresponding to stage-gates and course-correction decisions. MEL occurs during the life of Initiative and is used proactively to reflect on and adapt the Theory of Change, where appropriate	MELIA plan 6.2, Planned MELIA studies and activities 6.3, Measurable three-year (End of Initiative) outcomes 2.2	Credibility, Effectiveness, Legitimacy	3
The MEL plan is clear and seems comprehensive, with explicit links to End of Initiative outcomes. What is specified in the plan is superior to what is discussed in the impacts section earlier on in the proposal. There are two areas of some concern, however. First, there appear to be gaps in the "3-year targets and metrics" columns in Section 5 (impact statements). Why, and how will progress be measured in the absence of such targets and metrics? Another concern: it is a surprise to read (p. 65) that "The Initiative may facilitate an independent external review ..." Is it not a best practice to plan (and budget) for an independent external review?			

Criteria	Proposal Sections	QoR4D Elements	Consensus Score
17. Well-defined plan for Initiative-level evaluation and impact assessment based on expected end-of-Initiative outcomes and impact. Links between the impact assessment plan and indicators in the Theory of Change are clear	MELIA plan 6.2, Planned MELIA studies and activities 6.3, Full Initiative ToC 3.1, Work Package ToCs 3.2, Projection of benefits 2.7	Effectiveness, Relevance	3
Overall, the IA plan is clear, and links are appropriately made to the Theory of Change. While it is indicated (p. 65) that “Researchers will design causal impact assessments ...,” it is not clear who will do that—will they be independent specialists in quantitative IA methods? If not, and design and implementation of impact assessments are done by researchers directly involved in the project activities, this creates a conflict of interest and credibility of such studies would be compromised. And, while it is good that <i>ex-ante</i> baseline studies (disaggregated by gender) are planned as indicated in section 6.3 (pp. 66-67), the bulk of MELIA activities appear to focus on adoption of innovations and are not traced through to Projection of Benefits (Section 2.7) or linked to Action Area outcomes (pp. 41-43) or higher-level impact indicators, providing little or no basis for assessment of impact at the level of CGIAR Action Areas. An additional fundamental concern here is that there is no mention of compilation or creation of baselines in CWANA for indicators of results regarding CGIAR Action Areas.			

Additional Comments Not Presented Above

Criterion 4.a. (Work Package 1). Research methods under RQ1.03 are incorrect—they do not describe methods.

Criterion 4.a. (Work Package 2). It is surprising to read that the proposal assumes “agrobiodiversity hotspots exist, and can be identified, reached and accessions collected” ... hasn’t ICARDA been conducting agrobiodiversity exploration and collection in CWANA for at least 40-50 years? If these are not well-characterized already, why not? And what are the prospects for filling those gaps given high current levels of instability in CWANA?

Criterion 4.a. (Work Package 5). What will success look like for what typically are high-risk ventures, with only a few innovations likely to yield commercially viable returns? What is an optimum success rate for these ventures?

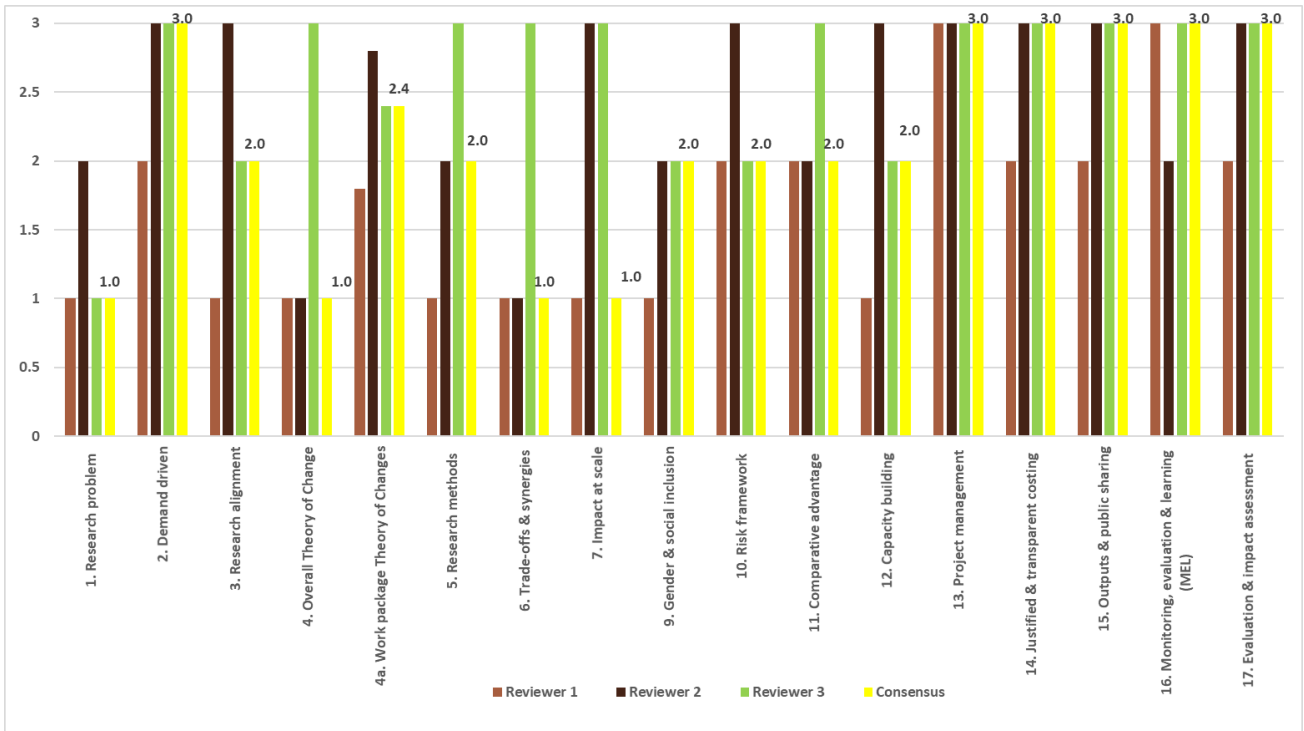
Criterion 5. (Research methodology and methods). It is not clear that our independent review team has the full skill set needed to evaluate what is in large part a Genetic Innovation project. And yet, there is so little information provided on breeding methods (and reference to supporting literature) that one also may wonder whether this could be evaluated by someone with cutting-edge expertise in breeding/genetics?

Criterion 9. (Gender and social inclusion). What does it mean to “empower refugees” and how would such efforts be viewed by national partner institutions?

Criterion 11. (CGIAR comparative advantage). Although it may have been an effort to signal a strong embrace of the “One CGIAR” concept, it was not helpful to list affiliations as simply “CGIAR” rather than indicating individual center affiliations of Initiative Design Team members and other personnel.

Criterion 14. Costing/budgets. Problems of specificity and transparency in the budget presented here are shared by other CGIAR proposals in this cohort, stemming in large part from the budget template structure required by CGIAR, which does not provide sufficient scope for detail necessary for reviewers.

The Figure below represents original scoring from each reviewer and consensus scoring for each criterion. *The consensus score across reviewers may not reflect the mathematical average.* For purposes of the Figure, the QoR4D criteria have been shortened. Please refer to p. 3 for full criteria definitions. Please note that criterion 8 on ethics was not scored.



*Three consensus scores were greater than 0.5 variance from the mathematical average. Please refer to criteria 4 and 6 above for rationale of these consensus scores (both with a variance of 0.7 from the mathematical average). Criterion 7 received a consensus variance 1.3, which is described in section "Areas of Divergence among Review Team and ISDC Resolution."

4. Fruit and Vegetables for Sustainable Healthy Diets (FRESH)

Review Summary and Actionable Recommendation

The proposal is well articulated and has SMART Work Packages with outcomes that are mostly appropriate. The activities are focused on locations with correctly identified low intake of fruits and vegetables (F&V) and the Theories of Change are, in the main, appropriate, but could benefit from linking with other significant components (explained in action points below) in the food system.

ACTION POINTS

Integrating or considering the following would enhance the proposal's overall QoR4D.

1. Apparent large overlaps between some FRESH Work Packages (1 and 5 especially) and several other CGIAR Initiatives. Research agendas, operationalization plans, and budgets need to be harmonized across all the relevant Initiatives to avoid duplication, replication, and waste of resources.
2. Highlight the role of internal and international trade in F&Vs more strategically. This also requires new product development (NPD) and product innovation, development of quality and safety standards for select F&Vs.
3. A broader set of postharvest technologies and value addition options are needed to extend shelf-life, to complement efforts striving for year-round F&V production and availability.
4. Mapping and leveraging complementary in-country Initiatives, for greater efficiency. For example, programs upgrading/introducing water infrastructure in traditional markets (water being a key food safety concern during marketing, preparation and/or using F&V [i.e., end-to-end]).
5. Spatial inequity in F&V availability/affordability/consumption and more details on planned gender inclusion strategies.
6. Considerable resources are devoted to analysis of diets and food environments, which are also focal areas of many other Initiatives. A strong collaboration and sharing of resources (for example, GAIN and SUN) from the outset could enable some rationalization of resources in these areas. An enabling environment policy workstream could be beneficial.
7. Provision of a disaggregated proposal budget to better assess if the activities and targets are realistic, and that provision for MELIA are adequate.

Overall Strengths of Proposal

A F&V focused Initiative itself is a great positive—for too long, the lack of a CGIAR mandate in this field had contributed to F&Vs having a much lower profile than their importance to global nutrition warranted. FRESH is advancing CGIAR's research on the underused F&Vs to improve food security and nutrition.

Work Package 6 aims to position the Initiative to achieve policy change is well conceived and quite critical to scaling. A major challenge with a F&V Initiative is that in most countries it is a lower priority. Even where it is a priority, F&Vs are viewed from the lens of commercial opportunities and exports, etc. Positioning F&Vs for nutrition will require a transformation of thinking in policy circles, which will require considerable dedicated engagement and facilitation, more so that for many other CGIAR Initiatives. Since Work Package 1 (consumer behavior) and Work Package 5 (food environment) have large intersections with many other CGIAR proposals, including SHIFT, Resilient Cities, TAFSSA, and probably many more, strengthening the internal coordination and rationalization of budgets across other relevant Initiatives could enable more resources to be dedicated to Work Package 6 and other more unique lines of inquiry under FRESH.

Consolidating already existing research areas in CGIAR will drive the One CGIAR concept. This proposal nicely brings together the specific expertise of The World Vegetable Center (WorldVeg) with broader strengths of CGIAR.

Overall Weaknesses of Proposal

The markets part of the portfolio could be considerably more ambitious and innovative. Placing internal (and international) trade in F&Vs at the heart of this work would enable strategic thinking about local production versus trade and avoid the Initiative getting boxed into the erroneous view that a shorter chain is the best way to improve long-term nutritional outcomes. There needs to be a much stronger emphasis on sound, science-based policy development. Inextricably linked to market stimulation/demand, is the need for parallel product development and value addition, including improvements in food safety (biological, chemical, and physical) and quality standards. Addressing these will increase relevance and effectiveness. There is also no sound rationale for recipe development. Some of the narrative about Food Environment research mostly being in rich countries is already a bit dated, since a lot of LMIC work has been coming out recently, and yet more is well underway.

The gender dimension of the project is not explicit, although there was an indication that data would be segregated by gender and the measurable outcomes Nos. 3 & 4 indicate that 50% of the participants would be women. It is unclear if participation by youth, farmers, etc., will be gender sensitive. How will FRESH build capacity of the next generation of researchers? Further, the reviewers argued that thinking about spatial dimensions of F&V consumption, e.g., rural versus urban, is an important equity aspect that the proposal is currently missing.

The capacity building targets need to be put into context, particularly with regard to national stakeholders/institutions. This would facilitate a more objective assessment of the value, relevance, and credibility of the planned capacity building interventions, and justification of priorities. This component is also viewed in terms of sustainability and supporting food environment transformation beyond the life of this project. The proposal is currently weak on proactive thinking about scaling-up pathways; some of these factors that can hinder scaling are included as assumptions, rather the Initiative design proactively addressing them. Investing further in Work Package 6 and making it scale-focused may be a good option. Without clear pathways to scaling (even if CGIAR is not understandably responsible for scaling) should be central to the inquiry. Without scaling pathways, the Initiative will risk just describing diets, food environments, and some policy processes and testing boutique pilot interventions that may have little potential for any impact at the population level in the medium term. Action on the ground needs to go together with appropriate policy development.

Areas of Divergence among Review Team and ISDC Resolution

Not applicable

Does the Initiative Align with the Cohesion of the Portfolio as Described in the Companion Document?

Within the proposal, there is some evidence of internal and external coherence, interdependencies, and management of uncertainties, though they need to be strengthened in parts. These are highlighted throughout the review. Resilience on the ground is highly dependent on local skills and knowledge of an integrated set of stakeholders to adapt to changing food systems and their environments.

This implies capacity building must be strategic and well targeted, along with adequate and sustained contribution of financial resources (commitment) provided by national governments.

Internal coherence and interconnectedness with other Initiatives: There is some evidence presented regarding various Initiatives FRESH would work with, but the proposal does not articulate how. The recognition of interdependencies across Initiatives is currently weak. As commented above, it must be deliberate and regularly reviewed before rolling-out all the relevant Initiatives. The risk of overlap, duplication, and lack of joined-up thinking is high, but this risk could be mitigated if such efforts are made before the roll-out of FRESH and the relevant Initiatives.

External coherence: While there is some evidence of stakeholder consultations, the proposal is weak on how it is positioned vis-à-vis several other strong global and regional research programs, especially related Work Package 1 and 5 and what its comparative advantage is.

The reviewers scored each of the 17 QoR4D criterion individually. The review teams then built a consensus score. In the majority of reviews, the consensus score was the same as the mathematical average of the three individual reviewer scores. The criteria that received between a 0 or 1 are presented with a rationale below. Although not required, in order to keep review reports brief and readable, some teams opted to record in their reports their consensus-built rationale for scores above 1. Following the criteria is a figure presenting the individual reviewer scores and consensus. Please refer to p. 5 for the Likert scoring definitions.

Criteria	Proposal Sections	QoR4D Elements	Consensus Score
1. Clearly defined research problem that addresses Impact Areas, is a high priority in the targeted geographies, is well aligned to shared, multi-funder priorities, and is well informed by previous research findings and evaluations	Challenge statement 2.1, Learning from prior evaluations and Impact Assessments 2.3, Impact statements 5	Relevance, Effectiveness	2
The proposal does reflect the priorities of the target countries and addresses the CGIAR Impact Areas. There is significant evidence on the current status of F&V consumption in the target countries, including its negative impact on economic development and other Impact Areas, which the proposal identifies. However, it would be			

Criteria	Proposal Sections	QoR4D Elements	Consensus Score
important to understand what prospects are expected for larger, poorer countries where the needs are arguably larger. Further, more evidence concerning shifting LMIC F&V consumption significantly via interventions, beyond a few low-hanging fruit areas like home gardens or school feeding is needed. The big wins for the future are potentially via strengthening markets (not just food environments), but this has not been studied adequately.			
2. Evidence that the Initiative is demand driven through codesign with key stakeholders and partners (Investment Advisory Groups, governments, private sector, funders) and research collaborators within and outside CGIAR	Participatory design process 2.6, Challenge statement 2.1, Work Package ToCs 3.2	Relevance, Effectiveness	2
There is good evidence that the Initiative is demand driven. Sri Lanka, Tanzania, and the Philippines had indicated interest in addressing their countries nutrition challenges that are related to the intake of F&Vs. Other CGIAR research groups are involved in the design of the research and have participated in previous related research endeavors. FRESH is collaborating with national and international partners and uses evaluations and impact assessments from the research activities of WorldVEg, nutrition-sensitive Agriculture, and learnings from previous CGIAR research programs. FRESH also intends to improve on CGIAR's competence in technology deliver through partnerships. However, there could be more information on how much <i>demand</i> and commitment there is from local governments, particularly when it comes to nutrition—consultation participation does not always equate to commitment. Government horticulture departments and ministries typically have a mandate around production and exports, with often little conception of F&V for delivering nutrition. An important question is whether these key scaling partners are willing to undergo a transformation in their missions, goals and structures, and the extent to which the target national governments will financially (and otherwise) support this.			
3. Research questions, objectives, outputs, and outcomes are aligned to the research problem, and are measurable with defined deliverables	Work Package ToCs 3.2, Measurable three-year (End of Initiative) outcomes 2.2, Priority-setting 2.4, Management plan 7.1	Relevance, Effectiveness	2
<p>For the most part, there is a good description of research questions, outputs, and outcomes, which are aligned. However, a greater engagement with trade—especially internal, but also international—should serve as a major plank in the strategy set for delivering F&V equitably. Ensuring inter-regional trade of F&V is strengthened so that F&V flows from surplus to deficit areas, and consumption is smoothed over space and time, would seem to be a potentially important piece in the puzzle. Therefore, viewing F&V consumption as a spatial problem may be important.</p> <p>The six Work Packages are related to one another and will elicit collaboration of the government, private sector, academia, other CGIAR Initiatives and the target population for mutual benefits. The Work Packages, if properly executed, will lead to achievement of the goals of the project.</p> <p>Compared to preharvest practices and the production environment, safety risks associated with F&V marketing (particularly within informal markets), processing and utilization downstream (postharvest) are inadequately understood and poorly quantified. With increased consumption of F&Vs, comes an increase in associated safety hazards and risks, thereby negating desired nutritional outcomes. It would deepen the QoR4D if the planned innovations would deliberately target improvements in food safety, and not leave it optional as is implied in the proposal.</p>			
4. Overall Theory of Change with intended outputs, outcomes, and impacts at scale clearly described. Assumptions are documented, causal linkages are clear, especially the role of partners in driving impact	Full Initiative ToC 3.1	Effectiveness, Relevance	2
<p>The overall Theory of Change could be strengthened. This is elaborated upon in 4a below.</p> <p>Some of the text about End of Initiative outcomes in section 2.2 is different compared to End of Initiative outcomes listed in this section. For example, the Theory of Change section says the End of Initiative outcomes is 12 new post-harvest innovations piloted, while section 2.2 talks about at least four post-harvest innovations ready for scaling.</p>			

Criteria	Proposal Sections	QoR4D Elements	Consensus Score
<p>4.a Individual work package Theories of Change (score individually)</p> <p>Work Package 1 Work Package 2 Work Package 3 Work Package 4 Work Package 5 Work Package 6</p>	Work Package ToCs 3.2	Effectiveness, Relevance	WP1: 2 WP2: 2 WP3: 2 WP4: 1 WP5: 2 WP6: 3
<p>Work Package 1 (consumer behavior) and WP5 (food environment) have large intersections with many other One CGIAR proposals, including SHIFT, Resilient Cities, TAFSSA, and probably many more. These are important areas at the heart of much future CGIAR work, but it does beg the question, how efficient a use of resources is this, especially considering that global research community already has a large amount of work going on relating to diets and food environments? The team are aware of plans to coordinate with SHIFT, and recognize strong coordination that addresses questions around innovation, repetitive work and shared resources will be crucial.</p> <p>It is timely the program aims to generate scientific data to explain the poor consumption of fruits and vegetables even when they are accessible and affordable, but some aspects of WP1 are not entirely convincing – ideas such as recipe development have not shown any signs of contributing to improving consumption at scale in this past.</p> <p>Work Package 4: As discussed previously, is light on trade. Opportunities for value addition and new product development where year-round production is not possible, or where yields are significantly greater than expected (to avoid postharvest losses) is an important research area to include, as a complimentary (and program risk reduction) activity for greater impact.</p> <p>Some of the narrative about Food Environment research (WP5) mostly being mostly in wealthier countries is already a bit dated, since a lot of LMIC work has been coming our recently, and yet more is underway.</p> <p>Work Package 6: This is a critical WP for scaling. It will be a lot of hard work to get this working effectively, with continuous investments needed in facilitating stakeholder convergence, especially traditionally minded government ministries.</p>			
<p>5. Research methodology and methods (and supporting activities) are fit-for-purpose, feasible, and assumptions and risks are clearly stated</p>	Work Package ToCs 3.2, Priority-setting 2.4, Innovation Packages and Scaling Readiness Plan 4.1	Credibility, Relevance, Effectiveness	2
<p>The Theory of Change for all Work Packages are coherent and aligned with the overall Theory of Change and Impact Areas. The science is sound, and the research questions and methods appropriate.</p> <p>Studying the nutrition implications of interventions in parts of value chains other than the consumer level are challenging. Modelling, e.g., partial equilibrium or spatial or systems models might be natural ways to approach these. It would also be interesting to understand if formative research will feature in any of the target locations, particularly for Work Packages 5 and 6 given that context is highly variable within countries. Lastly, some aspects of the research methods require improved details (e.g., Work Package 2). The challenge posed by lack of adequate field trials and poor understanding of the agroecology in relation to seed systems is real and requires attention (Work Package 3).</p>			
<p>6. Analysis of trade-offs and synergies across the CGIAR Impact Areas; ex-ante assessment of project benefits provides logical rationale for scaling of impacts</p>	Projection of benefits 2.7, Result framework 6.1, Impact statements 5, Innovation Packages and Scaling Readiness Plan 4.1	Effectiveness, Credibility	2
<p>Some <i>ex-ante</i> assessment of benefits seems a little optimistic, e.g., a high (50 to 80%) probability of having a lifesaving/transformational/substantial effect on the number of people meeting min. micronutrient requirements. Text says the probability is high due to the high priority given to healthy diets and noncommunicable diseases in study countries, but there's no evidence of such clear momentum in practice within the study countries. Again, this necessitates a strong policy component and scaling partners.</p>			

Criteria	Proposal Sections	QoR4D Elements	Consensus Score
<p>As noted previously, the gender potential is not very clear—should metrics be focusing on other aspects of socio-economic inclusion?</p> <p>Trade-offs are mentioned occasionally, but not elaborated, and dialogue with non-CGIAR associates on trade-offs and synergies concerning CGIAR Impact Areas would be beneficial, to get a broader view on scaling. This is also linked to criterion 7 below.</p>			
<p>7. Evidence that the Initiative will likely lead to impact at scale through integrated systems approaches that drive innovation in research and partnerships, including linking to and leveraging of other Initiatives within and outside CGIAR</p>	<p>Projection of benefits 2.7, Work Package research plans and ToCs 3.2</p>	<p>Effectiveness, Credibility, Relevance,</p>	<p>2</p>
<p>The project will promote nutrition, health and food security, poverty reduction, climate adaptation and environmental health and biodiversity. The Theory of Change and causal links are well defined in the proposed Initiative. However, linked to the need to increase post-harvest value addition, address negative consumer perception on safety and hazards associated with consumption of F&V, and importantly prioritize the real risks (particularly for vulnerable groups) from biological and chemical hazards, the food safety component could be strengthened, in order to realize the desired longer-term outcomes of improved consumption and nutrition. "Food is not nutritious, if not safe."</p> <p>The probability of the impacts being beneficial has been estimated to be 50-80%, with the exception of gender equality and social inclusion. Good links proposed with synergistic initiatives like SHIFT, Resilient Cities, etc. As the proposal notes, scaling cannot be funded by the Initiative, and so a lot depends on careful choice of scaling partners. Again, investing further in WP6 and making it scale-focused may be a good way to go.</p>			
<p>9. Research design and proposed implementation demonstrates gender and social inclusion that can be tracked in outcomes</p>	<p>Gender equality, youth & social inclusion 5.3, Projection of benefits 2.7</p>	<p>Legitimacy, Effectiveness</p>	<p>2</p>
<p>The method that will be used for gender inclusion is not clearly stated. It is mentioned that data will be segregated by gender, but how the gender data would be generated is not explicit. The ability of F&V interventions to be gender transformative is also unclear within the proposal, and good gendered insights to base interventions on are often not available. Further, perhaps partly because the proposal form focuses on gender and youth as the key dimensions of equity, the proposal does not adequately address other dimensions of equity that are critical. Arguably, a lot of inequity in F&V is spatial—e.g., rural/urban or equity based on economic distances from markets, and it would have been good to see such important aspects developed solidly. Much can be drawn on planned inclusion strategies from learnings cited from previous research.</p>			
<p>10. A risk framework that details main project risks and mitigation actions across areas of science, funding, operations, partnerships, ethics, and environment</p>	<p>Risk assessment 7.3</p>	<p>Credibility, Legitimacy, Relevance</p>	<p>3</p>
<p>The evaluation of risks is realistic. A significant risk warranting more focus, concerns inadequate incentivization of policy actors, and the potential impact of the program. If the "areas of weakness" highlighted in this review are strengthened, then detailed consideration of those associated risks and mitigation will be required.</p>			
<p>11. CGIAR capacity and its comparative advantage and appropriateness to lead the work is justified. This includes the skills, diversity and multi-/trans-disciplinarity of the research team and approaches to meeting gender and diversity targets</p>	<p>Comparative advantage 2.5, Initiative team 9.1, Gender, diversity and inclusion in the workplace 9.2</p>	<p>Relevance, Legitimacy, Effectiveness</p>	<p>2</p>
<p>This comes across as a WorldVeg led proposal with a CGIAR plus others supporting, even if on paper WorldVeg is not the prime organization. But this is entirely reasonable given the comparative advantage and long history of work that WorldVeg brings to the table. There is a solid CGIAR supporting cast. The comparative advantage of the other institutions is not at all clear. For examples, what do Institute for Development Studies or WUR contribute that cannot be found in CGIAR?</p> <p>The range of skills is generally appropriate, but a greater concentration of economists (not the same as value chain experts) to ensure key aspects—such as efficiency of short value chains versus strengthening internal/international trade, investment cases, incentives of value chain actors, nature of market failures, etc.—would strengthen the proposal. Public and private partners are also critical in providing the balanced and breadth</p>			

Criteria	Proposal Sections	QoR4D Elements	Consensus Score
of experts required across the different product types, food systems, and dynamic food environments, integrating the One Health approach.			
12. Capacity building within project teams, partners, and stakeholders captured in capacity development plan. This can include development of early career researchers and partner staff, support/empowerment for under-represented stakeholders, and building partner networks	Capacity development 9.3	Credibility, Legitimacy	2
Capacity building within project teams could be more explicit and inclusive especially for early career researchers who would provide continuity in research. The team may also want to consider developing specific training on interdisciplinary working, particularly where it comes to local partner institutions that may be strongly siloed. It is not clear (i.e., assessment of needs not fully articulated) if the proposed targets will be impactful long-term in the context of current skills and knowledge deficits across the target countries' institutions in those areas the intervention is targeting.			
13. Project management mechanisms and (if applicable) additional scientific oversight and governance measures effectively and efficiently support the Initiative objectives	Management plan and Risk assessment 7, Research governance 8.1	Legitimacy, Credibility	2
Project management is generically described. Further details on the external advisory panel who might contribute to the periodic MELIA would be helpful—would this also provide scientific oversight? The Initiative may benefit from marketing economics/value chain expertise outside of the CGIAR/WorldVeg group.			
14. Justified and transparent costing explicitly linked to expected Research for Development results	Financial Resources 10	Legitimacy, Effectiveness	1
Justification for the budget is not presented. This makes it difficult to evaluate if there is an appropriate balance of financial resources across the Work Packages, and whether those funds are indeed sufficient/realistic for the proposed activities. The proposal might consider consolidating funds allocated to diets and food environments across Initiatives, and for this Initiative, repurpose additional funds towards Work Package 6 and especially towards building political capital. The latter would also reorient more of the budget towards LMIC geographies.			
15. Anticipated research outputs (knowledge, technical, or institutional advances, specific technologies or products, policy analyses) are described and knowledge/gaps they will fill are evident. Protocols for open-data and open-access compliance are evident in plan (including budget)	Work Package research plans and ToCs 3.2, Open and FAIR data assets 8.2	Credibility, Effectiveness	2
The output and outcomes for most of the proposal are clearly stated, though research outputs could be enhanced if the knowledge gaps concerned with some of the Work Packages (see criterion 4a above for details) are addressed.			
16. Monitoring, evaluation & learning (MEL) plan for the Initiative is clearly defined, with flexibility to adapt. MEL plan supports effective management and learning, including baseline data collection, and evaluative and review processes corresponding to stage-gates and course-correction decisions. MEL occurs during the life of Initiative and is used proactively to reflect on and adapt the Theory of Change, where appropriate	MELIA plan 6.2, Planned MELIA studies and activities 6.3, Measurable three-year (End of Initiative) outcomes 2.2	Credibility, Effectiveness, Legitimacy	2
The monitoring and evaluation plan are clearly explained. Quarterly progress reports on Work Packages and annual reports will be strengthened with annual rapid assessments and stakeholder interviews to monitor the progress of the Theories of Change. It is noted that randomized controlled trials will be undertaken wherever possible. But there is need to keep in mind that many of the upstream evaluations may not be easily amenable to randomized controlled trials—realistically randomized controlled trials may only be applicable to a small proportion of the research being undertaken here. The preference for randomized controlled trials must not bind the evaluation questions that are asked—i.e., do not just look for the keys where the streetlights are. Alternative methods, including market			

Criteria	Proposal Sections	QoR4D Elements	Consensus Score
<p>simulation models that may not strictly be impact assessment might well need to come into play and need to be planned for early on in the project.</p> <p>The proposal should create provision for post-Initiative MEL, in order to facilitate impact assessment, capture unforeseen externalities, enhance sustainability, and reduce cost of future baseline studies. This implies sufficient capacity at stakeholder level, and appropriate stakeholder indicators which are also easy to capture during routine operations. It is not possible to assess if the budget is sufficient to cover the MELIA plan.</p>			
<p>17. Well-defined plan for Initiative-level evaluation and impact assessment based on expected end-of-Initiative outcomes and impact. Links between the impact assessment plan and indicators in the Theory of Change are clear</p>	<p>MELIA plan 6.2, Planned MELIA studies and activities 6.3, Full Initiative ToC 3.1, Work Package ToCs 3.2, Projection of benefits 2.7</p>	<p>Effectiveness, Relevance</p>	<p>2</p>
<p>The MELIA studies map reasonably on to the End of Initiative outcomes. Baselines of 'Impact assessment learning studies' for evaluating the end-to-end approach are noted, but it is unclear why this will not apply to all countries?</p> <p>The monitoring and evaluation and learning plan intends to use the already existing CGIAR Performance and Result System. Annual rapid assessments and stakeholder interviews intend to complement the progress reports, but it is not entirely clear who will be responsible for the annual rapid assessments, and whether the assessors will be an independent consultancy group. Adoption, process evaluation and impact studies will also serve as evaluation of the success of the activities.</p> <p>Again, it would be useful to look at the budget assigned to this component.</p>			

Additional Comments Not Presented Above

The review team wishes to expand on select Work Packages.

Work Package 2: The conservation of germplasm to improve diversity and strengthen seed systems. This Work Package will promote sustainability even in the face of climate change. The FRESH team could reconsider whether the timeline for this Work Package is achievable given the diversity of genetic materials involved as well as the need for research uptake activities by the end users (analyzing gene banks, inventories through monitoring biodiversity in farmers' fields could be a longer-term project?).

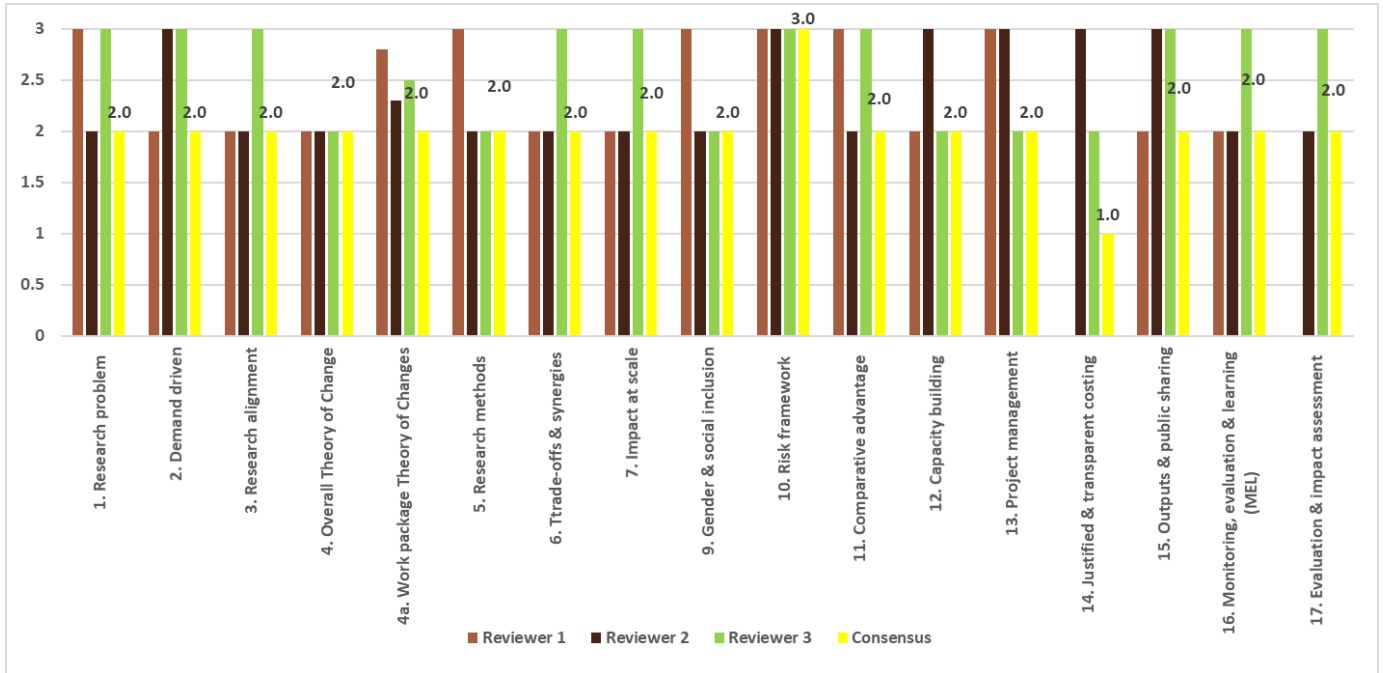
Work Package 3: Low adoption rates of improved F&V varieties are typical. The project may need to research more widely, the performance of selected breeds across a wider range of ecological zones to assess yields and other desired traits by farmers, consumers and processors. With regards to genetic innovation, the approach presented is well aligned, though there needs to be more evidence of the specific market intelligence and economic analyses (nationally, regionally, and globally) feeding into market demand, and thus, rationale/justification for targeting select crops for promotion (breeding, value addition, and/or postharvest technology support). There should be a causal link with operational logic for the delivery of CGIAR results at nested level. As such, perhaps the mechanisms for external cohesion across the planned target countries could be re-examined.

Work Package 4: There are numerous examples of dried F&Vs which could be promoted for increased food security (particularly in light of climate variability), product differentiation for local, regional or global markets, with the potential to improve incomes and livelihoods for many smallholder farmers, micro and small enterprises. These opportunities could be discussed further—recognizing, however, that systemic changes in these traditional value chains will need to be addressed across the pre/postharvest continuum.

Work Package 6: Scaling also needs to consider the role of extension services that are critical for communicating realities on the ground, constraints around social and cultural norms, and supporting farmers and other stakeholders on technology adoption. Experience shows there is considerable competition from several Initiatives within the same regions for the same extension services. Does the current proposal recognize this, and how do they propose addressing this issue, which could limit achieving targets across all Work Packages.

More insight on how FRESH will adjust if scaling up partners are unwilling or unable to finance the scaling process is required.

The Figure below represents original scoring from each reviewer and consensus scoring for each criterion. *The consensus score across reviewers may not reflect the mathematical average.* For purposes of the Figure, the QoR4D criteria have been shortened. Please refer to p. 3 for full criteria definitions. Please note that criterion 8 on ethics was not scored.



*Seven consensus scores were greater than 0.5 variance from the mathematical average. Please refer to criteria 1, 2, 11, 13, 14, and 15 (all with a variance of 0.7 from the mathematical average) and criterion 4 (variance of 0.5 from the mathematical average) above for rationale of these consensus scores.

5. Harnessing Digital Technologies for Timely Decision-Making across Food, Water, and Land Systems

Review Summary and Actionable Recommendation

The Initiative (DX1) centers upon the three challenge areas of (1) the digital divide, (2) inadequate information and (3) limited digital capabilities. Overall, DX1 is a critical avenue for ensuring beneficial impact. A large part of the proposal was articulated in an intentional, appropriate, explicit, and convincing way. To further improve the rigor, legitimacy, credibility and effectiveness of DX1, the review suggests the following eight actionable recommendations:

- 1. Clearer definitions of digital technologies:** While repeated throughout the proposal, the meaning of this term is nuanced and value laden. The title of the Initiative suggests digital technologies will be used to make more timely decisions, though the body of DX1 suggests many different uses for digital technology. It would be prudent to define "digital technologies" at the beginning of the proposal, as different actors may have very different interpretations of what digital technologies actually entail, and thus how they can be effectively promulgated for impact. Armed with this definition, the Initiative team will be better able to distinguish between aspirations of *generating positive outcomes per se* versus *development of processes and technology* for generating positive outcomes.
- 2. Improved linkage of DX1 Challenge Areas with CGIAR Impact Areas:** Review suggests that the Initiative team more clearly link the three DX1 challenge areas to the CGIAR Impact Areas in which these technologies can be used to improve food security, human health, and environmental outcomes. This is because technology development and deployment are not enough: new technologies and innovations must be coupled with appropriate science and extension that will enable delivery of regional benefits (the latter may be delivered by other CGIAR Initiatives).
- 3. Work Package detail:** More specifics would help readers better grasp the background and intent of these, particularly the descriptions of the proposed methods. There also appears to be overlap between work package outputs. Part of this could be the inclusion of greater detail on how digital technologies will be used for monitoring and planning food systems in advance: much of the proposal seems to outline more *reactive* uses of digital technologies, rather than being *proactive*.
- 4. Greater emphasis on iterative refinement of proposed innovations during the life of DX1:** The Initiative focuses more on provider offerings (e.g., service suppliers) and less on the needs of target end-users (e.g., farmers, governments). While DX1 does propose some investigation of demand-side issues, the main emphasis is on the supply-side (most of the 83 stakeholder groups consulted are not end-users, e.g., farmers or industry). More effort on the iterative refinement with next- and end-users during the life of DX1 (as opposed to only at the end) may help ensure that proposed innovations genuinely translate into beneficial and intended impact. A key example of this is in Work Package 1: more immediate impact could be obtained by integrating digital infrastructure providers and investors (e.g., telecommunications and phone service companies etc.) into the design process. If the aim of Work Package 1 is to catalyze access to digital technologies, it is essential that at least telecommunications providers be part of the codesign process.
- 5. Risk that proposed gender diversity, inclusivity and equity targets do not come to fruition:** It was very pleasing to see the focus of DX1 on engagement with women. The authors should consider adding risk mitigation measure/s to address scenarios in which women do not effectively engage in this Initiative. If men are more inclined to be involved in codesign, extension etc., the current DX1 *modus operandi* may obviate substantial opportunities for impact. Perhaps DX1 proponents could first prioritize gender diversity and inclusivity, but as a second priority provide scope for engagement with other members of society through which an obvious and straightforward path to impact exists. In the same vein, the MELIA metrics tend to focus on women and youth (e.g., p. 39) but say little about other marginalized members of society, e.g., the rural poor. This language could be broadened. Finally, the COVID-19 pandemic provoked significant changes in the need for digital technologies in many rural areas of the world. Consideration of how baseline values measured prior to the pandemic may have changed since the time of writing the proposal would be worthwhile.
- 6. Budget assumptions and justification:** Details underpinning budget assumptions seem completely absent; as such, reviewers had difficulty understanding and justifying proposed costs. In addition to some description of how the budget was derived, the Initiative team should consider adding cost-benefit analyses to determine if envisioned financial benefits resulting from access to new technologies AND agrifood systems extension outweigh not only the direct costs of DX1, but more importantly, the expected costs to end-users. Part of this is the risk that adopted technologies are economically unviable. Approaches to identifying adverse economic outcomes resulting from DX1 need greater consideration and documentation.
- 7. Legacy of digital technologies:** Part of DX1 includes the development of decision support tools. The Initiative team should clearly outline proposed processes for ensuring legacy of these tools after the Initiative

concludes: who and how will these tools be maintained and kept fit-for-purpose after DX1 concludes? Legacy through public-private partnerships and commercial hosts of digital tools could be more clearly elicited; this information would fit well in Work Package 5. The distinctions between development of *technology* per se and *conceptual processes* for developing technology (outlined above) would fit well here.

- 8. Attribution of impact associated with DX1:** Given the lofty aspirations of DX1, it may be difficult for proponents to attribute impact without strong counterfactuals that segregate and lend credence to the Initiative impact estimates. During the life of the Initiative, other technology developments and adoption will co-occur that are not associated with DX1 or even with CGIAR: to effectively attribute impact of DX1 from that of background developments, the authors need to clearly define counterfactual/control groups and questions (e.g., does capacity training result in improved digital literacy, and how is this measured?). In cases where control/counterfactual groups are not well defined and planned, could the authors highlight examples of MELIA evaluations on other CGIAR Initiatives that have rigorously assess the important but often recalcitrant outcomes they are aiming to change? Measurement and attribution of impact through comparisons of communities that are advanced in the use of digital technologies with those significantly lagging behind could be part of this.

Overall Strengths of Proposal

Challenge statements and priority settings: The Initiative has been codesigned with stakeholders: 83 groups were engaged in the design of this Initiative (2.4, Priority settings, 2.6 Participatory design process) but many of these groups were not end-users. The participatory design process used to develop the three challenge areas has been well considered and adequately explained. The three priority areas are very interlinked, e.g., women are less likely to have phones and more likely to have lower literacy levels due to lesser school attendance (social norms), resulting in reduced digital literacy and fewer women using mobile phones will reduce demand for appropriate digital content. Perhaps a sentence linking these would highlight this?

Aspirations of Work Packages 3 and 4: As described, these components are very compelling and may, if effective, overcome many low-information/low-data barriers that plague low- and middle-income countries (LMIC). Assuming DX1 actors successfully execute plans and with sufficient geographical resolution, Work Packages 3 and 4 hold significant promise in helping many people in case of emergencies caused by natural disasters. These Work Packages aligns very well with other CGIAR Initiatives described, highlighting the lengths that the developers of DX1 have went to in order to align the current Initiative with others.

Codesign with stakeholders within and outside CGIAR: The Initiative team clearly show they have engaged with 19 other Initiatives, which is very encouraging. This suggests that DX1 actors will ensure collaboration across other CGIAR Initiatives, thereby strengthening research methods, findings and their comparability, and avoiding duplication of efforts that seems to be all too commonplace in contemporary scientific institutions.

Overall Weaknesses of Proposal

Clarification of where local end-user needs are incorporated in the design process: DX1 primarily focuses on the supply-side (e.g., digital startups) and less so on the demand-side (end-users needs). Addressing both supply and demand issues are important for success, but oftentimes research and technology are delivered in a top-down fashion, rather than starting from end-user needs (e.g., smallholder farmers and rural poor) and how these may be overcome. The authors could clarify where the local demand is coming from (e.g., partners in South Africa, Botswana, Zimbabwe) and how it will be incorporated into the research codesign process.

Mitigation of risk that engagement targets for gender diversity, inclusivity, and equity regarding engagement are not met: It was extremely pleasing to see the focus on engagement with women and marginalized groups (e.g., section 2.2). Reviewers suggest the Initiative team also include risk mitigation measure(s) that address scenarios in which such groups do not or cannot engage in this Initiative. If it is only men or traditional end-users that engage, DX1 stands to obviate substantial potential impact. The intent of these sections could be better framed: first a priority on women and marginalized groups, and second a priority on other actors through which an obvious and straightforward path to impact also exists. Clarification of how baselines will be established would be an important addition; at present, DX1 assumes all regions could improve, with the implicit assumption of relatively similar baselines. This may not be the case.

Work Package 1: This Work Package aims to support policymakers and investors, although it only goes part way towards the impact needed to fully address the digital divide. More immediate impact could be obtained (within the life of the Initiative) by working directly with digital infrastructure providers and investors (telecommunications and mobile phone service companies etc.). Assuming that the "digital divide" challenge area of DX1 aims to enable access to digital technologies, it is essential that telecommunications providers be part of the codesign process. The aim should be not just to investigate rural digital infrastructure, it should be to work directly with digital infrastructure providers and enable investment and change to happen. This would help soften

the focus on research and agriculture in Work Package 1 and strengthen actions and iterative codesign with a broad array of stakeholders to better define bottle including digital infrastructure providers.

Areas of Divergence among Review Team and ISDC Resolution

Not applicable.

Does the Initiative Align with the Cohesion of the Portfolio as Described in the Companion Document?

External coherence (country, regional, global levels): This aspect is overall very good, although some reviewers questioned the choice of criteria used to select the target countries, as well as the problem definition through consultation with partners in the United States, Europe, and Australia (i.e., non-target countries). Greater emphasis on end-user demand and needs rather than top-down suppliers (such as agtech companies) will be critical for success of DX1.

Internal coherence including Impact Area Platforms: While the Work Packages were clearly defined, the inter-linkages between them could be more annunciated. The linkages between the research plan and the proposed methods also seem vague: this sometimes leads to a lack of continuity between the research plan, scientific methods, and outputs. Internal coherence could be improved by stating how methods and outputs from each Work Package will feed into the other. In terms of management, there is a need for a person in the Integrated Management Team to have oversight of all Work Packages to provide coordination and consistency (assuming this role does not already exist).

Interdependencies between other thematic/regional Initiatives: By definition, DX1 is more focused on methods (i.e., digital technology) rather than outcomes (generation of positive economic, environmental, social outcomes e.g., see Outcome boxes in each Work Package). As such, there is a tension between the creation of *outcomes* and the creation of *frameworks* (technology development and adoption) used to catalyze positive *outcomes*. To clarify what *outcomes* will be led by DX1 relative to those generated in other Initiatives (e.g., ClimBER, NEXUS Grains, MITIGATE+), the authors could add a column to the End of Initiatives outcomes in the tables on p. 46 that clearly delineate which outcomes are led by DX1, and which come from other Initiatives. This will help proponents better shape their intentions towards creation of beneficial outcomes for society, whether it be through alliances or from research within DX1 per se.

Management of funding uncertainties: This aspect was completely absent from the proposal. Perhaps this deficit is a short coming in the proposal template design.

Integrated results framework at Initiative, Action Area and CGIAR Levels: In general, this is well covered.

Measurement and reporting at multiple levels and timeframes: *Ex-post* impact assessments of Work Packages are covered, but *ex-post* assessments of the monitoring and evaluation (and how this assessment will be used to refine management going forwards) could be more transparent. Linkage of management to the Work Packages could be more detailed (metrics, aim, achievability and timing). The MELIA plan is reasonable overall, although it is difficult to judge how and where baselines will be collected (e.g., p. 54). "Causal impact evaluation" is relegated to the last year of the Initiative (e.g., p. 54); perhaps this would be better incorporated throughout the implementation process.

The Measurable End of Initiative outcomes for the "digital divide" challenge area seem to focus more on strengthening digital agrifood systems, rather than the digital divide per se, i.e., the people who have limited or nil access to mobile internet, digital services etc.

The reviewers scored each of the 17 QoR4D criterion individually. The review teams then built a consensus score. In the majority of reviews, the consensus score was the same as the mathematical average of the three individual reviewer scores. The criteria that received between a 0 or 1 are presented with a rationale below. Although not required, in order to keep review reports brief and readable, some teams opted to record in their reports their consensus-built rationale for scores above 1. Following the criteria is a figure presenting the individual reviewer scores and consensus. Please refer to p. 3 for the Likert scoring definitions.

Criteria	Proposal Sections	QoR4D Elements	Consensus Score
1. Clearly defined research problem that addresses Impact Areas, is a high priority in the targeted geographies, is well aligned to shared, multi-funder priorities, and is	Challenge statement 2.1, Learning from prior evaluations	Relevance, Effectiveness	2

Criteria	Proposal Sections	QoR4D Elements	Consensus Score
well informed by previous research findings and evaluations	and Impact Assessments 2.3, Impact statements 5		
<p>Challenge statement: The research problem clearly acknowledges the social norms and soft forces that underpin digital inclusion. The context surrounding the digital divide, information inadequacy and lack of digital capabilities provides a compelling need for promulgation of these technologies in LMICs.</p> <p>Clear and convincing research problems: The research questions are very broad and could be more explicit. It is implied that DX1 will address the stated underlying problems, either <i>per se</i> or as part of other Initiatives (e.g., MITIGATE+, ClimBER). The "Digital Divide" could be better rephrased as 'Lack of supporting telecommunications infrastructure' or similar.</p> <p>Learning from prior evaluations and impact assessments (section 2.3): This section speaks more about what this Initiative <i>will do</i> (e.g., sustainable adoption of digital technologies will be pursued...) rather than what the authors <i>have learnt from</i> prior evaluations.</p> <p>Impact statements (section 5): Women and youth are bundled together in the impact statements although each group may have very different prerogatives for wanting to be part of DX1. It would be useful to point this out. As well, climate risk and gender probably do not need to be linked in these sections.</p>			
2. Evidence that the Initiative is demand driven through codesign with key stakeholders and partners (Investment Advisory Groups, governments, private sector, funders) and research collaborators within and outside CGIAR	Participatory design process 2.6, Challenge statement 2.1, Work Package ToCs 3.2	Relevance, Effectiveness	2
<p>Overall: It is clear that many stakeholders were consulted in the design of DX1, including several focused explicitly on gender. It would be useful to delve further into local organizations in this space (around gender, digital inclusion, youth, etc.) in the geographies proposed, as soft forces at a local level often look very different to findings presented by multilaterals/international organizations.</p> <p>Consistency of groups consulted across the report: Ethiopia, Japan, Korea, and Indonesia are included in the list of groups consulted but these countries do not feature in target list at the beginning of the report. Similarly, several countries in the target list do not seem to be in the consultation set (from Mexico, Guatemala, Egypt, Kenya, Rwanda, Malawi, Bangladesh (Ganges Delta), India (Odisha), Nepal, Limpopo River Basin, only organizations in Kenya and Mozambique are included in the list). There may be potential for double counting. For example, Yara seems to be counted multiple times. It is not overt that this is demand-driven and codesigned by partners in the target regions. As stated, the text suggests that partners in South Africa, Botswana and Zimbabwe <i>will be</i> consulted, but have not yet been engaged in the design of this component.</p> <p>Clarification of where the localized end-user needs are incorporated into the design process: DX1 collaboration with Grow Asia as part of their learning event focuses mainly on the supply-side (digital startups) rather than the demand-side (what end-users need and want). Addressing both supply and demand issues are key to success, but oftentimes such projects are delivered in a top-down fashion, e.g., from companies looking for markets, rather than starting from, e.g., smallholder farmer needs and how to overcome them. The proposal authors could clarify where the local demand is coming from (e.g., partners in South Africa, Botswana, Zimbabwe), as well as how this will be incorporated into the research codesign process.</p> <p>Linking CGIAR Impact Areas with DX1 problem areas (section 2.6): While some technologies are mentioned (e.g., low-earth orbit satellites), DX1 does not go far enough in terms of describing how it will deliver food security, alleviate poverty, transform agri-food systems etc. The Initiative authors are suggested to better link the three DX1 problem areas to specific technologies/mechanisms in which they can be used to improve society. Having technology is not enough: new innovations <u>must</u> be coupled with appropriate codesign and extension that will deliver regional benefits. More details of collaborative linkages with other science-focused CGIAR Initiatives would be helpful in this regard.</p>			
3. Research questions, objectives, outputs, and outcomes are aligned to the research problem, and are measurable with defined deliverables	Work Package ToCs 3.2, Measurable three-year (End of Initiative) outcomes 2.2,	Relevance, Effectiveness	1

Criteria	Proposal Sections	QoR4D Elements	Consensus Score
	Priority-setting 2.4, Management plan 7.1		
<p>Work Package outputs and End of Initiative outcomes: Overall, the review team found these too vague. To be measurable, readers MUST be able to understand <u>what</u> is being measured. For example, “the digital agrifood ecosystem is strengthened through >5 policies, investments, and partnerships” could mean anything, from a partnership between two neighboring smallholder farmers, to an UNFCCC global policy aiming to raise adoption of digital technologies by the rural poor. Words such as design and boost without clear definitions lend themselves to be manipulated in MELIA work as they are too subjective. Clear descriptions of actual baselines and the timing of their measurement are also critically important for success. The authors should try to be more specific, measurable, achievable, realistic and time-bound (SMART) in describing End of Initiative outcomes.</p> <p>The agrifood startups, NRM organizations and governments referred to should be geographically pinned, because as written the EoI metrics could all be achieved all in the one country (so the goals of the Initiative are met) while many targeted countries miss out.</p> <p>Priority setting (2.4): Similar to the above, the reviewers suggest that the authors adopt more quantitative metrics relative to those currently listed e.g., “lower than normal” and “higher than normal.”</p>			
<p>4. Overall Theory of Change with intended outputs, outcomes, and impacts at scale clearly described. Assumptions are documented, causal linkages are clear, especially the role of partners in driving impact</p>	Full Initiative ToC 3.1	Effectiveness, Relevance	2
<p>Overall: While the text description supporting the Theory of Change (ToC) does list the assumptions as required in a theory of change, the three DX1 challenge areas are not described as connected components of a systemic problem. If appropriate, ToC could also include women's marginalization in digital ecosystems. The text seems more solutions-focused rather than problem-focused, lacks a systems lens, and tends to focus on component parts rather than holistic solutions.</p> <p>Theory of Change Work Package 1: As explained further below, we suggest the Theory of Change Work Package 1 should include digital infrastructure and telecommunications providers in the design of potential agrifood solutions. Greater involvement of end-users themselves in the design codesign is required.</p> <p>Theory of Change Work Package 2: would digital innovators be more focused on generating business and so catering for the masses versus diagnosing gender and inclusion gaps in their digital agri-food solutions? What incentives do digital innovators really have to address gender equality and inclusivity relative to their business-as-usual activities?</p> <p>Theory of Change Work Package 3: Theory of Change expected to occur during the Initiative is clearly laid out, but what about after the Initiative concludes? The Initiative team should consider legacy of digital tools developed by researchers—who will upgrade and maintain these tools after the Initiative concludes?</p> <p>Theory of Change Work Package 4: These seem more focused on outcomes rather than theories of change. For example, “Food systems stakeholders will use digital tools...” WHY and HOW will food systems stakeholders use digital tools?</p>			
<p>4.a Individual work package Theories of Change (score individually)</p> <p>Work Package 1 Work Package 2 Work Package 3 Work Package 4 Work Package 5</p>	Work Package ToCs 3.2	Effectiveness, Relevance	WP1: 1 WP2: 2 WP3: 2 WP4: 2 WP5: 2
<p>Overall: more specifics would help readers comprehend the Work Package aims, particularly methods leading up to the outputs. The Theory of Change diagrams seem more solutions driven, rather than thinking through the needs/problem statements. The text explains the Theory of Change very well; these sentiments could be more clearly laid out in the Theory of Change diagrams.</p> <p>The review team also suggest changing terms such as “digital extension apps” to “digital extension services” or broadening the text to any mobile-enabled initiative (e.g., IVR, SMS blasts, low tech solutions etc.), as not everyone has a smartphone.</p> <p>Work Package 1: This aims to support policymakers and investors although only goes part way towards articulating the processes needed to address the digital divide. Much more immediate impact could be obtained by working directly with digital infrastructure providers and investors (telecommunications and mobile phone</p>			

Criteria	Proposal Sections	QoR4D Elements	Consensus Score
<p>service companies, etc.). Assuming that the 'digital divide' part of DX1 aims to enable access to digital technologies, it is essential that telecommunications providers be part of the codesign process. The aim should be not just to investigate rural digital infrastructure, it should be to work directly with digital infrastructure providers and enable investment and change to happen. This is less about research and agriculture, and more about action and iterative codesign with a broad array of stakeholders, including digital infrastructure providers. Outcomes 1.1 and 1.2 would probably fit better in other Work Packages. It was also unclear what some terms mean, e.g., "convergent parallel design approach of mixed methods"—what does this mean? And how does this link to causally assessing the effects of digital access on inclusion, poverty, food security, and jobs?</p> <p>Work Package 2: We were not convinced as to the Theory of Change redesign of digital advisory services to improve gender responsiveness. How can we be sure that advisory services will actually do this? The review team did not have a clear picture of the specific changes yielded from the processes described, and/or other previous/current programs that are/have addressed the issue of gender inclusiveness. This seems important and would probably not need to be created from scratch.</p> <p>Work Package 3: Outcome 3.1 seems more like a case study while outcomes 3.2 and 3.3 seem more generic. Should these outcomes be more consistent? Work Package 3 is a fantastic initiative but the timeline seems ambitious given all of the constituent parts it embodies.</p> <p>Work Package 4: More detail is needed. How do food value chain actors fit into any of these outputs? What is the process for identifying how to optimize sampling frames suggested in RQ4.1?</p> <p>Work Package 5: this touches on the demand-side of problem development to encourage use and uptake, and design iteration is emphasized. This is to be applauded. With respect to data sharing on p.33, some other examples of data sharing procedures that CGIAR is learning from and/or drawing upon as an example of partially restricted access could be useful (e.g., DHS data sharing policies). The Digital Crisis Response Framework seems apt and would be a great way to combine the different strands together. Overall, Work Package 5 still comes across as vague (e.g., the review team had no knowledge of the South x South Collaboration Lab) Again, more details would be helpful.</p>			
<p>5. Research methodology and methods (and supporting activities) are fit-for-purpose, feasible, and assumptions and risks are clearly stated</p>	<p>Work Package ToCs 3.2, Priority-setting 2.4, Innovation Packages and Scaling Readiness Plan 4.1</p>	<p>Credibility, Relevance, Effectiveness</p>	<p>2</p>
<p>Major issues: support of policymaker's efforts probably only goes part way towards the intent needed to address the digital divide. Would more immediate impact be obtained (within the Initiative) by working directly with digital infrastructure providers (telecommunications and mobile phone service companies, etc.)? If the digital divide is about enabling access to digital technologies, the telecommunications providers must be part of the codesign process.</p> <p>The methods as described do not create a rigorous framework for testing the questions claimed as "research questions" for each Work Package. There is no mention of control groups, counterfactuals, etc., and so it is DX1 actors plan to disentangle the impact of their efforts compared with other simultaneous background trends that will co-occur as this Initiative is rolled out.</p> <p>Minor issues:</p> <ul style="list-style-type: none"> The assumptions and risks underpinning the methods could be more explicit. Such additional detail would be best placed in the descriptions of the Work Packages. There is a tendency in digital development for people to focus on emerging and "bright shiny object" tech as opposed to simpler solutions; indeed, the latter are often overlooked (e.g., leveraging existing social groups, use-case driven approaches, community radio, etc.), which in some cases can be a better fit and more sustainable. This often results in the <i>solution driving the problem</i>—similar to the focus on the supply side rather than the demand side outlined above. As such we suggest that the Initiative team tone-down language where appropriate. For example: "A range of forward-looking solutions will be considered, including emerging communication technologies currently being field-tested by DX1 partners like low earth orbit satellites and optical wireless" Budget support in some cases seems limited if the aim is to engage many organizations and initiatives. Perhaps 'light-touch' is more accurate than 'dedicated' in statements such as this: "2.5% of our budget (US\$750,000) over three years to provide dedicated support for 20 Core Innovations, 5 Innovation Packages, and 1 Portfolio Management System." 			

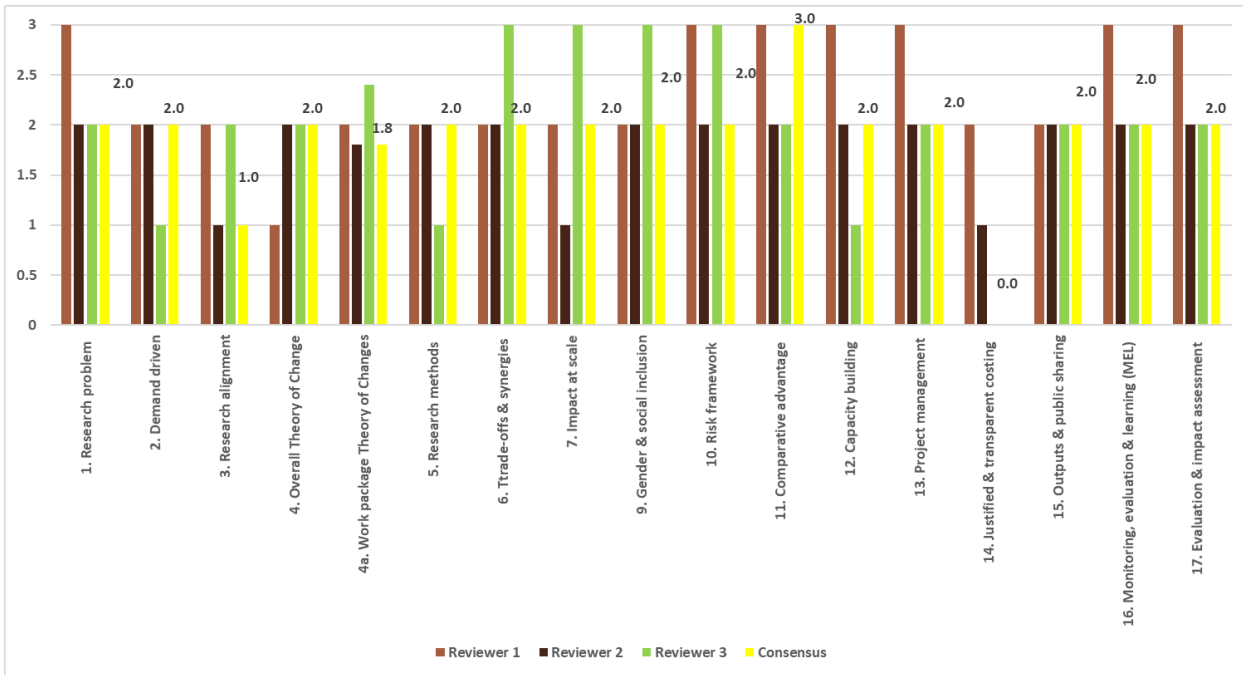
Criteria	Proposal Sections	QoR4D Elements	Consensus Score
<ul style="list-style-type: none"> The assumption on income-related impacts appears to be based on a single paper that uses methods (propensity score matching on cross-sectional data) that are not currently considered sufficiently robust for causal claims. "CTA-reported value" is used in several places as a supporting assumption underpinning probability of success, but the term does not seem to be defined. The methods do not distinguish between mobile subscriptions and individuals, which are distinct entities (a single individual can have multiple mobile subscriptions). The impact of some interventions may thus be inflated if based on mobile subscriptions. Explanation of the <i>South x South Collaboration Lab</i>, including how it operates and how it is funded, would help readers better understand the priority settings. 			
<p>6. Analysis of trade-offs and synergies across the CGIAR Impact Areas; ex-ante assessment of project benefits provides logical rationale for scaling of impacts</p>	<p>Projection of benefits 2.7, Result framework 6.1, Impact statements 5, Innovation Packages and Scaling Readiness Plan 4.1</p>	<p>Effectiveness, Credibility</p>	<p>2</p>
<p>Projection of benefits (2.7): It was great to see synergies identified between the 19 Initiatives ahead of time. However, the anticipated impact of the digital technologies in each of the other thematic Initiatives could be more transparently defined. This includes <i>how</i> digital technologies will be effectively used for public good. In the table at the top of p. 10, the metric for gender equality and youth and social inclusion should also include the number of youths and marginalized members of society who have been effectively engaged. Youths could also include men, not just women (section 2.7.3). The proponents need to consider <i>inclusivity</i> (youths and marginalized members of society) in addition to gender equality (number of women). The high number defined in the gender targets may be too ambitious as it assumes that the attraction rates beyond the first 40% share will be just as easy to attain. This is more likely to be the lower hanging fruit, and the ease of achieving further targets may come up against harder to address barriers, such as infrastructure and social norms.</p> <p>Scaling readiness (4.1): This seems vague. The only real pathway for scaling is ostensibly the collaboration with five digital extension and financial innovators.</p> <p>Impact statements (section 5): The climate adaptation and mitigation text focusses only on extreme events (flood and drought). In practice, seasonal climatic forecasts (months in duration) will be more useful to end-users more often compared with forecasts of extreme events (which would be less frequent).</p> <p>Trade-offs and synergies across CGIAR Impact Areas: Trade-offs and synergies across biophysical, social, and economic indicators are considered (e.g., "investigating trade-offs in different land and water management scenarios related to multiple objectives, including food security, resilience, environmental health, and biodiversity" in 2.7.5), though this may be strengthened if the Initiative team also listed trade-offs between CGIAR Impact Areas?</p>			
<p>7. Evidence that the Initiative will likely lead to impact at scale through integrated systems approaches that drive innovation in research and partnerships, including linking to and leveraging of other Initiatives within and outside CGIAR</p>	<p>Projection of benefits 2.7, Work Package research plans and ToCs 3.2</p>	<p>Effectiveness, Credibility, Relevance,</p>	<p>2</p>
<p>The approach across of coordinating across CGIAR Initiatives is very good and will help maximize learning and use of funding within the organization. However, we are doubtful as to how much impact can be made with the dedicated support put aside for this (\$750,000) given the broad range of initiatives and partners it seeks to support.</p> <p>Impact at scale. The projected number of people benefitting from DX1 may be overestimated. The number of people purported to benefit in section 2.7 is over 6 million. In contrast, the measurable 3-year (End of Initiative) outcomes in section 2.2 are only 1,000 food system actors and 10-20 agrifood research-for-development organizations. As the measurable values in section 2.2 are considerably lower, it would be good to see some more justification (and calculations in an appendix) and the assumptions made to calculate the benefits at scale. Scaling will be critically important for realizing the main benefits of this Initiative; the authors should more carefully elucidate how scaling will be enabled.</p> <p>2.7.1. Projected benefits for nutrition, health, and food security: access to technology (mobile phones, internet coverage, etc.) is not enough. This needs to be coupled with appropriate scientific advice (through extension and</p>			

Criteria	Proposal Sections	QoR4D Elements	Consensus Score
<p>training) and must consider cost: benefit. Will the benefits resulting from access to these new technologies AND agrifood systems science outweigh the cost to individuals?</p> <p>2.7.4. Climate adaptation and mitigation: flood risk seems to be the only climate adaptation considered. What about drought, heat waves, bushfires, seasonal rainfall, soil moisture, river levels, etc. to which digital technologies could provide advance warning or more timely real time information for? This section should go beyond just flood risk.</p>			
<p>9. Research design and proposed implementation demonstrates gender and social inclusion that can be tracked in outcomes</p>	<p>Gender equality, youth & social inclusion 5.3, Projection of benefits 2.7</p>	<p>Legitimacy, Effectiveness</p>	<p>2</p>
<p>Section 5.2 contains the first full description of the term triangulation (meaning North-South-South collaboration). This definition, together with a better explanation of north and south, should appear much earlier in the proposal.</p> <p>Section 5.3: does the focus on women mean that men who want to engage with this Initiative will be deliberately excluded or avoided? If yes, this would mean that much impact stands to be lost? Would it be better to prioritize women and youth, etc. but then allow traditional end-users (perhaps more likely men) to be engaged as a tier-2 priority?</p> <p>How do the authors define youths? People younger than 18 years of age? What happens if a large group of the target audience are, e.g., 19-20 years old? Will these people be excluded from codesign?</p> <p>The gender equality section has a clear focus on women, less so youths and even less marginalized members of society. How will the rural poor and peasant farmers be engaged and what metrics are there for social inclusion?</p> <p>The target of 40% is high: the geographic spread (not 100% clear) will be determining factor in achieving this, as some geographies will have more challenging gendered social norms than others (for example, South Asia).</p>			
<p>10. A risk framework that details main project risks and mitigation actions across areas of science, funding, operations, partnerships, ethics, and environment</p>	<p>Risk assessment 7.3</p>	<p>Credibility, Legitimacy, Relevance</p>	<p>2</p>
<p>The risk framework is relatively clear and logical. Risk of low participation from women, youth, and marginalized members of society needs to be accounted for, as does the risk that target end-users fail to engage as part of this Initiative (e.g., limited uptake by policymakers, etc.). However, "lack of incentives to collaborate" appears a significant risk. The review team suggests that this is amended to "lack of <i>appropriate</i> incentives to collaborate," as this is not about volume of incentives, but rather context specific incentives, particularly to marginalized groups for whom previously offered mass incentives have not appealed.</p>			
<p>11. CGIAR capacity and its comparative advantage and appropriateness to lead the work is justified. This includes the skills, diversity and multi-/trans-disciplinarity of the research team and approaches to meeting gender and diversity targets</p>	<p>Comparative advantage 2.5, Initiative team 9.1, Gender, diversity and inclusion in the workplace 9.2</p>	<p>Relevance, Legitimacy, Effectiveness</p>	<p>3</p>
<p>Initiative team: The type of work described in DX1 is not contained within a single government or private actor (lacking the agency, resources, and/or incentive to do so). Instead, the proposed work requires people who can encourage, convene, and connect these actors with resources while maintaining a focus on the public interest. If anyone can accomplish these goals, it is CGIAR staff. The team should, however, clearly identify which frameworks they proposed building off, and carefully define the scope at the outset in order to define sub-programs that are indeed specific, measurable and achievable. Could the team also include communications/extension staff and regional extension officers to help inform and validate the research? It will also be important to ensure that the voices of women/other marginalized groups are brought forth through the engagement of local organizations who work directly with DX1 proponents.</p> <p>Comparative advantage (2.5): while the justification for CGIAR's comparative advantage is present, this section lacks supporting evidence. Statements such as "CGIAR's strong technical capabilities" and "CGIAR's earned reputation" will not convince readers without supporting evidence (citations, URLs, papers, websites, etc.).</p>			
<p>12. Capacity building within project teams, partners, and stakeholders captured in capacity development plan. This can include development of early career researchers and partner staff, support/empowerment</p>	<p>Capacity development 9.3</p>	<p>Credibility, Legitimacy</p>	<p>2</p>

Criteria	Proposal Sections	QoR4D Elements	Consensus Score
for under-represented stakeholders, and building partner networks			
<p>It may be worth adding a mid-Initiative awareness session on CGIAR's values, code of conduct around gender/marginalized group. The lessons of such sessions when conducted during the busy startup phase of projects are often lost; reaffirming this guidance could be more useful.</p> <p>Further detail on the capacity building of junior Initiative team members, together with partners and stakeholders could be added.</p> <p>Could the Initiative team specify which inclusive leadership training courses are planned and why it is sufficient/appropriate for the roles they have—and across the geographies they plan to work in? It is great that these training programs are mentioned, although details of what is offered are lacking (there is also wide variation in the quality of these programs). A list of training courses with end-users on all planned products from DX1 would be helpful, as would a table outlining who, how, when and where capacity building activities are planned.</p>			
13. Project management mechanisms and (if applicable) additional scientific oversight and governance measures effectively and efficiently support the Initiative objectives	Management plan and Risk assessment 7, Research governance 8.1	Legitimacy, Credibility	2
<p>Management (section 7): The review team encourages the Initiative team to add or elicit geographical representation in the Integrated Management Team (IMT) from the target countries. If all of the IMT are located in one region, significant opportunities for input and leadership from local expertise as part of the IMT may be missed. As per criteria 10, project management of "lack of incentives to collaborate" should closely monitor the effectiveness of the incentives offered, adjust accordingly on an iterative basis. The second greatest risk "biased data, unreliable information, and inaccurate predictive analytics results" is appropriate when considering implicit mitigation strategies.</p> <p>Governance (8.1): Reviewers suggest that a better linkage with the IMT in this section. The IMT structure outlined in section 7 appears to have an IMT per Work Package, wherein one person from each team will coordinate. Should there be an individual whose sole responsibility is oversight across these teams? Would coordination be challenging without such a single point dependency that ensures and incentivizes collaboration and consistency across work packages?</p>			
14. Justified and transparent costing explicitly linked to expected Research for Development results	Financial Resources 10	Legitimacy, Effectiveness	0
<p>No justification for the budget is given, even though the budget values are large and presumably high risk. For instance, how do the authors justify spending \$4.7M in Bangladesh but only \$0.48M in Mexico? Without adequate supporting details of assumptions and spending it is difficult to understand how this budget was derived.</p> <p>Reviewers strongly encourage the Initiative team to consider adding a benefit: cost ratio of the target spending, from the national level down to individual research themes, proposed innovations and developments in digital infrastructure.</p>			
15. Anticipated research outputs (knowledge, technical, or institutional advances, specific technologies or products, policy analyses) are described and knowledge/gaps they will fill are evident. Protocols for open-data and open-access compliance are evident in plan (including budget)	Work Package research plans and ToCs 3.2, Open and FAIR data assets 8.2	Credibility, Effectiveness	2
<p>FAIR datasets: this section is specific and well considered. Protocols for open-data and -access compliance are evident in in section 8.2; this could be detailed in the budget.</p> <p>Work Package research plans: see comments for section 4 above.</p>			
16. Monitoring, evaluation & learning (MEL) plan for the Initiative is clearly defined, with flexibility to adapt. MEL plan supports effective management and learning, including baseline data collection, and evaluative and review processes corresponding to stage-gates and course-correction decisions. MEL occurs during the life	MELIA plan 6.2, Planned MELIA studies and activities 6.3, Measurable three-year (End of	Credibility, Effectiveness, Legitimacy	2

Criteria	Proposal Sections	QoR4D Elements	Consensus Score
of Initiative and is used proactively to reflect on and adapt the Theory of Change, where appropriate	Initiative) outcomes 2.2		
<p>MELIA Plan (section 6.2): Overall, the MELIA plan seems reasonable; it is pleasing to see that agile frameworks, iterative development, and flexibility are mentioned in this section. It is however difficult to judge how and where baselines will be collected (p. 54). "Causal impact evaluation" seems to be relegated to the last year (p. 54), perhaps this would be better incorporated throughout the implementation process and included in design. The South x South collaboration is mentioned again here, though some readers may not know what this term refers to.</p> <p>Measurable End of Initiative outcomes: While the aspirations of the digital divide challenge area are relatively clear, the End of Initiative outcomes focus more on strengthening digital agrifood systems, rather than the digital divide per se, i.e., the people who have limited or nil access to mobile internet, digital services, etc.</p>			
<p>17. Well-defined plan for Initiative-level evaluation and impact assessment based on expected end-of-Initiative outcomes and impact. Links between the impact assessment plan and indicators in the Theory of Change are clear</p>	<p>MELIA plan 6.2, Planned MELIA studies and activities 6.3, Full Initiative ToC 3.1, Work Package ToCs 3.2, Projection of benefits 2.7</p>	<p>Effectiveness, Relevance</p>	<p>2</p>
<p>Overall, proposed plans for Initiative-level evaluation are good. Reviewers particularly like the tables in section 6.1.3 (starting p. 46) with the systematic explanation of result, indicator, location, target audience, and timing.</p> <p>Section 6.2.2 (page 52) suggests that the Initiative Impact assessment is embedded in Work Package 1, but in the said Work Package, there is no mention of impact assessment. If these aspects could be added to the Work Package 1 description, including the addition of relevant detail to the Figure on p. 16 (section 3.2.1), it would help readers better understand how impact assessment fits within this Work Package. Another question under 6.2 is whether the gap analysis demonstrates gender inclusion regarding the use of digital services?</p> <p>Although monitoring elements in 6.3 are reasonable, it is unclear who will be performing the causal impact assessments in 6.2.2. Causal assessment is more successful when planned and integrated from the outset, rather than being a post-hoc analysis after the work has been completed.</p> <p>Given DX1 goals, it will be difficult for proponents to assess impact without a strong counterfactual that would lend credence to the Initiative impact estimates. During the life of the Initiative, other digital developments and adoption will take place that have nothing to do with DX1 or even CGIAR: the Initiative team needs to be able to segregate this background change from that specifically attributed to DX1. To do this, the Initiative could involve counterfactual/control groups and questions (e.g., does capacity training result in improved digital literacy or not, and how is this measured?). In cases where control/counterfactual groups are not well defined and planned, could the authors highlight examples of other Initiative evaluations that rigorously assess the important but often recalcitrant outcomes they are aiming to change?</p>			

The Figure below represents original scoring from each reviewer and consensus scoring for each criterion. *The consensus score across reviewers may not reflect the mathematical average.* For purposes of the Figure, the QoR4D criteria have been shortened. Please refer to p. 3 for full criteria definitions. Please note that criterion 8 on ethics was not scored.



**Four consensus scores were greater than 0.5 variance from the mathematical average. Please refer to criteria 3, 10, and 11 (all with a variance of 0.7 from the mathematical average) and criterion 14 (variance of 1.0 from the mathematical average) above for rationale of these consensus scores.*

6. HER+: Harnessing Gender and Social Equality for Resilience in Agrifood Systems

Review Summary and Actionable Recommendation

This Initiative proposal, HER+, focuses on crucial barriers to gender equality that limit women's engagement within agrifood systems. The Challenge Statement articulated the knotty interplay of the different forces (women's agency, empowerment, and social norms) and proposes apposite approaches to shifting these different dynamics. This concise structuring runs through the proposal, offering a comprehensive MEL plan to track progress while illustrating links with other CGIAR Initiatives. HER+ provides a welcome complement to other Initiatives by aiming to understand what normative barriers prevent women from seizing economic opportunities, with the promise that this will make targeting women more effective. The plan to assemble existing insights about social norms that adversely affect women's capacity to achieve economic resilience will be very useful. The review team is less optimistic about the prospects for affecting these norms in a meaningful way—and at scale—within the given timeframe. Adverse reactions and efforts to regain lost ground can emerge as a response to improvements of women's economic outcome in one dimension. Work Package 1 and 4 are particularly ambitious, as they aim at affecting mindsets, which would take time to occur and may be difficult to sustain over time, as resistances will be met.

On actionable recommendations, the quality of sections and Work Packages is uneven: in Work Package 4, it is not clear whether outcomes are to be achieved by changing institutions or through the less arduous route of making policies more gender responsive. This needs to be explained in greater depth. Further, and given the purposive selection of target countries, the initial conditions in women's positions and participation in different segments of the local agrifood systems, in the institutions that matter for the governance of food security and nutritional outcomes within each of these systems and of the local climate-induced challenges, will be highly heterogeneous (section 2.4): while this enriches the scope for learning, it also means that scaling challenges will vary across highly diverse contexts. A generic Theory of Change weakness is the lack of attempt to conceptualize (in Theory of Change form) the catalysts for and hurdles to scaling. This could be addressed by explaining the activities envisaged per country, the types of local partners (and what those engagements look like), and what sustainable progress/partnerships will look like.

Given the proposal's emphasis on norms, possible backlash, bureaucratic and other buy-in and local partnerships will all be crucial: the proposal is much too vague about each of these. For a proposal on this topic, research governance should include a robust tracking system to monitor intended and unintended, both positive and negative, consequences of the Initiatives under the research program. A clear Theory of Change that documents and addresses ethically adverse events should be developed *a priori*. An avenue to new knowledge about transformative initiatives would be to generate new evidence on the diverse, cross-country, and cross-Work Package nature of the scaling challenges, even if transformations fail. Long-term revisits and in-depth process evaluations would be key. Further, some specific methodological effort will be required to measure attitudinal changes. Finally, the capacity building section is weak and the overall budget details provided fall well short of what is required to assess the resource requirements of a sizeable research initiative like this.

Overall Strengths of Proposal

This Initiative tackles a crucial issue, that of promoting women's access to technologies and resources, both as a specific objective and as a means to contribute to the adoption of climate resilient practices in AFS. This will be achieved in part by changing norms and attitudes towards women's role by empowering them. HER+ provides a welcome complement to other Initiatives by aiming to understand what normative barriers prevent women from seizing economic opportunities, with the promise that this will make targeting women more effective. The planned effort to assemble existing insights about social norms that adversely affect women's capacity to access economic resilience will be very useful.

Planned MELIA studies and activities. HER+ has an appropriately sequenced and convincing MELIA plan suggesting a variety of methods which include a good balance of qualitative and quantitative assessments, (surveys, interviews, randomized control trials and discussion) with inbuilt points to stop, reflect, and adjust course if necessary, allowing the HER+ team to actively learn throughout the Initiative, and hopefully detail this experience on this important theme.

The Work Package 2 focus on sociotechnical innovation bundles (STIBs) echoes CGIAR specialization strongly and although the diversity of women's initial conditions regarding control and influence over the use of, e.g., climate-smart agricultural technologies is challenging. These are challenges that the team should be particularly well placed to tackle.

Overall Weaknesses of Proposal

The reviewers are less optimistic about the capacity to affect norms in a meaningful way within the timeframe of the Initiative, if at all. Adverse reactions and efforts to regain lost ground can emerge as a response to improvements of women's economic outcome in one dimension. Work Package 1 and 4 are particularly ambitious, as they aim at affecting mindsets. The ambitious changes will take time to happen, as resistances will be met, and achieved progress may be difficult to sustain over time. Several assumptions should have already been thoroughly thought through, especially given the experience in these countries, the workshops held to date and so on. The proposal builds on existing initiatives and impact assessments of these initiatives to date. Of course, some of these risks can't be completely mitigated or there are potentially not well-known risks, but it leaves the impression that the authors are not confident about their contextual knowledge, appropriateness, and feasibility of some of the ideas proposed. Some the assumptions in the Theory of Change should be re-thought to bring them into the ambit of program delivery (for example assumptions 2 and 4 in Work Package 1, assumption 2 in Work Package 2). Planning some long-term revisits would be key. Work Package 4 is not coherent, underdeveloped and needs to be articulated with greater clarity and purpose, including clues about the relevant institutions—within each target country's food system—where women's voices and influence will be enhanced. A specific suggestion is to sharpen the text boxes explaining outputs 4.1 to 4.5 in the Work Package 4 Theory of Change diagram on p. 27.

The proposal does not articulate promising gender transformative approaches or science, technology and innovation to be tested to foster gender equality. The overall proposition relies on successfully identifying STIBs and designing Work Package 3 (social protection programs with complementary program components) in the priority countries. There is, however, no assessment of the current status of these approaches/innovation pipelines to demonstrate that relevant STIBs or programs for Work Package 3 would indeed be developed, tested, and adopted within the indicated project timeframe. While high-quality research outputs will be produced, the viability of the scaling ambitions for the norm-shifting, gender-transformative efforts can be questioned. A generic weakness, in the overall Theory of Change and for each of the four proposed workstreams, is the lack of attempt to systematically theorize or conceptualize (in Theory of Change form) the catalysts for and hurdles to scaling. Scaling challenges are likely to vary across each of the target countries and for each theme addressed under each workstream. An unaccounted avenue for valuable new research knowledge would be to conceptualize and generate new qualitative and quantitative evidence on the diverse, cross-country, and cross-Work Package nature of these scaling challenges. Even if the Initiative turns out to be less successful than expected the value of learning from transformative efforts that fail should not—given the very limited state of current knowledge—be underestimated.

A clear and more convincing strategic plan for capacity strengthening beyond occasional support for MSC and PhD fellowships could be codeveloped with other key stakeholders. Of interest, a capacity strengthening plan jointly developed with relevant institutions in the target countries will ensure that new capacity will fill important gaps and also transfer the relevant skills to the countries/institutions. A strategic approach to institutional capacity strengthening at national and regional levels requires serious attention, commitment, and resources. An important aspect of such an approach would be equitable partnerships with regional and national level institutions, the articulation of which is also missing

Areas of Divergence among Review Team and ISDC Resolution

In addition to the points made above, the ISDC review lead had additional substantial concerns. For an Initiative budget of \$28 million, the proposal is remarkably light on detail and vision in terms of the evolution of the field in the next decade. The Initiative team is strong and has decades of experience and is highly regarded. However, this Initiative feels incremental rather than visionary.

Below, ISDC review lead illustrates these concerns with three key points:

1. The proposal stays within the comfort zone and falls short of being at the frontier of this research area in the next decade. Here are some key examples:
 - The field has moved from gender equality to gender equity.
 - The body of research is moving from an exclusive focus on women to inclusion of men and other intergenerational power issues that affect women's endowments/resources and agency, shaping gender equity at various levels.
 - Further, while the title includes social equality, the proposal quickly moves on to say that this phase of research would focus on only gender equality. While focus and building on successes of CGIAR's gender research in the past is a very valuable approach, this Initiative could seize this opportunity to catalyze an inspiring research agenda for the next decade, firmly building on the lessons learnt to date. For example, while attempting to research all aspects of social equality could be challenging within the timeframe for this phase of the initiative (although the budget ask is rather ambitious), the research team could have considered the next generation of research on intersectionality of gender

- with other multiple social identities and how they shape risks and opportunities (a research agenda that is nascent, but fast gaining traction).
2. Often the proposal lacks any scientific rationale for the focus of proposed work and the hyperlinks to websites with plethora of document links are unhelpful; the reviewers should not be expected to wade through to try and understand the rationale for the Work Packages proposed. A few examples are below:
 - o While the framework (figure 1) and the immediate text below alludes to climate change, pandemics, and conflict, the rest of the proposal focuses only on climate change and gender without any rationale.
 - o Even within climate change, the focus is on economic resilience only. Is the assumption that economic resilience to climate change leads to other impacts that are claimed (and not gender disaggregated gender-sensitive in table 1; most rows and columns just say "people")? There is no scientific rationale (and many studies are done by the CGIAR scientists) to claim that interventions solely focused on economics (economic growth, poverty reduction, or economic resilience) without intentionality and operationalization of that intentionality into programs and policies would result in change in other impact areas (e.g., nutrition or equity).
 - o Why cash or food transfers in Work Package 3? What is the rationale for social protection programs to feature so centrally in a research agenda that is about how to improve gender equality in agriculture food systems to have meaningful impacts on CGIAR's five Impact Areas? While the proposal alludes to complementary interventions, it is not clear what these interventions might be (and why these) how and why cash/food transfers are needed. Given that many studies were already done by CGIAR on impact of social protection programmes, what specific lessons will be taken forward from this research and how?
 3. The gender jargon is neither helpful nor intuitive. Why are certain activities, interventions, or approaches gender transformative versus empowering women versus gender-responsive? Clear definitions and rationale would improve this ambiguity. Overall, the proposal leads one to think that this research will be executed employing usual top-down research approaches (despite the use of the term "participatory") than aspiring towards research processes that are empowering and transformative across various groups of people and communities of practice this Initiative aims to work with.

While there is much potential in this proposal, a hope is that a more forward-thinking research agenda (building on past successes), with a strong impact theory, could be fostered in the future iterations of this Initiative.

Does the Initiative Align with the Cohesion of the Portfolio as Described in the Companion Document?

From what is presented, this important Initiative is well aligned and integrated with the CGIAR Impact Areas, the SDGs and other CGIAR Initiatives.

External cohesion: While HER+ has demonstrated some consultative processes in designing this Initiative, the depth and nature of these is unclear. External coherence could be addressed by articulating how HER+ is positioned and complements other Initiatives at global and regional levels. It also does not articulate its comparative advantage vis-a-vis other global players on gender or climate change research and in STIB development and delivery.

Interconnectedness of thematic and regional integrated Initiatives: There is some evidence presented regarding various initiatives HER+ would work with. The recognition of interdependencies across initiatives is currently weak. As commented above, it must be deliberate and regularly reviewed during the implementation.

The reviewers scored each of the 17 QoR4D criterion individually. The review teams then built a consensus score. In the majority of reviews, the consensus score was the same as the mathematical average of the three individual reviewer scores. The criteria that received between a 0 or 1 are presented with a rationale below. Although not required, in order to keep review reports brief and readable, some teams opted to record in their reports their consensus-built rationale for scores above 1. Following the criteria is a figure presenting the individual reviewer scores and consensus. Please refer to p. 3 for the Likert scoring definitions.

Criteria	Proposal Sections	QoR4D Elements	Consensus Score
1. Clearly defined research problem that addresses Impact Areas, is a high priority in the targeted geographies, is well aligned to shared, multi-funder priorities, and is well informed by previous research findings and evaluations	Challenge statement 2.1, Learning from prior evaluations and Impact Assessments 2.3, Impact statements 5	Relevance, Effectiveness	3

Criteria	Proposal Sections	QoR4D Elements	Consensus Score
<p>The research problem is clearly defined and suggests a robust understanding of the many competing challenges that women face, and importantly, their interlinkages. While the agenda is high priority in the targeted geographies, it is also very ambitious. The knowledge gaps identified appear to be well informed by prior evaluations and impact assessments and deep-rooted norms made evident in past research have been distinctly articulated. Figure 1 illustrates this interplay well. Alignment with multi-funder priorities is quite clear, raised, by the UNFSS. The contextual variation across the target countries is intentional and reflects purposive sampling. It is not always clear how this contextual diversity will inform the research agenda and how it will and can be effectively managed. On granularity, the food systems terminology is often too general to provide sufficiently clear directions about depth and specific sub-research foci and what to expect about the choice of conceptual and methodological approaches.</p>			
<p>2. Evidence that the Initiative is demand driven through codesign with key stakeholders and partners (Investment Advisory Groups, governments, private sector, funders) and research collaborators within and outside CGIAR</p>	<p>Participatory design process 2.6, Challenge statement 2.1, Work Package Theory of Changes 3.2</p>	<p>Relevance, Effectiveness</p>	<p>3</p>
<p>An extensive consultation process mobilizing donors, research and implementation partners, and other CGIAR Initiatives (17 of them) has informed the design of this Initiative. This is true for each of the four Work Packages. The Initiative team have thus made good use of international fora to understand recurring issues and constraints confronting women in AFS. This may be expected to happen (be implemented) through partners (unclear), but it would be great to see some pulse-checks and consultations with local partners, who can speak directly about community experiences and developments on the ground. One word of caution is that the growing emphasis on decolonization and participation provides a reminder of the need to ensure consultative and participatory processes that are regularly revisited and updated to match current expectations.</p>			
<p>3. Research questions, objectives, outputs, and outcomes are aligned to the research problem, and are measurable with defined deliverables</p>	<p>Work Package Theory of Changes 3.2, Measurable three-year (End of Initiative) outcomes 2.2, Priority-setting 2.4, Management plan 7.1</p>	<p>Relevance, Effectiveness</p>	<p>2</p>
<p>Research questions, objectives, outputs, and outcomes are coherent with the research problem, and deliverables are well defined. However, measurement of outcomes is tricky for some of the Work Packages (1 and 4 in particular) given the type of impact expected (changing social norms and the strengthening of women's voice and agency). The proposed counting approach is appealingly simple but lacks depth: a general challenge is to measure not only how many people are aware of normative constraints but also whether this is likely to imply meaningful improvements in terms of women's empowerment. With outcome 1.2 (from other workplans), will there be some form of assessment or scale of whether some norms are more malleable than others? Given how deep-seated most gender social norms are, it is important to understand whether some can be more easily shifted than others—even with a broad timescale attached.</p> <p>Another concern is how credible evidence on options to transform gender norms is. The cited literature appears to cover efforts that are intensive in time and resource use with highly localized impacts. The lack of micro-detail is a recurring weakness: For the last Work Package, the Initiative outcome objective states that government and other organization "in at least three LMICs use HER+ learning and guidance to better understand how social innovations, organizational strategies, and government and private-sector policies can effectively increase the voice and agency of women in AFS governance, and their resilience to CC." This is too vague to be meaningful.</p>			
<p>4. Overall Theory of Change with intended outputs, outcomes, and impacts at scale clearly described. Assumptions are documented, causal linkages are clear, especially the role of partners in driving impact</p>	<p>Full Initiative Theory of Change 3.1</p>	<p>Effectiveness, Relevance</p>	<p>2</p>
<p>The overall Theory of Change suggests that improving knowledge in the four directions (Transform, Empower, Protect, Voice) will contribute to mainstreaming gender-equitable resilience into AFS, thereby empowering women to take more active in decision making in AFS: this will lead to improvements in all five impact areas of interest when implemented at scale in the period following the three years of this Initiative. The assumptions behind the first step (reaching women empowerment) are spelled out and mainly linked to uptake by</p>			

Criteria	Proposal Sections	QoR4D Elements	Consensus Score
<p>stakeholders. What is needed for the second, major, step that links empowerment to progress towards the SDGs is less clear, maybe because it will matter more for the period 2024–2030. Given the time it takes to shift GTS, it would be great for Figure 3.1 to illustrate “the longer-term outcomes expected by 2030” showing how these work plans are sowing the seed for ongoing change, not just that between 2022–2024.</p>			
<p>4.a Individual work package Theory of Changes (score individually)</p> <p>Work Package 1 Work Package 2 Work Package 3 Work Package 4</p>	<p>Work Package Theory of Changes 3.2</p>	<p>Effectiveness, Relevance</p>	<p>WP1: 2 WP2: 3 WP3: 3 WP4: 2</p>
<p>Each of the Work Package details the relevant Theory of Change. Graphs are complete and assumptions are well described. The synergies across Work Packages and with other CGIAR Initiatives are clearly highlighted.</p> <p>Work Package 1: The gender transformative work is ambitious but then narrowed down to “<i>norms that block women’s access to financial services and entrepreneurship opportunities.</i>” Is this the plan? The diagram is dense and challenging to read: the narrative was more straightforward/explanatory. Is there a way to include local actors in 2.2 (Measurable 3-year outcomes)? They are key for the learning/iterative journey, but in the Theory of Change civil society organizations appear mostly in the latter stages (outcomes). Assumption 4 (speed of attitudinal changes) and assumption 5 (GTAs bring visible gains in terms of productivity, nutrition, and income security) are very strong. While identifying GTA options in the first phase which can then be implemented between 2024 and 2030 makes good sense, can we be confident that the resources required to generate these changes are available and can be mobilized also at scale?</p> <p>Work Package 2: This was difficult to understand—could the structuring of the figure (p. 19) illustrating Work Package 2 better show the temporal aspects? It was hard to understand the interplay and sequencing. The narrative is clear, but as it stands the figure is a difficult overview to follow. It would be great to see this Work Package dig into specific constraints that impact women’s uptake of technology (literacy, affordability, etc.) not always addressed under norms but deeply influenced by them: while sociotechnical innovation bundles (STIBs) echo CGIAR specialization quite strongly, a concern here is the diversity of women’s initial conditions with regard to say and influence over the use of, e.g., climate-smart agricultural technologies. Where and how does fragility and conflict come in?</p> <p>Work Package 3: Not clear enough what types of social protection will be considered. Is this, and again, mainly about cash transfers and credit? Great that the dissemination of findings is not one dimensional but focusses on packaging this data to different audiences and making it context specific for different stakeholders.</p> <p>Work Package 4. This Work Package does not mention working with specific government partners—are these relationships and agreements existing as per Work Package 3? Alongside outcome 4.1 (toolkits) it would be good to include and engagement or dissemination strategy. Toolkits can be great, but important to get stakeholders using them, which can take some support or nudging. On p. 27 it becomes clear that the focus will be on strengthening women’s voice and agency in climate related AFS governance at the community level. This seems to conflict with, e.g., the Gender Equity in AFS Governance Index to track women’s voice and agency in climate-related governance which also and presumably, involve other institutions than community-level ones? The body of work that will address the governance and policy challenges is not articulated with the necessary clarity for pathways 1 and 2 in Work Package 4. For a proposal with a budget of this size, outcomes 4.1 to 4.3 do not meet the clarity and rigor standards it would be reasonable to expect.</p>			
<p>5. Research methodology and methods (and supporting activities) are fit-for-purpose, feasible, and assumptions and risks are clearly stated</p>	<p>Work Package Theory of Changes 3.2, Priority-setting 2.4, Innovation Packages and Scaling Readiness Plan 4.1</p>	<p>Credibility, Relevance, Effectiveness</p>	<p>3</p>
<p>The array of methods is complete and fit for purpose. Assumptions and risks are well documented, as are priority areas—again well illustrated in Figure 1. One assumption that underpins other assumptions, is that stakeholders will engage, work alongside, use data, collect data, etc., without support or handholding from CGIAR. These partners will have their own agendas, and it is important to not assume that the assigned inputs will happen organically. It will likely take time and resources to ensure the desired outcomes. For example, “<i>EoI outcome above is that (A1) Stakeholders will engage with promising social innovations, organizational strategies, and policies for expanding women’s voice and agency.</i>”</p>			

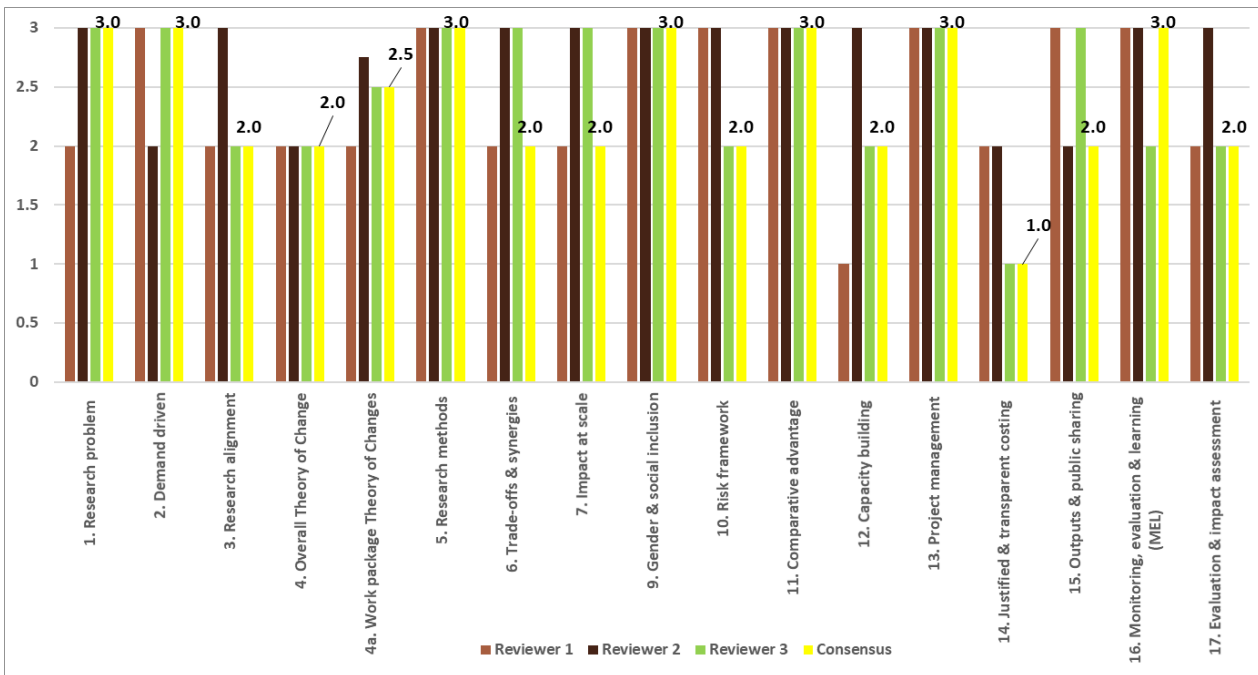
Criteria	Proposal Sections	QoR4D Elements	Consensus Score
6. Analysis of trade-offs and synergies across the CGIAR Impact Areas; ex-ante assessment of project benefits provides logical rationale for scaling of impacts	Projection of benefits 2.7, Result framework 6.1, Impact statements 5, Innovation Packages and Scaling Readiness Plan 4.1	Effectiveness, Credibility	2
<p>The <i>ex-ante</i> assessment of project benefits is particularly transparent and seems rather conservative. Synergies across impact areas are clear. No trade-offs are underlined.</p> <p>No major trade-offs noted—this appears to be a proposal where the Work Packages are very integrated and interlinked, with clear ‘pause points’ to assess scalability. Designed well for purpose.</p> <p>The Results Framework is described in considerable and useful detail. A generic weakness, in the overall Theory of Change and for each of the four proposed workstreams, is the lack of attempt to systematically theorize or conceptualize (in Theory of Change form). Consider and engage with the catalysts for and hurdles to scaling. Scaling challenges are likely to vary across each of the target countries and for the themes addressed under each workstream. A so far unaccounted for avenue for generating valuable new research knowledge would be to conceptualize and generate new qualitative and quantitative evidence on the diverse, cross-country, and cross-Work Package nature of these scaling challenges. For such new knowledge generation, carefully designed process evaluations would be key.</p>			
7. Evidence that the Initiative will likely lead to impact at scale through integrated systems approaches that drive innovation in research and partnerships, including linking to and leveraging of other Initiatives within and outside CGIAR	Projection of benefits 2.7, Work Package research plans and Theory of Changes 3.2	Effectiveness, Credibility, Relevance,	2
<p>Links and collaboration with other CGIAR Initiatives are described in the Theories of Change. They are numerous and clearly relevant. The Initiative will rely on an integrated system approach, leveraging on these other initiatives to achieve its impact at scale.</p> <p>While the risks are significant, there is real scope for both the impact, and ongoing gains from this Initiative. There is robust integration with other CGIAR Initiatives, and this work could play an important role in future gender inequalities/AFS programming. The range of partners planned for Work Package 3 (social protections)—governments and development NGOs is promising—it would be good to see (it is not clear from the proposal) whether similar partners will be engaged for policy and governance work detailed in Work Package 4.</p> <p>It is quite possible that CGIAR teams work with an in-built understanding of how scaling can be achieved based on the in-house experience in working with relevant stakeholders in charge of implementation and scaling efforts in the countries of interest. In the absence of a clearer articulation of how scaling is expected to work and of the nature of the challenges that scaling of the work under the four Work Packages may encounter, how much confidence can we have in the scaling ambitions?</p>			
9. Research design and proposed implementation demonstrates gender and social inclusion that can be tracked in outcomes	Gender equality, youth & social inclusion 5.3, Projection of benefits 2.7	Legitimacy, Effectiveness	3
<p>This is a robust proposal with potential to offer significant learning and positive outcomes for gender equality. Gender inclusion is the central objective, and all Work Packages should contribute. Although there might be spill overs for work on youth and social inclusion, this is not the focus here. Outcomes are coherent with the objective, but some are difficult to measure (e.g., outcome 2.1 on a “better understanding” of the role of women by stakeholders). Even the actual use of evidence produced by the Initiative by the various scaling partners is going to be difficult to assess and counting those who claim to do this is likely to be a (very) poor proxy of actual impact. Given that the first few years will entail intensive learning with much expected back and forth on outcomes to focus on. The Results Framework would thus be expected to change significantly during the initial program phase.</p>			
10. A risk framework that details main project risks and mitigation actions across areas of science, funding, operations, partnerships, ethics, and environment	Risk assessment 7.3	Credibility, Legitimacy, Relevance	2

Criteria	Proposal Sections	QoR4D Elements	Consensus Score
<p>The Risk Framework seems comprehensive and well thought through. Risk 1 (changing norms and values at organizational level can take time and meet resistance) is major. Maybe it would be necessary to consider more explicitly the risk linked to the fact that individuals (men): 1) oppose changes that weaken their relative power or 2) develop new strategies to compensate the potential gains of women induced by gender sensitive policy interventions. Indeed, it is unlikely that they'll let go of their domination without reactions.</p> <p>Another significant threat is that gender and social norms may limit the uptake of system-level changes with partners (for all Work Packages). While donors are keen to work on GTS, they are put off by then speed in which these norms change, and which does not fit tidily in to a three- to four-year project cycle. This can be somewhat mitigated by expectation setting, informing the donor upfront, and keeping the donor in the loop— but remains a risk. Great to see the proposal noting that working with partners already focused on GTS can alleviate partner drop off. A perhaps underestimated risk relates to reading too much into the evidence on Initiatives to shift norms to date and the intensity of effort required and scale at which such initiatives have succeeded: yes, norms do change, but this often takes more time than that available here and through less social engineering than HER+ aspires to.</p>			
<p>11. CGIAR capacity and its comparative advantage and appropriateness to lead the work is justified. This includes the skills, diversity and multi-/trans-disciplinarity of the research team and approaches to meeting gender and diversity targets</p>	<p>Comparative advantage 2.5, Initiative team 9.1, Gender, diversity and inclusion in the workplace 9.2</p>	<p>Relevance, Legitimacy, Effectiveness</p>	<p>3</p>
<p>The Initiative requires a large array of skills both in agronomy and social sciences, and the CGIAR team covers this broad spectrum. Probably the best positioned organization to assemble such a team. The cohesion and depth of this proposal offers much confidence in the quality of the team who will be supporting/carrying out the work. Also includes some strong research partnerships with academia, and a diverse team. Good to read there is an intention to bring men into the team also.</p>			
<p>12. Capacity building within project teams, partners, and stakeholders captured in capacity development plan. This can include development of early career researchers and partner staff, support/empowerment for under-represented stakeholders, and building partner networks</p>	<p>Capacity development 9.3</p>	<p>Credibility, Legitimacy</p>	<p>2</p>
<p>The Initiative seems committed to capacity development, although the corresponding plan is not very detailed and concentrated on gender and inclusion topics. More general capacity building in management and/or research skills could also be embedded in the activities. Further and what are the specific plans and prospects for capacity building among local partners in the target countries? Has sustainability planning for partner networks been considered? Good opportunity to create gender champions of non-gender focused staff and consultants, and useful in the ongoing mainstreaming of gender.</p>			
<p>13. Project management mechanisms and (if applicable) additional scientific oversight and governance measures effectively and efficiently support the Initiative objectives</p>	<p>Management plan and Risk assessment 7, Research governance 8.1</p>	<p>Legitimacy, Credibility</p>	<p>3</p>
<p>The timeline appears well sequenced and spaced over the three-year period. MELIA takes place throughout the Initiative, with several points for assessment and reflection. It might be useful (under MELIA) to factor in some stakeholder consultations throughout this three-year period, to ensure structured check-ins and pulse-checks with key partners take place. The transformative agenda is welcome, but also very ambitious: it seeks to tackle steep and highly diverse challenges within a relatively short time window. While having confidence in the experience and capacity of the individuals in charge, it is also important to not underestimate the challenges of managing this well.</p>			
<p>14. Justified and transparent costing explicitly linked to expected Research for Development results</p>	<p>Financial Resources 10</p>	<p>Legitimacy, Effectiveness</p>	<p>1</p>
<p>The budget template is inadequate for anyone to judge the relevance of the budget to the objectives. Clear—but very high level, and so hard to say what is justified. Would be good to see detail on the Work Package/geography cross over, i.e., what is the budget for Work Package 1 in Nigeria? Given the information provided, it is difficult to provide deep commentary on the appropriateness of the individual budget lines.</p>			

Criteria	Proposal Sections	QoR4D Elements	Consensus Score
15. Anticipated research outputs (knowledge, technical, or institutional advances, specific technologies or products, policy analyses) are described and knowledge/gaps they will fill are evident. Protocols for open-data and open-access compliance are evident in plan (including budget)	Work Package research plans and Theory of Changes 3.2, Open and FAIR data assets 8.2	Credibility, Effectiveness	3
<p>Protocols for open-data and -access compliance are described in section 8.2, with pledge adherence to the Open and FAIR Data Assets Policy. Research outputs are described and the knowledge gaps they fill are evident in the way they motivate the entire initiative. As noted previously, it would be good to have “depth/severity” level in output 1.2 (qualitative assessments)—some are going to be much easier to shift than others, and this could also prove a useful tool in keeping donors informed and engaged. In Work Package 2, much new evidence will likely be produced through research. It might be a good opportunity to encourage more junior, potential gender champions to get very involved on the subject matter and contribute to/publish some of this research. On 3.2, new research findings on how social protection design and complementary programs can help women is very useful. The design of the product (how it is perceived/experienced) is as important as the content of the product itself. It would also be very useful to detail how these insights were attained, how design could also act as a constraint. There should be some in-person and tailored support 4.1 (toolkit) and 4.5 (guide) to ensure that people understand a) how they are relevant for their specific work and b) how to use them. Otherwise, they may end up under-used or used at all.</p>			
16. Monitoring, evaluation & learning (MEL) plan for the Initiative is clearly defined, with flexibility to adapt. MEL plan supports effective management and learning, including baseline data collection, and evaluative and review processes corresponding to stage-gates and course-correction decisions. MEL occurs during the life of Initiative and is used proactively to reflect on and adapt the Theory of Change, where appropriate	MELIA plan 6.2, Planned MELIA studies and activities 6.3, Measurable three-year (End of Initiative) outcomes 2.2	Credibility, Effectiveness, Legitimacy	3
<p>Generally, very good, with some early data collection that can inform the initiative and help adapt. Aside from that, a lot of the MELIA activities will be completed by 2024 and by mainly used in the second phase of this initiative (the action phase). Now, the targets are all set in terms of number of occurrences (breadth) rather than anything reflecting depth, which would be more difficult but potentially more useful.</p> <p>It is unclear whether the causal impact evaluations will all be based on randomized control trials (and mixed methods).</p> <p>Clear end of year deliverables, and plenty of systematic pulse-checks and assessments with the MEL timeline. Given the quite considerable risks related to the gender-transformative initiatives—and the hurdles to scaling these may encounter—running thoughtful and well-designed process evaluations alongside the randomized control trials would make a great deal of sense. A general concern—and given the focus on shifting gender norms—is how such changes will be measured. Reported attitudinal changes in small-scale studies are subject to bias concerns and one important contribution of a well-resourced program like this would be to make progress on how to measure social norm changes more robustly and convincingly.</p>			
17. Well-defined plan for Initiative-level evaluation and impact assessment based on expected end-of-Initiative outcomes and impact. Links between the impact assessment plan and indicators in the Theory of Change are clear	MELIA plan 6.2, Planned MELIA studies and activities 6.3, Full Initiative Theory of Change 3.1, Work Package Theory of Changes 3.2, Projection of benefits 2.7	Effectiveness, Relevance	2
<p>There is a clear roadmap for the impact assessment at the Initiative level, guided by the clearly defined expected benefits. It is based on the Work Packages Theories of Change. Now, the Initiative evaluation and learning questions are very broad, two of them aimed at assessing whether the Initiative leads to a “better understanding” of the role of social protection and of social innovations in boosting women’s climate resilience. It is difficult to envision how this will be measured. There is going to be a wealth of learning that emerges from this Initiative. If it is possible to share these insights more broadly, it would be good to ensure that there is a strong plan to package and share the results. For gender-transformative initiatives, effort intensity and commitment are</p>			

Criteria	Proposal Sections	QoR4D Elements	Consensus Score
likely to affect impact: while feasible at a small-scale level, working with a dedicated NGO, impact is likely to be much harder to achieve at scale when working, e.g., with public sector administrators in the country. This should thus not come as a great surprise. While there is some acknowledgment of such concerns in the risk assessment, a systematic and more thoughtful analysis of the scaling challenges would have been helpful.			

The Figure below represents original scoring from each reviewer and consensus scoring for each criterion. *The consensus score across reviewers may not reflect the mathematical average.* For purposes of the Figure, the QoR4D criteria have been shortened. Please refer to p. 3 for full criteria definitions. Please note that criterion 8 on ethics was not scored.



*Five consensus scores were greater than 0.5 variance from the mathematical average. Please refer to criteria 6, 7, 10, 14, and 15 (all with a variance of 0.7 from the mathematical average) above for rationale of these consensus scores.

7. NATURE+: Nature-positive Solutions for Shifting Agrifood Systems to More Resilient

Review Summary and Actionable Recommendation

This Initiative is highly welcome. It is high time for the CGIAR to start building its own capacities and producing knowledge on nature positive solutions (NPSs). However, the proposal needs a more detailed problem statement from which region-specific knowledge gaps/research questions can be derived. Currently, the problem description does not justify the research activities proposed in Work Packages 1-4.

The proposal describes the problems associated with industrial agriculture yet the focus (as it should be) is on smallholder systems. Therefore, the problem statement needs to be reframed to identify the type of problems small holder agriculture faces that will benefit from NPS. Moreover, it is not clear how the research in this proposal integrates with others dealing with agronomy innovations in farming systems (i.e., transformational agroecology, excellence in agronomy).

The use of public/private partnerships to market and sell Nature+ products and generate incomes for smallholders is lacking with an assumption that investors will be state agencies and the exclusion of direct corporate buy-in. The development of business models in Work Package 4 could be added to Work Packages 2 and 3 as this is a strong lever for change that will multiply economic benefits and provide models for Work Package 5.

The development of demonstration areas is an excellent approach to transfer knowledge to the broader community; however, these activities may lead to demonstrations of failure as they are relatively under-resourced compared to other activities. It is suggested that food production be at the center of the model (Fig 1). Then the impacts of climate change, water, biodiversity, etc. are organized around achieving sufficient quality food with a minimum impact on these components, through resource use efficiency, management of trade-offs and optimization of inputs.

There is no evidence on the "many years of research on NPSs" or the "CGIAR leadership" in this area, as claimed in the proposal. On the contrary, CGIAR is well known for the promotion of nature-negative solutions from its creation. It has had many positive outcomes in terms of agricultural productivity, no doubt, and it is also true that there are examples of research that contributed knowledge on biologically mediated solutions. But these are mostly exceptions. Even today, CGIAR continues to push for chemical fertilisers, GMOs, crop hybrids that need high N and P inputs, herbicides, etc. (i.e., what they call "excellence in agronomy"). CGIAR need to build legitimacy. A way to start, is to be modest and honest about the team's capacities to lead the way on NPS, which is really a new area for CGIAR and not yet internalized at system level.

Seeking broader partnerships to gain concrete NPS capacities is highly advisable.

The scorings and comments below are meant to help the Initiative's team build a stronger proposal as there are several weak points, gaps, and overlaps. Yet, the intentions, the dedication and the capacities of the scientific team are not in question. As reviewers, we understand the difficulties associated with bringing together a large consortium of stakeholders and CGIAR scientists from different continents, in the context of the COVID-19 pandemic, and try to build a coherent yet comprehensive proposal.

Overall Strengths of Proposal

The systematic review of CGIAR NPS-relevant innovations that identified several innovations that can be deployed quickly and cost-effectively with partners on the ground in the first cycle (2022–2025).

It is very positive to see the diversity of partners engaged in this consortium. One obvious gap in the consortium are the science and development members of the worldwide agroecology movement. While NPS is a new area for CGIAR, this has been the playground of the agroecology movement for more than 25 years, not only in research but also in practice, advocacy, and policy action.

Work Package 3 will provide on-ground demonstrations of activities that may be implemented more broadly across the region. The restoration activities provide excellent opportunities for resourcing teams locally that may then teach others to implement successful activities. Monitoring of crop and soil biodiversity in these areas will provide empirical data that may be used to "sell" the benefits of Nature+ innovations to the broader community.

Overall Weaknesses of Proposal

The proposal is difficult to follow as the work programs (conserve, manage, restore, recycle, and engage) address the innovation areas (biodiversity, soil, water, and waste) in different ways with considerable overlap between Work Packages.

While biodiversity conservation is clearly important, the rationale for conservation of crop genetic resources providing Nature+ solutions in Work Package 1 is unclear. Conservation without utilization will have no impact by 2025. This Work Package appears to replicate existing resources from other CGIAR programs as many of the focus crops (sorghum, potato, rice, cattle, pigs) have established, global and well-resourced crop improvement programs with extensive germplasm diversity collections and a focus should be on sharing these genetic resources among partners rather than creating lists. The use of existing informal seed systems in Work Package 2 to deliver genetic diversity uses existing systems so it is unclear how this will elicit change. Participatory breeding protocols have been developed for other systems and novel aspects of this Work Package are unclear. The establishment of on-farm demonstrations for orphan crops as part of the Work Package 3 should be better connected to Work Packages 1 and 2. Engagement of industry to codevelop orphan crops is completely lacking. FAO and many other organizations have compiled excellent descriptions of genetic resources that should be the basis of gap analyses. There is no consideration of forest genetic resources.

There is a need to clearly define problems (e.g., problem tree) as much as possible per country/region, with identifiable knowledge gaps, and consequent/justified research needs (in the light of a possible solution tree). Currently, the proposal enumerates the generic, well-known problems associated with industrial agriculture (biodiversity loss, environment, poor nutrition, deforestation, etc.) and simply states that NPSs are all you need to address them. This simplification has two major flaws:

- It describes the symptoms, not the root causes of the problems; hence it is not clear how this action will contribute to solving the real problems
- By focussing their critique on "industrial agriculture" they fail to identify the other problems that render smallholder farming systems also unsustainable.

Areas of Divergence among Review Team and ISDC Resolution

Major average discrepancies from scoring by the three reviewers were found for criteria 1, 2, 3, 4, 9, 11, 12, 15 and 16. Lowest average scores were those of criteria 4a (Work Package 1), 7 (effectiveness) and 12 (capacity development). One reviewer tended to score 3 on most criteria, even those in which he/she was critical of in the narrative evaluation. Although the discrepancies in terms of scores seem sharp between the other two reviewers, the narrative evaluation behind the scorings coincides in most points. Such is also the case when comparing across the three reviews. For example, for criterion 1: while one does not find the problem clearly defined in order to justify the research activities (score 1), the second reviewer finds it well aligned with the UK COP26 priorities but points to deficiencies, e.g., in terms of private sector engagement (score 3); and the third points to missing questions and lack of stock-taking from previous CGIAR Initiatives (score 2). These three assessments are not contradicting each other. They only differ in the scores provided. So, the differences in scoring suggest wider discrepancies than what exists among reviewers. After discussion consensus scores were reached for all except criteria 1, 3, 7, and 11 and these scores were agreed after discussion with the ISDC review lead.

Does the Initiative Align with the Cohesion of the Portfolio as Described in the Companion Document?

Lateral cohesion: This Initiative is part of the overarching results framework that comprises the 32 Initiatives that will be deployed by the CGIAR. Many of the scientists participating in this Initiative will also participate in other CGIAR Initiatives, and several Initiatives will be implemented in exactly the same target regions (and likely with the same households and communities). However, the messages conveyed by these different Initiatives are often contradictory. This will create confusion among the partners and beneficiaries on the ground. If CGIAR wants to take up NPSs as its main approach to agricultural innovation and rural development, then the entire portfolio of 32 Initiatives should follow a nature-positive approach. It is not advisable to promote different practices that may contradict each other (at all, but specially not) in the same locations.

Vertical cohesion: The Initiative is structured on the basis of the Theory of Change framework and hence it aligns well with the CGIAR result framework (cf. Fig. 6 and Annex 1 in Companion Document) which is basically an application of the classical Theory of Change. However, this model works well when Initiatives are top-down and/or unidirectional, moving from activities, to outputs, outcomes, and impacts, assuming causalities between these steps and an if-then logic. This model is however not always applicable—almost never—to co-innovation approaches, as proposed by this Initiative, which are bottom-up, adaptive, and where the role of researchers and development agents is to broker knowledge, to create a dialogue to support self-investment, a sense of ownership and a risk-taking attitude by local actors. Problems are addressed as they emerge, and solutions are developed through trial and error, experimenting together between different actors of a platform. This adaptive dynamic, inherent to managing projects as complex adaptive systems, cannot be captured in the classical Theory of Change structure.

The reviewers scored each of the 17 QoR4D criterion individually. The review teams then built a consensus score. In the majority of reviews, the consensus score was the same as the mathematical average of the three individual reviewer scores. The criteria that received between a 0 or 1 are presented

with a rationale below. Although not required, in order to keep review reports brief and readable, some teams opted to record in their reports their consensus-built rationale for scores above 1. Following the criteria is a figure presenting the individual reviewer scores and consensus. Please refer to p. 3 for the Likert scoring definitions.

Criteria	Proposal Sections	QoR4D Elements	Consensus Score
<p>1. Clearly defined research problem that addresses Impact Areas, is a high priority in the targeted geographies, is well aligned to shared, multi-funder priorities, and is well informed by previous research findings and evaluations</p>	<p>Challenge statement 2.1, Learning from prior evaluations and Impact Assessments 2.3, Impact statements 5</p>	<p>Relevance, Effectiveness</p>	<p>1</p>
<p>The problem is only partly described but is too simplistic and not well characterised. The diagram in Figure 1 shows that the focus of the problem is on the impacts of industrial agriculture. This is not comprehensive enough and fails to consider other problems that cause smallholder systems to be unsustainable. Besides, next to the "problem" diagram of Figure 1, it would be nice to see a "solutions" cloud that underpins the activities proposed by this Initiative.</p> <p>Problems need to be characterized as much as possible per region or at least per "groups of countries" amongst the target ones. The type of problems that require NPSs are not the same in Burkina Faso or Colombia, for example, the latter being an OECD country already.</p> <p>Can the work be informed by lessons learned in various CGIAR programs such those achieved by DIVERSITAS in the past, particularly in terms of the long history of participatory, multisectoral approaches to research? The Initiative should show connections to the other new Initiatives dealing with soil degradation and with agronomy, for example.</p> <p>The proposal of "scaling NPS best practices in technologies, innovations, and tools" appears to be well aligned to funder priorities give, "Ninety-two governments and over 90 high-level corporations pledged urgent action and investment at UKCOP26 to" be Nature+. The Impact statements are well written (though grammar is occasionally poor and a "partner annex" is not available).</p> <p>Fixing this part of the proposal is of high importance: a poor problem definition leads to unjustified research questions and to a weak Theory of Change.</p>			
<p>2. Evidence that the Initiative is demand driven through codesign with key stakeholders and partners (Investment Advisory Groups, governments, private sector, funders) and research collaborators within and outside CGIAR</p>	<p>Participatory design process 2.6, Challenge statement 2.1, Work Package ToCs 3.2</p>	<p>Relevance, Effectiveness</p>	<p>2</p>
<p>There is evidence that the Initiative was built around the UN Food Summit priorities and terminology. There is not enough evidence that the demand came from the target countries. On pgs. 10-12, a consultation procedure is mentioned on which the reviewers cannot say much with the details provided. But consulting national organizations is not enough to assume that there is demand for this Initiative among stakeholders. There is, however, a demand from nature conservation organizations, but since they are also part of this consortium (which is a positive aspect), this cannot be part of the demand. There is no evidence of codesign with national research organizations. What are the differentiated approaches that will derive from these partnerships with UNEP, WWF, and FOLU. Each has an influence and a comparative advantage that needs to be clarified.</p>			
<p>3. Research questions, objectives, outputs, and outcomes are aligned to the research problem, and are measurable with defined deliverables</p>	<p>Work Package ToCs 3.2, Measurable three-year (End of Initiative) outcomes 2.2, Priority-setting 2.4, Management plan 7.1</p>	<p>Relevance, Effectiveness</p>	<p>1</p>
<p>Questions are well developed in the science narrative of each work program, but the outputs are not well quantified or measurable for most of the Work Package. There is a missed opportunity to integrate soil biodiversity with water management activities to measure change from Work Package 4. In general, a problem with scoring this criterion is that the quality of the research proposed depends largely on the problem</p>			

Criteria	Proposal Sections	QoR4D Elements	Consensus Score
<p>identification. Since this is not clear enough (beyond enumerating the well-known “symptoms” associated with industrial agriculture), it is hard to assess to what extent the research questions derive from an actual analysis of the problem, or from the preferences, disciplines and previous (and current) work of the members of the team. For example, the focus on agrobiodiversity in Work Package 1, is it justified from the problem analysis or is it just the type of work normally done by the specific team of Work Package 1? Where, in the problem statement (2.1), do we see any justification that agrobiodiversity is the key to solving the problems at stake?</p> <p>Another concern is the project entry point. The title never mentioned biodiversity but the whole rationale is built upon achieving conservation objectives. Integrating food and biodiversity objectives should be at the core of this proposal.</p> <p>The proposal would benefit from a clearer problem statement that derives in a set of logical research questions to be prioritized and addressed by each Work Package, so that the proposed research within each Work Package is justified. The lateral connectivity between Work Packages needs also to be strengthened.</p>			
<p>4. Overall Theory of Change with intended outputs, outcomes, and impacts at scale clearly described. Assumptions are documented, causal linkages are clear, especially the role of partners in driving impact</p>	<p>Full Initiative ToC 3.1</p>	<p>Effectiveness, Relevance</p>	<p>2</p>
<p>Theory of Change summary is difficult to follow. High-level outcomes are well described in the Theory of Change summary as: a 30% increase in food, land, and water productivity, at least 20% increased income from NPS-focused value chains, reversal of soil degradation, and a 50% increase in systems biodiversity, directly and indirectly benefiting ~69 million people by 2030. Some of the more specific outputs already exist (informal seed systems, participatory breeding protocols, etc.) and benefits from comparisons of how existing systems may be applied in other region is a missed opportunity. The role of partners is less clear with general statements like “willing to fund” or “applying practices.” The production of quality food for improved nutrition is missing in the Theory of Change. Following from the previous criterion, the causal link between the research proposed and the outcomes (and problems) is not always explicit.</p> <p>It is good that the Initiative aims to capitalize on existing knowledge generated through CRPs, for example. However, it would have been interesting to see a summary of these readily applicable NPS innovations in the proposal. NPS “innovations” are presented in annex 3 (reviewers are not supposed to need to access the annexes to score the proposal), which is an Excel file with a long list of NPS techniques. For example, #19: “Increasing water productivity through irrigation and use of water saving technologies, expanding or intensifying irrigation/ /water delivery through canals/groundwater/ Spate Irrigation; Improving irrigation efficiency ; Water saving technologies are any technology that purportedly saves water. They could be technologies like drip and sprinkler irrigation, faucets and sanitary systems with low water consumption, agronomic practices that saves water.”</p> <p>This does not look like an “NPS innovation.” First, because these are not CRP innovations, they’ve existed for more than 50 years now. Second, because this is too generic to assess what the project will deliver, where, and how? Are they going to bring irrigated cropping systems to northern and central Burkina Faso, for example? With what water? How many hectares? Which crops? Drip irrigated? Pumped or gravity? This needs further work and planning to be credible.</p> <p>How will the partners drive impact? This appears to be a weakness.</p>			
<p>4.a Individual Work Package Theories of Change (score individually)</p> <p>Work Package 1 Work Package 2 Work Package 3 Work Package 4</p>	<p>Work Package ToCs 3.2</p>	<p>Effectiveness, Relevance</p>	<p>WP1: 1 WP2: 2 WP3: 2 WP4: 2 WP5: 2</p>
<p>The Work Packages need to be better linked to problem statement and End of Initiative outcomes; especially Work Package 1. The research proposed needs to be better justified. Plenty of what is proposed has already been done under the different CRPs. What’s new here? What’s the added value of this Initiative with respect to previous research?</p> <p>A good example in the research plans is the way the first research question of Work Package 2 is addressed with a concrete set of methods and outputs. It is easy to link up these research activities and outputs with the End of Initiative Outcomes. It would be useful if all Work Packages and research questions were presented in this way; this would strengthen the research plan and Theory of Change.</p>			

Criteria	Proposal Sections	QoR4D Elements	Consensus Score
<p>Work Package 3 demonstration sites are excellent initiatives that are required to quantify benefits of actions. The ability to elicit change by influencing smallholders is questionable if direct linkages to agro-industry are not developed. Restoration is the main pathway. But the expected restoration action is missing. How for instance the Work Package 3 intends to restore land, for what impact, which land is targeted, how restoration in that context will be based on their outputs such as restoration actions with communities. There is a great deal of cost effectiveness and entrepreneurship in relationship with the value chain actors, that alone justify that the restoration include aspect beyond waste recycling.</p> <p>Work Package 1: what is meant by best-bet conservation practices and how are these identified? It is hard to see the youth specificity of the pathway 2. The guidelines and evidence output can be applied to any group. On the narrative side, the traditional practices could be a great entry point for the best-bet practices compendium. However, the aspect of food production comes as an underpinning of the conservation. Food production should be a deliberate outcome, articulated with agrobiodiversity and water conservation. As a matter of coherence with other One CGIAR, it is important to check if there no duplication with another project about knowledge platform for genetic resources.</p> <p>Work Package 4: The issue of markets and governance are central to this Work Package. The term recycling is limiting to express the various dimensions of this Work Package. What do you mean by a catalog of codesigned business models for rural typologies? How do you differentiate the two pathways: Scaling up NPS and Adoption of NPS. These two pathways are linked. A large deal of circularity is already embedded in rural practices. The concept is new, but the practical use of circular system is not new for these traditional systems, particularly in Africa.</p> <p>Work Package 5: This Work Package is well presented, but the narrative should have been first about partnership for scaling up and the process to be established in that regard. Only with a suitable partnership we can envision a better governance at context, address market drivers and funding are part of the responses to the barriers of adoption and scaling up.</p>			
<p>5. Research methodology and methods (and supporting activities) are fit-for-purpose, feasible, and assumptions and risks are clearly stated</p>	<p>Work Package ToCs 3.2, Priority-setting 2.4, Innovation Packages and Scaling Readiness Plan 4.1</p>	<p>Credibility, Relevance, Effectiveness</p>	<p>2</p>
<p>There is no reason to doubt the research capabilities of the Initiative team and their state-of-the-art methodologies. The review team wished to see additional aspects about how sustainable intensification improved food production for nutritious crops. Methods for Work Package 1 and 5 are vague in the Theory of Change, however these are clarified and well summarized in the 7.2 Gantt chart.</p>			
<p>6. Analysis of trade-offs and synergies across the CGIAR Impact Areas; ex-ante assessment of project benefits provides logical rationale for scaling of impacts</p>	<p>Projection of benefits 2.7, Result framework 6.1, Impact statements 5, Innovation Packages and Scaling Readiness Plan 4.1</p>	<p>Effectiveness, Credibility</p>	<p>2</p>
<p>Breadth depth and probability are well defined in 2.7. Given that “the results of NPS co-creation processes are unknown at this stage,” this is well developed in the proposal. The Results Framework is particularly strong at providing a summary of scaling impacts.</p> <p>However, the nutrition, health, and food security aspects (p. 13) are not clearly articulated in this proposal. It was an expected entry point but at the end we see more connections with biodiversity and conservation.</p> <p>On 2.7, it will be important to be more precise about the aspect of livelihoods the project will influence. The breadth of the project is about number of HH members who benefit. What exact benefit can they expect from NATURE+ Impact Areas? In the Depth section, it is mentioned income increase that they MAY have because of increase productivity. The aspect of increased food productivity was quite shallow in the proposal.</p> <p>Another point of attention is that NPSs may sometimes entail more labor. This should not become an extra burden for women or disadvantage groups. Please make this more implicit as a trade-off to be embraced. Also, bioeconomic opportunities are often ripped by the wealthier males in a community.</p>			

Criteria	Proposal Sections	QoR4D Elements	Consensus Score
7. Evidence that the Initiative will likely lead to impact at scale through integrated systems approaches that drive innovation in research and partnerships, including linking to and leveraging of other Initiatives within and outside CGIAR	Projection of benefits 2.7, Work Package research plans and ToCs 3.2	Effectiveness, Credibility, Relevance,	1
<p>The indicators of outcomes and impact seem unrealistic. How can this Initiative in such a short period of time (2025) achieve the outcomes/impact that the CRPs failed to achieve in 10 years? What are the substantial changes in approach and methods brought by this Initiative to substantiate such claims? (Yet, the reviewers agree with the assumptions made for probability estimation in section 2.7—this is not the issue here.)</p> <p>For a CGIAR outsider, it is difficult to see the linkages without delving into the copious annexes/attachments that are supplied. For example, there are several references to agroforestry systems and developing lists of suitable germplasm with no reference to the World Agroforestry Center or FAO databases for suitable germplasm. Much of the work on managing crop diversity is being actively managed by other CGIAR programs. Additionally, there seems to be overlaps with some Genetic Innovation Initiatives (e.g., Genebanks). Are links with such Initiatives expected? Can this Initiative achieve impact at scale if key partners are missing? Similarly, what are the links with Initiatives such as Transformative Agroecology or Excellence in Agronomy? The three Initiatives are likely going to work in the same communities in some countries—would this not lead to conveying contradictory messages, undermining the whole idea of One CGIAR?</p>			
9. Research design and proposed implementation demonstrates gender and social inclusion that can be tracked in outcomes	Gender equality, youth & social inclusion 5.3, Projection of benefits 2.7	Legitimacy, Effectiveness	2
<p>This is addressed, although it seems to be an afterthought with “stakeholders” replaced by “women and men” in many parts of the proposal. Simply counting “the number of women, youth and indigenous people employed in NPS value chains, the number of women in decision making positions, or the number of plans developed with the inclusion of women, youth and indigenous people perspective,” will be meaningless without a baseline of total people involved (it is a percent of total). The change in numbers would be a more appropriate metric, assuming there is an appropriate baseline. The review team expected in this Gender Equality aspect a breath based on the number of women engaged in nature-based production or women and youth higher number of actors across the value chains, not just demography statistics. This will make it more consistent with the Depth section below.</p>			
10. A risk framework that details main project risks and mitigation actions across areas of science, funding, operations, partnerships, ethics, and environment	Risk assessment 7.3	Credibility, Legitimacy, Relevance	2
<p>General risks such as “natural disaster, civil unrest and diseases outbreaks” (Work Packages 1 to 5) and “lack of capacity in the CGIAR or other partners to operationalize such innovative agenda” (Work Packages 1 to 5) with generic responses should be replaced with Work Package specific risks. There are several collateral risks related to pressure on land and resources and mostly those related to competition for water resources. The market risks are not minimal.</p> <p>When proposing to demonstrate $B/C > 1$ for any NPS alternative proposed, the Initiative team should also propose to evaluate B/Cs in the long run, both averages and st.dev.</p>			
11. CGIAR capacity and its comparative advantage and appropriateness to lead the work is justified. This includes the skills, diversity and multi-/trans-disciplinarity of the research team and approaches to meeting gender and diversity targets	Comparative advantage 2.5, Initiative team 9.1, Gender, diversity and inclusion in the workplace 9.2	Relevance, Legitimacy, Effectiveness	1
<p>Capacities to lead work on NPSs and bioeconomy is not something for which CGIAR is well known. It is actually a new theme for CGIAR. Capacity building and teaming up with experienced partners, and certainly with resource people from the agroecology movement, is highly recommended.</p> <p>Lack of a precise problem statement, that outlines the challenges of implementing, e.g., circular bioeconomies, indicates that the team is not well versed in these areas.</p> <p>Legitimacy would be very compromised if this Initiative appears as “a new bag to sell old grain” (i.e., that the ethno-botanist, the modeller, the economist, the gender specialist, etc. all continue to do the work they like to do, but now under the NPS umbrella).</p>			

Criteria	Proposal Sections	QoR4D Elements	Consensus Score
<p>The Initiative team appears to be balanced, however, one geneticist and one soil biodiversity scientist to examine diversity of orphan crops and soil types in the five regions does not appear to be sufficient and teams should be specified.</p>			
<p>12. Capacity building within project teams, partners, and stakeholders captured in capacity development plan. This can include development of early career researchers and partner staff, support/empowerment for under-represented stakeholders, and building partner networks</p>	<p>Capacity development 9.3</p>	<p>Credibility, Legitimacy</p>	<p>1</p>
<p>This section needs further work. This section was written under the assumption that the Initiative team has the capacities to lead this project and train, educate, and develop capacities among stakeholders (a classical CGIAR attitude).</p> <p>It is a good Initiative for CGIAR to embark on NPSs and circular bioeconomy, but they need to be trained first, and/or hire the necessary capacities. They do not have any legitimacy to start training stakeholders on themes that are new to them.</p> <p>Besides, there are key resource people in some of the target countries that can train the Initiative's team on such topics.</p> <p>The One CGIAR is a good opportunity to renew the CGIAR's portfolio and field of expertise, specially to introduce themes as potentially transformative as NPSs or agroecology, but this needs to be taken seriously, it is not just a change of label. A humble attitude is needed. Develop first the necessary capacities within CGIAR centers, before embarking in developing capacities among partners and stakeholders.</p> <p>No reference to developing early career researchers through recruitment into the teams. It very much a top-down approach that may lead to copious extension documentation that is used by few. The demonstration areas of Work Package 3 will provide excellent opportunities for participatory learning. However, it is unclear how these demonstrations will be shared beyond the local districts where they are developed.</p> <p>Capacities need to be developed not only on specific areas of research but mostly in the area of tradeoffs and options for ecosystems management through trans-disciplinarity lens.</p>			
<p>13. Project management mechanisms and (if applicable) additional scientific oversight and governance measures effectively and efficiently support the Initiative objectives</p>	<p>Management plan and Risk assessment 7, Research governance 8.1</p>	<p>Legitimacy, Credibility</p>	<p>3</p>
<p>Well-articulated project coordination in general. However, there is a lack of inclusion of the smallholder or cluster leaders in the management team and a single "representative" from each partner country is inadequate.</p>			
<p>14. Justified and transparent costing explicitly linked to expected Research for Development results</p>	<p>Financial Resources 10</p>	<p>Legitimacy, Effectiveness</p>	<p>2</p>
<p>Difficult to judge. Work Package 3 appears to be under-resourced if effective demonstration areas are to be developed and maintained; failure to resource these demonstrations will provide permanent demonstrations of ineffectiveness</p> <p>The budget seems to be small for the list of countries targeted. If the food production with sustainable intensification is duly addressed, this project may require additional funding. Maybe a strategy for co-funding may be considered.</p> <p>The budget is partitioned in the same amount per country. This seems unreal, given the different costs and logistics involved in each country.</p> <p>The amount of \$500,000 for upscaling across five countries does not seem to be sufficient.</p>			
<p>15. Anticipated research outputs (knowledge, technical, or institutional advances, specific technologies or products, policy analyses) are described and knowledge/gaps they will fill are evident. Protocols for open-data and open-access compliance are evident in plan (including budget)</p>	<p>Work Package research plans and ToCs 3.2, Open and FAIR data assets 8.2</p>	<p>Credibility, Effectiveness</p>	<p>2</p>

Criteria	Proposal Sections	QoR4D Elements	Consensus Score
<p>If the research outputs will be like those presented as “NPS innovations” coming from the CRPs in annex 3 then worry is warranted. Knowledge gaps are not well explained in problem statement and hence research needs are assumed to be the same across all countries.</p> <p>The Gantt chart provides the clearest summary of outputs. Developing new databases rather than using existing databases (i.e., FAO and WAI) for germplasm is a poor use of resources.</p> <p>Apart from the weak food security dimension, the overall project outputs are very ambitious.</p>			
<p>16. Monitoring, evaluation & learning (MEL) plan for the Initiative is clearly defined, with flexibility to adapt. MEL plan supports effective management and learning, including baseline data collection, and evaluative and review processes corresponding to stage-gates and course-correction decisions. MEL occurs during the life of Initiative and is used proactively to reflect on and adapt the Theory of Change, where appropriate</p>	<p>MELIA plan 6.2, Planned MELIA studies and activities 6.3, Measurable three-year (End of Initiative) outcomes 2.2</p>	<p>Credibility, Effectiveness, Legitimacy</p>	<p>2</p>
<p>While the plans are generic and suitable, it was left to the studies and activities to provide detail as to what will be done. It is unclear how conservation activities will be monitored. Unsure why Burkina Faso and Vietnam are excluded from plans?</p> <p>One of the reviewers was particularly worried about the lack of attention paid to quality food production using NATURE+ approaches. The food production side is lost amidst the biodiversity conservation aspect. The selected countries (Burkina Faso, Colombia, India, Kenya, and Vietnam) are places of many international programs and projects. Kenya is a hub of projects and international research programs. On p. 15, the link between agrobiodiversity and food production seems to be weak. In the Depth section the issue of conservation is reiterated, this type of agriculture is meant for production of food that has never been strongly stated. Table on pgs. 15 and 16 may include some impacts on food production.</p>			
<p>17. Well-defined plan for Initiative-level evaluation and impact assessment based on expected end-of-Initiative outcomes and impact. Links between the impact assessment plan and indicators in the Theory of Change are clear</p>	<p>MELIA plan 6.2, Planned MELIA studies and activities 6.3, Full Initiative ToC 3.1, Work Package ToCs 3.2, Projection of benefits 2.7</p>	<p>Effectiveness, Relevance</p>	<p>2</p>
<p>All the boxes are ticked correctly, but the Initiative does not seem feasible.</p> <p>See our earlier points in criteria 3 and 4.</p> <p>With a budget of 25 million over three years (to be spent mostly on personnel?), the Initiative will improve food and nutrition of 27 million people in five countries, will allow increasing the income of 66 million people between 15 and 50%, will enhance inclusiveness for 13 million women and 5 million young people, etc. It’s hard to believe. Having illusory plans is equivalent to having no plan at all. It’d be more useful to design a proper project, with detailed activities, specific research questions and a consequent budget, and reachable outputs/outcomes.</p>			

Additional Comments Not Presented Above

The review team would like to make some observations about Nature+ and two other Initiatives that have the potential to achieve excellent outputs and impact but as currently presented seem to overlap and propose different solutions to common challenges (noted above).

This proposal and the Transformational Agroecology Initiative have the potential to produce contradictory findings to other Initiatives such as Excellence in Agronomy. CGIAR needs to have a discussion on farming systems and approaches to uncover possible differences of opinions. This Initiative would benefit from the recommendation provided to Transformational Agroecology reproduced here verbatim: *"It would have been more useful to select case studies (countries/regions) where seeds of agroecological innovations—both technical and organizational—are already in place, so that the CGIAR researchers learn about agroecology and its approach before attempting to 'promote' it. It is therefore recommended that the authors link up with on-going Initiatives and organizations supporting agroecology innovation, especially from the Global South, where transformations have been taking*

place for about two decades. CGIAR scientists would then have the opportunity to first learn from these actors, learn about transformational processes, political, social, and technological. This would equip them with valuable insights to be able to write—together with these new partners—a proposal that could be truly transformative” (p. 66, *Review of 19 Initiative Proposals*). The same advice is applicable for this Initiative.

Also, Nature+ will promote neglected and underused plants and crop wild relatives seemingly in isolation from some of the Genetic Innovation Initiatives. The research in these Initiatives will take place in the same regions and as stated in the cohesion section has the potential for stakeholder confusion or worse. Nature+ contains good science but it should be better integrated and justified by a more informative problem definition, as stated, and we recommend inclusion of external partners with a track record in NPS research.

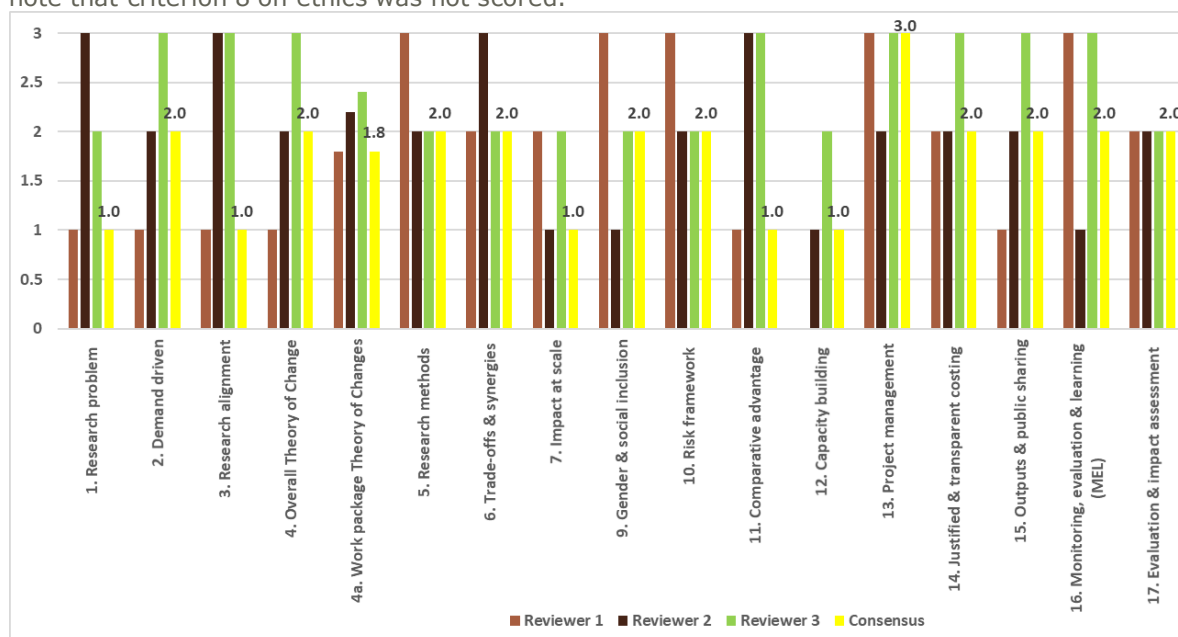
Since this Initiative is proposing (on a technical level) the same type of research that is proposed in the Agroecology Initiative, perhaps it is a good idea to try merge the two proposals, with sufficient time to rework their plans and engage new key stakeholders and resource people in CGIAR.

A two-stage proposal building is a mechanism that has been contemplated as part of this process, as explained in the Companion Document. When time is short, proposals tend to be mostly written by a small group of volunteers who often find themselves having to make decisions on methods and approaches that may fall outside of their core area of expertise. The review team believes this proposal would benefit from a second stage of reflection, learning across disciplines and partners, and more detailed planning that accounts for differences across the proposed target regions.

- Team up with outsiders with experience in agroecology and NPSs
- Explore stronger links and possibly mergers with Transformative Agroecology and Excellence in Agronomy, to avoid contradictory messages
- Build capacities on agroecology and NPSs within CGIAR before embarking on training other stakeholders
- Link up with other Initiatives that deal with genetic resources—especially Work Package 1
- Develop clearer and functional interactions between Work Packages (and avoid overlaps)
- Justify why and how NPSs can work for smallholder farmers (avoid extrapolating the problems of and solutions for industrial agriculture)

The reviewers encourage the team to rework the Initiative following the suggestions proposed and those from external experts with experience in agroecology and NPSs. One CGIAR needs a strong team and solid programs on both these themes.

The Figure below represents original scoring from each reviewer and consensus scoring for each criterion. *The consensus score across reviewers may not reflect the mathematical average.* For purposes of the Figure, the QoR4D criteria have been shortened. Please refer to p. 3 for full criteria definitions. Please note that criterion 8 on ethics was not scored.



*No consensus score was greater than 0.5 variance from the mathematical average.

8. Resilient Aquatic Food Systems for Healthy People and Planet

Review Summary and Actionable Recommendation

The Resilient Aquatic Food Systems for Healthy People and Planet (RAqFS) Initiative is highly relevant with potential to make significant contributions to CGIAR Impact Areas and more broadly to the 2030 Strategy. The Initiative addresses fundamental problems and challenges to the development of resilient aquatic food systems in the next 10 years. The Initiative places emphasis on development and scaling of innovations, and on creating the ecosystems necessary for research results to become innovations that are used and create value by stakeholders (farmers, producers, institutions, policymakers) in an integrated and inclusive way.

CGIAR's ability to convene a multidisciplinary team as well as its earlier experience in this area is a clear source of comparative advantage in addressing identified systemic challenges. The Initiative has been developed through a participatory process, responds to demand, and is targeted to maximize benefits.

The Initiative has clear potential to bring about significant outcomes and contribute to and provide a foundation for longer term impacts. This is substantiated through the overall project Theory of Change and complementary work programs with largely appropriate and ambitious End of Initiative outcomes and by convincing pathways to longer-term impact.

The review team has made the following recommendations, which are linked to identified weaknesses or cross-cutting remarks at the end of this review.

1. Strengthen the project management mechanisms to provide for a more responsive and adaptive approach and to ensure engagement of the senior management in the participating centers.
2. Strengthen the explanation of the role and engagement of partners and stakeholders and ensure that the related assumptions that underpin impacts pathways are addressed in a continuous and proactive manner.
3. Implement actions to guarantee the future sustainability of the project's outputs and outcomes including notably at the small-holder level.
4. Improve definition and tracking of indicators to measure direct and indirect benefits and impacts.
5. Ensure integration of the work package on new varieties (Work Package 4) with other Work Packages.

Overall, this is a strong proposal with the minimum review rating of 1 applied to just one criterion (13, project management). A few comments have been included under the individual criteria to assist in possible revision of the proposal.

Overall Strengths of Proposal

Contribution to Impact Areas and longer-term outcomes. The RAqFS Initiative is highly relevant and with potential to make significant contributions to each of the five CGIAR Impact Areas described in the CGIAR 2030 Research and Innovation Strategy and more broadly to the 2030 Agenda. The Initiative is timely in view of identified threats. CGIAR's ability to convene a multidisciplinary team is a clear source of comparative advantage in addressing identified systemic challenges.

The Initiative has clear potential to bring about significant outcomes and contribute to and provide a foundation for longer term impacts. This is substantiated through the overall project Theory of Change, appropriate and ambitious End of Initiative outcomes and by convincing pathways to longer term impact. Research questions, objectives, outputs, and outcomes are previously well identified.

Aquadata. The Aquadata Work Package (1) offers an exciting and much needed focus on creating and tracking the data that is fundamental to the improvement and sustainable management of aquatic food systems. The approach here is well reasoned and positioned to lead to improvements in data collection that will support food system transformation and address the knowledge gaps in tracking aquatic food system shifts.

Cross-sectoral working and learning. The Initiative incorporates cross-sectoral learning and an effort to build on strategies that have worked well in other CGIAR settings to foster interdisciplinary and participatory approaches. The integrated approaches of Work Package 3 and 5 can be expected to foster innovative solutions in complex settings that can be taken up by policymakers and used to inform improved aquatic foods systems.

Overall Weaknesses of Proposal

Management mechanisms, scientific oversight, and governance. The governance and oversight system as described does not sufficiently address overall project coordination and management in the CGIAR network context. There is limited information on i) how the leadership team will liaise intersessionally, between its roughly

annual meetings, to coordinate work amongst WPs, ii) on involvement of senior management in the participating CGIAR centers, and iii) on how external expertise, champions, or partners may be involved. Governance measures are described only in terms of compliance with CGIAR policies. This may be strengthened through a mechanism for engagement of senior management in the participating centers and, potentially, an external advisory board. The proposed leadership team of coordination staff should meet more than once annually.

Involvement of partners and stakeholders. The proposal includes a sound commitment to working with stakeholders and named partners and has clearly considered lessons arising from earlier CGIAR work. There is limited specification regarding the engagement and role of partners. The project assumptions for the causal pathways in the overall Theory of Change and in individual Work Packages refer to stakeholder commitment to codevelop solutions and partners engagement in scaling, and receptiveness of scientific evidence. The reviewers considered that these assumptions should be considered as drivers that need to be actively pursued at all stages of the project through ongoing dialogue and other mechanisms.

Integration amongst work programs. While cross-sectoral working and several aspects of internal coherence are viewed as strengths there are areas of the proposal that could benefit from stronger integration. In particular, the proposal would benefit from a more detailed strategy for integration of the work programme on AquaGenetics with Work Packages 2, 3, and 5 (addressing partners, foodscapes, and national innovation platforms) which can be expected to inform pathways for uptake and impact. In addition, making further linkages across CGIAR settings will leverage the lessons learned from other agricultural products—both regarding the development of improved strains, and how that process can best support their eventual uptake.

Areas of Divergence among Review Team and ISDC Resolution

Not applicable.

Does the Initiative Align with the Cohesion of the Portfolio as Described in the Companion Document?

External coherence: The Initiative is clearly demand driven. The proposal demonstrates a strong commitment to engagement of partners, many of which are identified. But their collaboration is frequently treated as an assumption and there is limited information on how this will be practically achieved. RAqFS has dedicated Work Packages on partnerships and AquaLabs (national innovation platforms) that provide useful platforms for developing and exploring mechanisms for coherence and sustainability.

Internal coherence: This is apparent in the overall Theory of Change, which reflects the complementarity of the five programs as well as the anticipated collaboration amongst Work Packages (which is stated rather than described). Two of the Work Packages (3 and 5) offer strong integrative potential at a landscape or national level. It is not clear to what extent this potential will be realized in view of limited overlap in the choice of focal countries.

Interdependencies: RAqFS is a partnership between five CGIAR entities. RAqFS anticipates collaboration with seven System Transformation Initiatives (AE-I; SHiFT; NPS; HER+; MITIGATE+; Markets and Value Chains; Foresight; Digital Technologies), two Resilient Agrifood Systems (Resilient cities; One Health) and four RII (AMD; ESA; TAFS-WCA; TAFSSA). Several of the individual work programs specify the intended collaboration but do not describe this in detail. National consultations in some target countries were coordinated with other Initiative Design Teams. The Initiative clearly supports the CGIAR Performance and Results Management Framework and includes cross-cutting budgets and human resources organized around each of the Impact Areas.

The reviewers scored each of the 17 QoR4D criterion individually. The review teams then built a consensus score. In the majority of reviews, the consensus score was the same as the mathematical average of the three individual reviewer scores. The criteria that received between a 0 or 1 are presented with a rationale below. Although not required, in order to keep review reports brief and readable, some teams opted to record in their reports their consensus-built rationale for scores above 1. Following the criteria is a figure presenting the individual reviewer scores and consensus. Please refer to p. 5 for the Likert scoring definitions.

Criteria	Proposal Sections	QoR4D Elements	Consensus Score
4. Overall Theory of Change with intended outputs, outcomes, and impacts at scale clearly described. Assumptions are documented, causal linkages are clear, especially the role of partners in driving impact	Full Initiative ToC 3.1	Effectiveness, Relevance	2

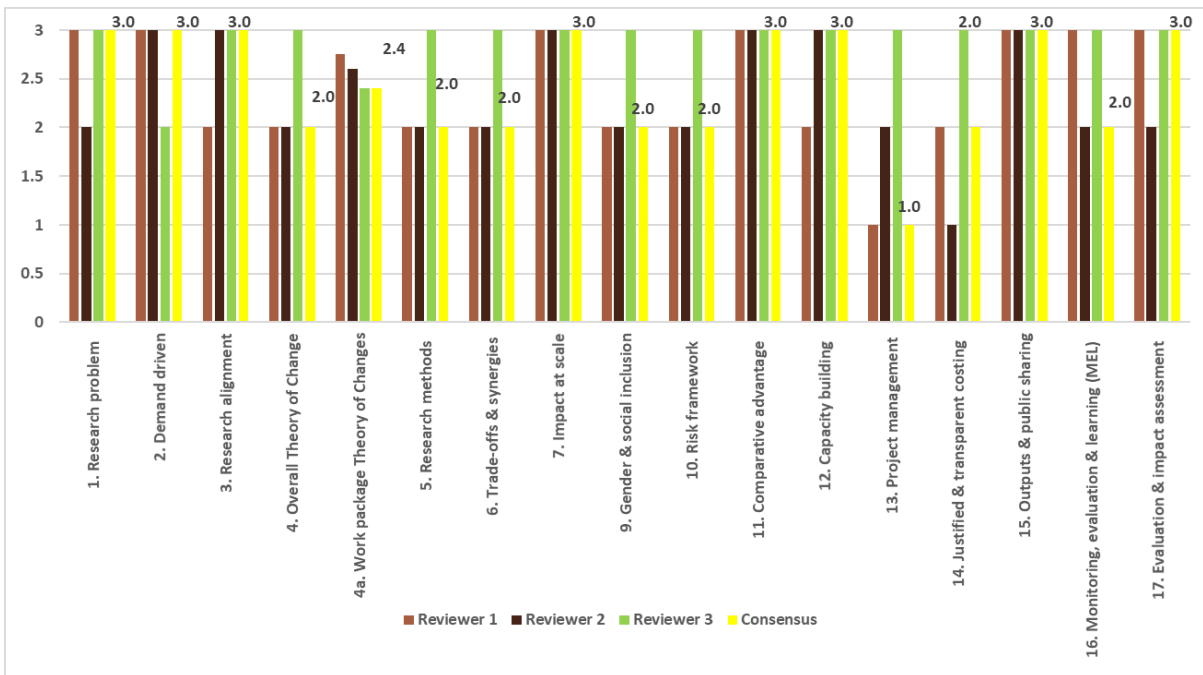
Integration: While cross-sectoral working and several aspects of internal coherence are viewed as strengths, there are areas of the proposal that could benefit from stronger integration. In particular, the proposal would benefit from a more detailed strategy for integration of the work program on AquaGenetics with Work Packages

Criteria	Proposal Sections	QoR4D Elements	Consensus Score
<p>2, 3, and 5 (addressing partners, foodscapes, and national innovation platforms) which can be expected to inform pathways for uptake and impact.</p> <p>Involvement of Partners and Stakeholders: There is limited specification regarding the engagement and role of partners. While this may be an artefact of the limited word count in the proposal template, the reviewers considered that further detail in this area would offset concerns around the related assumptions. (See criterion 5.)</p>			
<p>4.a Individual work package Theories of Change (score individually)</p> <p>Work Package 1 Work Package 2 Work Package 3 Work Package 4 Work Package 5</p>	<p>Work Package ToCs 3.2</p>	<p>Effectiveness, Relevance</p>	<p>WP1: 3 WP2: 2 WP3: 2 WP4: 2 WP5: 3</p>
<p>The Theories of Change for Work Package 2 includes some very ambitious intermediate outcomes that may not be feasible in the Initiative timeframe, particularly given starting point in RQs (as distinct to development practice, see also 5). Assumptions should be specified for the Work Package.</p> <p>The Work Package 3 impacts in the selected geographic areas would be enhanced by inclusion of training actions especially in pathways 1 and 2. The rice-fish decision support tool could be further specified (whose and which decisions). Work Package 4 would benefit from a description of how the work will engage with other work programs particularly Work Package 3 and 5. It would also be useful to specify how the Work Package will address the stressor identified under the challenge statement: aquatic animal diseases that destabilize production and antimicrobial resistance.</p>			
<p>5. Research methodology and methods (and supporting activities) are fit-for-purpose, feasible, and assumptions and risks are clearly stated</p>	<p>Work Package ToCs 3.2, Priority-setting 2.4, Innovation Packages and Scaling Readiness Plan 4.1</p>	<p>Credibility, Relevance, Effectiveness</p>	<p>2</p>
<p>Assumptions: The reviewers considered that the assumptions related to stakeholder commitment to codevelop solution, partners engagement in scaling, and receptiveness of scientific evidence should be considered as drivers that need to be actively pursued at all stages of the project through ongoing dialogue and other mechanisms. There is a high level of risk associated with non-compliance with these assumptions.</p> <p>Risk: The three-year project timeframe is unlikely to be sufficient to bring about durable change in stakeholders' and partners' behavior and at worst risks generating expectations and disruption at a local level. The proposal refers to earlier work and a possible follow phase and could provide further information on how continuity or exit strategies will be managed in the final year.</p>			
<p>6. Analysis of trade-offs and synergies across the CGIAR Impact Areas; ex-ante assessment of project benefits provides logical rationale for scaling of impacts</p>	<p>Projection of benefits 2.7, Result framework 6.1, Impact statements 5, Innovation Packages and Scaling Readiness Plan 4.1</p>	<p>Effectiveness, Credibility</p>	<p>2</p>
<p>The impact statement related to environment should distinguish the expected contributions of aquaculture and improved capture fisheries given the potential for conflict and competition between these sectors which is partly acknowledged in a proposed study on trade-offs.</p>			
<p>9. Research design and proposed implementation demonstrates gender and social inclusion that can be tracked in outcomes</p>	<p>Gender equality, youth & social inclusion 5.3, Projection of benefits 2.7</p>	<p>Legitimacy, Effectiveness</p>	<p>2</p>
<p>Gender is comprehensively considered in the narrative proposal; participation of youth is widely referred to, but the rationale for this and related strategies are developed only to a limited extent. There is only very limited</p>			

Criteria	Proposal Sections	QoR4D Elements	Consensus Score
consideration how benefits will be defined and tracked at the intervention level with related indicators tending to refer to inputs or immediate outcomes. (See also criterion 16.)			
10. A risk framework that details main project risks and mitigation actions across areas of science, funding, operations, partnerships, ethics, and environment	Risk assessment 7.3	Credibility, Legitimacy, Relevance	2
The risk framework is quite limited for an Initiative of this magnitude. A specific risk related to sustainability of outcomes and impacts is raised under criterion 5.			
12. Capacity building within project teams, partners, and stakeholders captured in capacity development plan. This can include development of early career researchers and partner staff, support/empowerment for under-represented stakeholders, and building partner networks	Capacity development 9.3	Credibility, Legitimacy	3
The development plan includes only a limited identification of approaches, with room for improvement on training. Training for trainers is a crucial tool to achieve the sustainability of the results, both for the technical ones, and the governance and organizational ones. Training in data collection and data management (Work Package 1), governance and associations creation and management (Work Package 2), production systems (Work Package 3), genetic techniques (Work Package 4), or entrepreneurship (Work Package 5) are a few examples of training initiatives that will facilitate the use of the knowledge achieved and the continuity of the initiatives started and scaled, despite the end of the Initiative.			
13. Project management mechanisms and (if applicable) additional scientific oversight and governance measures effectively and efficiently support the Initiative objectives	Management plan and Risk assessment 7, Research governance 8.1	Legitimacy, Credibility	1
There is limited information on i) how the leadership team will liaise intersessionally, between its roughly annual meetings, to coordinate work amongst Work Package s and ensure continuous, integrated, and adaptive management, ii) on involvement of senior management in the participating CGIAR centers, and iii) on how external expertise, champions, or partners may be involved. Governance measures are described only in terms of compliance with CGIAR policies.			
14. Justified and transparent costing explicitly linked to expected Research for Development results	Financial Resources 10	Legitimacy, Effectiveness	2
The reviewers considered that the budget information in the proposal is insufficiently detailed to allow them respond to the review question. The rating in this area reflects that this is presented according to the required template.			
16. Monitoring, evaluation & learning (MEL) plan for the Initiative is clearly defined, with flexibility to adapt. MEL plan supports effective management and learning, including baseline data collection, and evaluative and review processes corresponding to stage-gates and course-correction decisions. MEL occurs during the life of Initiative and is used proactively to reflect on and adapt the Theory of Change, where appropriate	MELIA plan 6.2, Planned MELIA studies and activities 6.3, Measurable three-year (End of Initiative) outcomes 2.2	Credibility, Effectiveness, Legitimacy	2
The proposal presents convincing and plausible pathways for long term benefits. However, there is a disconnect between the 2030 projected benefits and the End of Initiative outcomes, Work Package outcomes, and Work Package outputs level in particular related to gender and more broadly to consideration of benefits. The End of Initiative indicator on gender describes the number of actions (number of GTAs developed and number of partners adopting...), but not the real performance of these actions. Few of the Initiative indicators refer to beneficiaries and only one of these indicators refers to disaggregated tracking of benefits for men, women, and youth. There is limited information how Work Package 1 (Aquadata) will support impact studies at the level of local or national interventions.			

The Figure below represents original scoring from each reviewer and consensus scoring for each criterion. *The consensus score across reviewers may not reflect the mathematical average.* For purposes of the

Figure, the QoR4D criteria have been shortened. Please refer to p. 3 for full criteria definitions. Please note that criterion 8 on ethics was not scored.



*Two consensus scores were greater than 0.5 variance from the mathematical average. Please refer to criteria 13 (1.0 variance) and 15 (0.7 variance) above for rationale of these consensus scores.

9. Resilient Cities Through Sustainable Urban and Peri-urban Agrifood Systems

Review Summary and Actionable Recommendation

This is a proposal on an important topic, and the proposed Initiative has the potential to place this CGIAR and partners group at the center of what will undoubtedly become a major preoccupation for food systems research in the future. An important strength is the systems approach taken that considers peri-urban (UPU) production, consumption, and waste synergistically and considers mutually reinforcing solutions. However, the current version of the proposal also has some weaknesses, which include an unclear exposition of the precise contours of the existing knowledge that it is building on, imbalance in stakeholders that were consulted, insufficient consideration of how field impact and scaling will be prioritized and realized in practice, and insufficient attention to strengthening trade as a complementary way alongside UPU production to deliver equitable and sustainable nutrition to cities.

Actionable recommendations are:

- Flesh out major directions indicated by previous research and how the most promising areas from previous research will be fast tracked for early field implementation while research is undertaken on less well-understood areas.
- Strengthen the components of the proposal that are concerned with implementation and scaling, paying attention to networks of actors that will be engaged, ensuring strong co-design with the range of partners (particularly governments at various levels, NGOs, producer, trader and consumer associations), and considering whether budget allocations for field implementation and scaling are sufficient.
- Consider strengthening trade aspects, including elements such as transport, storage and marketing innovations.

Overall Strengths of Proposal

The systemic approach to studying urban food systems, linking production, consumption, and waste; considering both demand as well as supply side challenges and looking for mutually reinforcing solutions, is sound.

The proposal appropriately places priority on food access for the poorer sections of the population, and on the critically important but underserved informal sectors in the food and urban and peri urban farming sectors. It is pleasing to see focus trained on traders and their important but understudied and underappreciated role in local food systems.

Research outputs and the results framework contain good levels of detail.

Overall Weaknesses of Proposal

The greatest weakness of the proposal is its excessive focus on primary production and on government and NGO actors, when UPU agrifood systems are overwhelmingly based on the post-harvest components and driven by private sector decision makers.

The proposal does not engage sufficiently with internal/international trade, and with associated aspects such as transport, storage, manufacturing, food service, and marketing innovations (e.g., digital logistics) as key dimensions of urban food system resilience. While there is a market component to Work Package 2 that has some important elements in it, it would have been helpful for the initiative to consider the balance between UPU production and trade from a strategic perspective, giving greater attention to strengthening trade. There is a case to be made for short food chains and UPU production in specific circumstances, but it should not be taken as axiomatic that shorter chains are more efficient from any perspective—economic or environmental.

The proposal does not sufficiently explain how it intends to translate ideas into field results and scaling. While some aspects of partnership with international organizations are good, codesign (and co-funding) with a wider network of partners, and particularly with a range of NGOs, private sector partners and various levels of governments might have set the stage better for scaling.

The proposal could demonstrate more clearly how it is building on and adding value to research that already exists or is currently happening globally, within as well as outside the CGIAR. It notes that it is building on a significant body of work by CGIAR and others but does not clarify specific directions suggested by that body of work and how this new initiative will set about implementing or adding value to those ideas.

A case in point is Work Package 4 on nutrition, which is unusually exploratory, particularly for an area that has seen much research in the last decade. It proposes work on diets and food environments that many other One CGIAR initiatives are also proposing, and that a global community is already making considerable progress on. Rather than reinventing the wheel with a substantial (\$4m) budget, this Work Package might have been better

envisaged as pooling resources with other initiatives. Proposing tools for UPU dietary measurement, for example, seem a bit of a detour for this initiative—these are already well-studied in other initiatives and communities.

Areas of Divergence among Review Team and ISDC Resolution

Not applicable

Does the Initiative Align with the Cohesion of the Portfolio as Described in the Companion Document?

Yes, to some extent. Certainly, urban food systems resilience is a key component of the Resilient AgriFood Systems Action Area, and this proposal has the potential to be synergistic with a range of Initiatives, ranging from those relating to livestock, fish, and vegetables, to the markets and value chains and sustainable healthy diets programs, on gender, as well as the regional ones. That said, these potential synergies are not described in any detail, and are mostly restricted to scattered mention in the Work Package description tables. There are many areas where strong partnerships with related Initiatives from the outset could exploit complementarities and result in more efficient use of resources.

One example relates to Work Package 4 as described elsewhere in this report—a close collaboration on diets and food environments with other Initiatives would ensure resources are used efficiently across the portfolio. Partly this reflects excessive focus on UPU primary agricultural production, although UPU agrifood systems feature primarily consumers and post-harvest workers in the value chain. Even in lower-middle income countries, a sizeable majority of consumer food expenditures accrues to actors' post-harvest and that share is surely highest in UPU. Paying appropriate attention to the broader entry points and impacts of UPU agrifood systems will require careful coordination with country and regional offices, with the markets and gender global initiatives (and others), to devote sufficient attention to these crucial components of the systems. This is not historically an area of CGIAR strength, so likely needs more creative and far-reaching partnerships but upstream (to ARIs) and downstream, and especially with private sector actors, including large firms that heavily influence UPU agrifood systems.

The reviewers scored each of the 17 QoR4D criterion individually. The review teams then built a consensus score. In the majority of reviews, the consensus score was the same as the mathematical average of the three individual reviewer scores. The criteria that received between a 0 or 1 are presented with a rationale below. Although not required, in order to keep review reports brief and readable, some teams opted to record in their reports their consensus-built rationale for scores above 1. Following the criteria is a figure presenting the individual reviewer scores and consensus. Please refer to p. 5 for the Likert scoring definitions.

Criteria	Proposal Sections	QoR4D Elements	Consensus Score
1. Clearly defined research problem that addresses Impact Areas, is a high priority in the targeted geographies, is well aligned to shared, multi-funder priorities, and is well informed by previous research findings and evaluations	Challenge statement 2.1, Learning from prior evaluations and Impact Assessments 2.3, Impact statements 5	Relevance, Effectiveness	2
<p>The proposal articulates the distinct problems of urban food systems and the associated research opportunities well.</p> <p>Less convincing is the description of how previous research underpins the agenda proposed here. Some of the previous work (e.g., Urban Harvest) goes back 20 years, yet what the lessons from research from then onwards are and how they lead to the questions posed here are unclear.</p>			
2. Evidence that the Initiative is demand driven through codesign with key stakeholders and partners (Investment Advisory Groups, governments, private sector, funders) and research collaborators within and outside CGIAR	Participatory design process 2.6, Challenge statement 2.1, Work Package ToCs 3.2	Relevance, Effectiveness	1
<p>The proposal team has held several consultations, but a high proportion of consultations seem to have been with international organizations, with limited representation from government agencies at various levels, NGOs, social</p>			

Criteria	Proposal Sections	QoR4D Elements	Consensus Score
investors, producer, trader and consumer groups/networks, key city networks such as Smart Cities and Food Cities and private sector entities. It would have been particularly useful to get a sense for the extent to which this area is a strong priority for governments and influential businesses in agrifood value chains, as this will be critical to scaling. Another question is about the extent to which this was a co-design process with stakeholders, rather than presentation of a plan followed by discussion.			
3. Research questions, objectives, outputs, and outcomes are aligned to the research problem, and are measurable with defined deliverables	Work Package ToCs 3.2, Measurable three-year (End of Initiative) outcomes 2.2, Priority-setting 2.4, Management plan 7.1	Relevance, Effectiveness	2
<p>There is broadly a reasonable match of research questions and outputs with the research problem.</p> <p>However:</p> <ul style="list-style-type: none"> (i) The proposal fails to engage with a key strategic aspect: striking a good balance between trade and UPU production. Currently it seems to be taken as axiomatic that shorter value chains and more UPU production is the way to go. The scientific evidence is clear that value chain length has little explanatory power over the cost of a healthy diet, environmental impacts of a diet, or other desired outcomes. Trade will and should remain vital to resilient cities. Often, strengthening trade might be the most cost-effective way to deliver a nutrient or food to an urban population. From that perspective, it would have been good to see more explicit recognition of the balance between trade and local production reflected in the proposal. Yes, there is occasional scattered mention of rural-urban linkages. And there is a markets WP, but this seems much more focused on improving wet market sites (which is important) and less on facilitating (internal and international) trade in a strategic fashion to make cities resilient. (ii) A greater emphasis might have been expected on (a) what are incentives, social structure and primary influencers of the UPU producers, traders & businesses. Often projects fail because of incorrect assumptions about what the real requirements are for farmers and traders to change their behavior; and (b) better identification of the new successful techniques, technologies and business models emerging in the UPU space. The rapid rise of national and regional retail and food service businesses exerts considerable influence on UPU agrifood systems but is largely overlooked here. (iii) The End of Initiative outcomes imply additionalities that are likely to prove very difficult to measure ("new investment, increased availability" of safe food, etc.) 			
4. Overall Theory of Change with intended outputs, outcomes, and impacts at scale clearly described. Assumptions are documented, causal linkages are clear, especially the role of partners in driving impact	Full Initiative ToC 3.1	Effectiveness, Relevance	2
The overall Theory of Change is reasonable, with some innovative ideas such as facilitating investment in public goods through microfinance institutions. But the Theory of Change also has some weak elements relating to how to drive change at scale. It would make for a more compelling narrative if the team listed some potential methods that they intend to use to facilitate change. What translational methods, communication channels, and incentive structures are likely to work?			
4.a Individual work package Theories of Change (score individually) Work package 1 Work package 2 Work package 3 Work package 4	Work Package ToCs 3.2	Effectiveness, Relevance	WP1: 2 WP2: 2 WP3: 2 WP4: 1 WP5: 3
1: This Work Package plays to the CGIAR's strengths and is generally strong. It would have been helpful to have a) greater consideration of incentives and risks relating to introducing new production models, including radically new approaches well-suited to UPU systems, such as vertical farming, cellular and plant-based animal food substitutes, and b) better explanation of why some activities are located in some countries but not others.			

Criteria	Proposal Sections	QoR4D Elements	Consensus Score
<p>2: Work Package 2 is generally okay, but could have benefited from greater attention to trade, leveraging digital marketing pathways. explicit training in financial literacy, integration of middlemen into the program, and ideas about food safety testing and certification.</p> <p>3: This is an important Work Package given the importance of urban waste management. Some more specifics might have been helpful, e.g., any plans to tackle UPU human and livestock waste management, which is a major problem for many growing cities. Note that while recycling waste may result in many viable businesses, the potential for recycled grey water as an income generating enterprise in its own right is doubtful. A range of start-ups (and some large companies, like Buhler) are investing heavily in waste recovery for circular feeds systems based on insect proteins, but this proposal makes no apparent connection to such Initiatives.</p> <p>4: This Work Package is oddly exploratory, and it seems to go over the same ground that many other One CGIAR proposals as well a large non-CGIAR research community are dealing with. A number of CGIAR Initiatives seem to be seeking to do more or less similar work in slightly different geographies, and there is a danger of wasting resources in doing so. All Initiatives note that they will collaborate with other Initiatives in doing this work, but this has to be a firm, closely managed process from the start. Note also that much innovation in this area is already being engineered by Initiatives such as IMMANA and the AN4H Academy.</p> <p>5: The suggested mode of operating is sensible. The multi-stakeholder platform will be critical to scaling by helping build a sense of momentum for change and innovation and functioning as a catalyst for positive change through a wide range of invited partners. This is probably a key space for coordination with regional and country offices.</p>			
<p>5. Research methodology and methods (and supporting activities) are fit-for-purpose, feasible, and assumptions and risks are clearly stated</p>	<p>Work Package ToCs 3.2, Priority-setting 2.4, Innovation Packages and Scaling Readiness Plan 4.1</p>	<p>Credibility, Relevance, Effectiveness</p>	<p>2</p>
<p>The majority of methods and activities involve production of tools, literature reviews, stakeholder consultations, feasibility studies, etc. This is understandable to some extent given this is the first phase of activities. Still, more might have been expected by way of causal impact assessment of interventions, at least in cases where the previous portfolio of work has provided clear direction for intervention development. The actual methods that will be used to evaluate the work are unclear at this stage.</p>			
<p>6. Analysis of trade-offs and synergies across the CGIAR Impact Areas; ex-ante assessment of project benefits provides logical rationale for scaling of impacts</p>	<p>Projection of benefits 2.7, Result framework 6.1, Impact statements 5, Innovation Packages and Scaling Readiness Plan 4.1</p>	<p>Effectiveness, Credibility</p>	<p>2</p>
<p>Section 2.7 is clear, although it is confusing to have so many different ways of telling reviewers about outcomes. This section would be improved by linking each benefit to a specific Work Package—this was done for some benefits and not for others. As with the outcomes section, it is difficult to see how additionality might be proven for these large initiatives (i.e., diseases averted—what is the counterfactual?)</p> <p>The benefits projections and impact statement are explained in detail, but we cannot find much discussion of tradeoffs. The descriptions seem to suggest that there will be plenty of win-wins across multiple dimensions, but of course things are not quite that straightforward. An example is the balance between UPU production versus regional trade and the relative cost-effectiveness or environmental footprints, which is not addressed. Quite often short chains do not equate to efficient or cost-effective. Similarly, sometime resource intensiveness of UPU production with limited land may be higher than current baselines. A more careful discussion of these aspects is desirable.</p> <p>The enduring conundrum of development projects is between (i) small boutique projects which tend to improve the lives of a lucky few, and projects which generate almost imperceptible change across large number of beneficiaries. (ii) public good-focused projects, which are a comfortable space for donor funding but not necessarily effective or to work through the private sector, an approach which harnesses greater dynamism, effectiveness and investment, but can raise reputational risk issues, and (iii) the balance between a more blue-sky scientific approach, with its advantages of solid evidence-based foundation, but with weaknesses on its practicality, slow release and patchy uptake, compared with a greater focus on outreach, building on existing</p>			

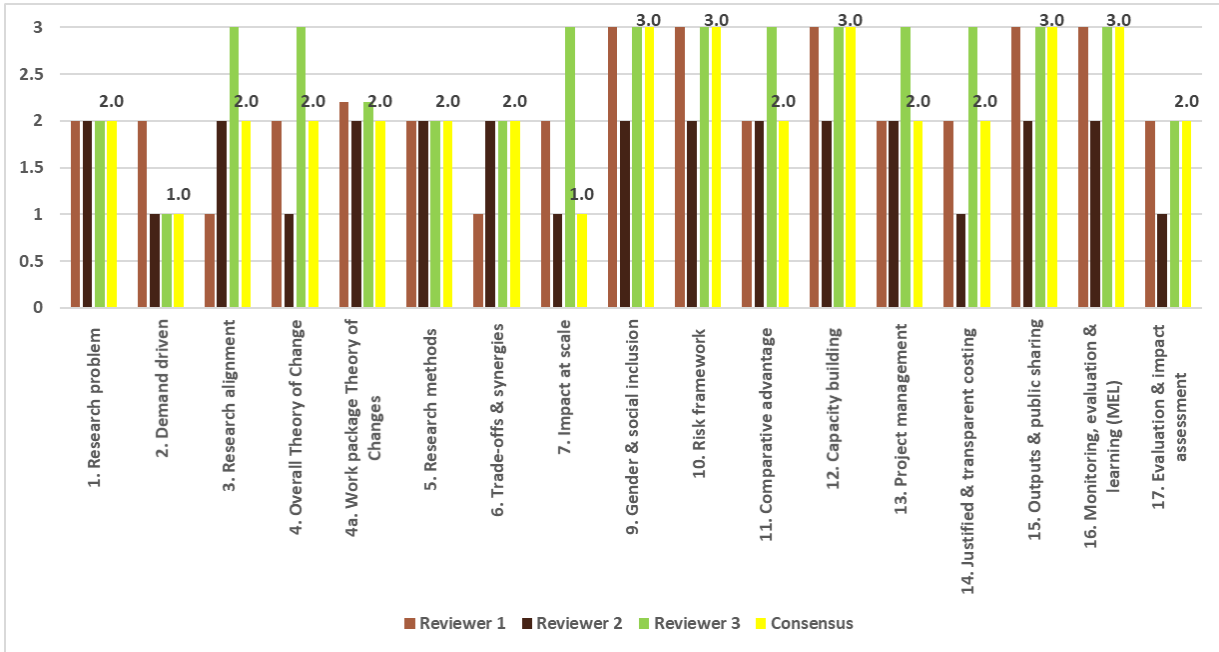
Criteria	Proposal Sections	QoR4D Elements	Consensus Score
<p>innovations. This project appears to have a sensible balance in category (i), would benefit from (ii) a stronger balance of promoting club goods, and (iii) lifting the outreach on the emerging field-based innovation. Such an approach should result in an acceleration and amplification of impact on the ground but avoid being criticized for supporting one company over another.</p>			
<p>7. Evidence that the Initiative will likely lead to impact at scale through integrated systems approaches that drive innovation in research and partnerships, including linking to and leveraging of other Initiatives within and outside CGIAR</p>	<p>Projection of benefits 2.7, Work Package research plans and ToCs 3.2</p>	<p>Effectiveness, Credibility, Relevance,</p>	<p>1</p>
<p>The proposal is well set up towards delivering research, but it is not yet clear that it is optimized towards delivering impact at scale in the field. One reflection of this is in the comment around consultations so far skewed towards international organizations, but patchy in terms of codesign with various levels of local governments, city initiatives, NGOs, businesses, etc., which will be critical to impact at scale. Another manifestation is in the comment that there isn't sufficient description of how the proposal builds on previous work around urban food systems, what innovations and solutions from that earlier research will be taken forward for implementation now, and the networks of actors who will be engaged in delivering that field impact. Instead, the focus is largely on fresh research (which is also important) and less on taking existing research into the field. Thirdly, we also think that the budget might need reconsideration in terms of improving scaling potential.</p>			
<p>8. Research design and proposed implementation demonstrates gender and social inclusion that can be tracked in outcomes</p>	<p>Gender equality, youth & social inclusion 5.3, Projection of benefits 2.7</p>	<p>Legitimacy, Effectiveness</p>	<p>3</p>
<p>The project design has a number of elements that will enable it to deliver positive results to some of the more disadvantaged sectors of society, including a focus on enterprises that are labor intensive and hold good potential for the involvement for women and youth. There is an additional aspect of equity that the team may want to consider, relating to spatial inequity within cities. Equitable access to nutritious food is often a spatial problem in cities, and it would be good to see a spatial lens applied to this problem.</p>			
<p>9. A risk framework that details main project risks and mitigation actions across areas of science, funding, operations, partnerships, ethics, and environment</p>	<p>Risk assessment 7.3</p>	<p>Credibility, Legitimacy, Relevance</p>	<p>3</p>
<p>The risk framework is satisfactory. Effectiveness of the Initiative will depend on being able to create a network of influential champions from the different sectors. Even then, it may prove very difficult to compete against the status quo of vested financial interests. The capacity to deliver is an enduring problem. In practice, these issues can be partially circumscribed by dividing the roles. The government monitors the progress of the project, and implementation of activities is outsourced to service providers.</p> <p>Risk #2 on "limited capacity of urban authorities..." is particularly important—possibly could also turn out to be limited commitment, given competing agendas and political realities. Management of this on the basis of a sound strategy will be extremely important.</p> <p>The team might like to consider identifying commercial and market risks and ways to mitigate those risks. The majority of new businesses fail. Studies on agricultural investing have shown that it is irresponsible to expose the poor to the considerable financial risks of embarking on a new business. It is important not to expose smaller-scale farmers, or businesspeople, to any significant financial risks. Hence, the importance of de-risking start-ups, building on well-resolved business models and proven technologies, and putting greater emphasis on creating more, better-paying, full-time jobs, many of which come from large firms, not micro-and small enterprises.</p>			
<p>10. CGIAR capacity and its comparative advantage and appropriateness to lead the work is justified. This includes the skills, diversity and multi-/trans-disciplinarity of the research team and approaches to meeting gender and diversity targets</p>	<p>Comparative advantage 2.5, Initiative team 9.1, Gender, diversity and inclusion in the workplace 9.2</p>	<p>Relevance, Legitimacy, Effectiveness</p>	<p>2</p>
<p>Urban food systems are not what would immediately come to mind as a comparative advantage of CGIAR. For example, in the Aug 2021 special issue of <i>Food Policy</i> on "Urban food policies for a sustainable and just future" only 1/46 authors declared a CGIAR affiliation. It is not immediately clear that the team has a good complement of economists, food industry experts, and spatial scientists who would be quite important to such a proposal.</p>			

Criteria	Proposal Sections	QoR4D Elements	Consensus Score
<p>A suggestion is that early-stage scientific ideas need to be shared with a broader community than researchers and academics. The network of potential partners operates in fields beyond the relatively narrow field of agricultural research such as digital experts, engineers, policy developers, formal and informal private businesspeople, farmer and NGO communities, financiers, logistics operators, etc. This will involve understanding and adapting to multiple cultural differences in language, communication, incentives, and the sense of urgency that these sectors have.</p>			
<p>11. Capacity building within project teams, partners, and stakeholders captured in capacity development plan. This can include development of early career researchers and partner staff, support/empowerment for under-represented stakeholders, and building partner networks</p>	<p>Capacity development 9.3</p>	<p>Credibility, Legitimacy</p>	<p>3</p>
<p>The capacity development plan is quite short but seems to touch on all of the relevant categories of support. It might be nice to include some specific examples of the type of development opportunities that junior staff might be able to participate in. There may also be an opportunity here to study the impact of leadership/diversity training.</p> <p>The team may want to consider developing specific training on interdisciplinary working, particularly where it comes to local partner institutions that may be strongly siloed.</p> <p>COVID has opened up a myriad of lower cost ways of training and sharing that the team will want to leverage as much as possible to deliver a rich set of capacity development opportunities.</p>			
<p>12. Project management mechanisms and (if applicable) additional scientific oversight and governance measures effectively and efficiently support the Initiative objectives</p>	<p>Management plan and Risk assessment 7, Research governance 8.1</p>	<p>Legitimacy, Credibility</p>	<p>2</p>
<p>Generally, the proposal presents a good overall plan. It would be good to have more on additional scientific (external) oversight. Even if there is a strong scientific team at the helm of this project, independent advice can be invaluable.</p> <p>It is vital that the measurement of processes (e.g., partner satisfaction, the buildup of commercial activities, and the progress of farmers and traders on the pathway to adopting new technologies and behaviors) are tracked. This enables management to adjust, respond, refine, and improve the way the project operates.</p> <p>This, along with other One CGIAR Initiatives, will be going beyond research to driving the delivery of field results to a greater extent than ever before. As past CGIAR reviews have highlighted, this implies the need to expand team skillsets to include a broader range of expertise and experiences, particularly in terms of management capacity geared for field delivery. This capacity is not yet apparent in the team, which still appears researcher heavy. The Initiative Design Team will want to give greater thought to the project's management team, and to personnel who have experience in managing the process of change.</p>			
<p>13. Justified and transparent costing explicitly linked to expected Research for Development results</p>	<p>Financial Resources 10</p>	<p>Legitimacy, Effectiveness</p>	<p>2</p>
<p>There isn't a sufficient breakdown and detailed explanation of expenditure plans, making it difficult to comment on how appropriate components of the budget are. However, we found this to be a problem with most proposals we have reviewed across our group and is likely constrained by proposal guidance and requirements. That said, we feel the proposal team may want to consider investing more on scaling activities, perhaps by repurposing (or pooling with other initiatives) some of the funds currently put towards Work Package 4.</p>			
<p>14. Anticipated research outputs (knowledge, technical, or institutional advances, specific technologies or products, policy analyses) are described and knowledge/gaps they will fill are evident. Protocols for open-data and open-access compliance are evident in plan (including budget)</p>	<p>Work Package research plans and ToCs 3.2, Open and FAIR data assets 8.2</p>	<p>Credibility, Effectiveness</p>	<p>3</p>
<p>Planned outputs are listed in good detail, and the availability of data follows reasonable scholarly expectations.</p>			
<p>15. Monitoring, evaluation & learning (MEL) plan for the Initiative is clearly defined, with flexibility to adapt. MEL plan supports effective management and learning, including baseline data collection, and evaluative and</p>	<p>MELIA plan 6.2, Planned MELIA studies and activities 6.3,</p>	<p>Credibility, Effectiveness, Legitimacy</p>	<p>3</p>

Criteria	Proposal Sections	QoR4D Elements	Consensus Score
review processes corresponding to stage-gates and course-correction decisions. MEL occurs during the life of Initiative and is used proactively to reflect on and adapt the Theory of Change, where appropriate	Measurable three-year (End of Initiative) outcomes 2.2		
<p>The results framework is considerably detailed. One thing to watch out for is the *depth* to which an outcome is achieved, given the indicators are largely counts. For example, the number of cities 'adopting' evidence-based approaches is an indicator, but the intensity of adoption is what will make a difference. The MELIA framework would do well to define meaningful intensity indicators early on in the Initiative</p> <p>We would suggest that Process indicators are included in the MEL. In addition, to improving the overall management of the program and its ability to ultimately deliver results on the ground, it will demonstrate progress to funders, partners and the governments.</p>			
<p>16. Well-defined plan for Initiative-level evaluation and impact assessment based on expected end-of-Initiative outcomes and impact. Links between the impact assessment plan and indicators in the Theory of Change are clear</p>	<p>MELIA plan 6.2, Planned MELIA studies and activities 6.3, Full Initiative ToC 3.1, Work Package ToCs 3.2, Projection of benefits 2.7</p>	<p>Effectiveness, Relevance</p>	<p>2</p>
<p>In the planned studies and activities, it seems like an end-line measurement (targets) of the key outcomes is missing. Perhaps it is implicit in the sections 3.1/3.2/2.7, since in those sections it says that certain indicators will be assessed. It would be nice to include this final assessment as part of the MELIA table.</p> <p>Rigorous causal studies of impact assessment remain unclear. Sections 6.2 and 6.3 have some text on this, but it is a bit vague. Since IA studies will inevitably need data over time (going beyond the initial 3-year period), studies will have to get underway pretty early in the project. More concrete early indications with some hint of outcomes, design approach, etc., might have been expected.</p>			

Review is continued on next page.

The Figure below represents original scoring from each reviewer and consensus scoring for each criterion. *The consensus score across reviewers may not reflect the mathematical average.* For purposes of the Figure, the QoR4D criteria have been shortened. Please refer to p. 3 for full criteria definitions. Please note that criterion 8 on ethics was not scored.



*No consensus scores were greater than 0.5 variance from the mathematical average but criterion 7 with a variance of 1.0. Please refer to criteria 7 above for rationale of the consensus score.

10. Securing the food systems of Asian Mega-Deltas for Climate and Livelihood Resilience

Review Summary and Actionable Recommendation

Sustainable agriculture in the Asian Mega-Deltas (AMD) is a huge and critically important issue, with implications for regional food security, coupled with attractive features in the proposal in nutrition and gender, with the intention of creating a holistic and synergistic impact. The proposal is strongest at what CGIAR is traditionally good at. However, the reviewers have doubts about the Initiative's understanding on how to: create, manage, and drive change through a multi-stakeholder coalition of partners; deliver field results and impact at scale; and be realistic about what can be achieved by 2025. A wide range of stakeholders have been consulted but it is not clear from the proposal how deeply the partners and stakeholders have been engaged in the design or whether they are an integral part of the project team, which are essential ingredients for achieving impact at scale.

Overall Strengths of Proposal

The proposal is addressing a huge, critically important, and mounting issue that is seriously threatening food security in the region and well as the livelihoods of the millions living in the AMD.

The proposal is strongest at what CGIAR is traditionally good at, and on the aim of creating a portfolio of activities that should lead to holistic and is mutually synergistic outcomes.

Within the international agricultural development arena, the CGIAR network has clear strengths in terms of the trust and respect that they are held, and the extraordinary quality of their researchers. This provides a strong starting point for any program.

Overall Weaknesses of Proposal

Reviewers are worried that the targets cannot be achieved by 2025 given the inevitable ramp-up period with a large, new Initiative. Whilst it is appreciated that the pressure is on researchers to produce ambitious outcomes, the proposal would appear to suffer from a "Planning Fallacy," i.e., being unrealistically optimistic and not considering previous experience.

There are doubts about the Initiative's understanding on how to deliver results on the ground and at scale (although this refers to this proposal, this would appear to be a challenge across the whole of CGIAR). The methods and description of activities would be more compelling if they were more detailed and specific. A successful program will require the type of embedded partnerships, management skills, expertise and experience that hitherto not been part of the CGIAR's traditional mode of operating. Suggestions have been made on how this important aspect of success can be better achieved.

It would appear that the Initiative has not been able to absorb and reflect in the project design sufficient external issues and opportunities. Examples would be (i) governments' existing plans and policies for their deltas, (ii) the existing innovative work by the CGIAR network and others (both national research programs as well as field-based innovations, engineering solutions) in the region, or (iii) concerns how much genuine co-creation was involved in the project's design.

Areas of Divergence among Review Team and ISDC Resolution

All three external reviewers agreed that this is an extremely important topic and liked its breadth of approach. The reviewers would have preferred a proposal that is able to sketch out with greater precision the specific activities envisaged that would be part of the program. Further, more detailed explanation of how the Initiative envisages it will actually create and translate research results into changing behavior on the ground at scale would be beneficial.

Two reviewers commented that the CGIAR project team is too researcher based, and with insufficient regional representation with doubts that a true co-creation of the project's design was undertaken. Further, these two reviewers believed the outcomes by 2025 are unrealistic.

One reviewer particularly commented on the need for the project team to have the necessary skills, tools, and techniques to successfully manage and motivate a broad coalition for change across the public, private, and producer sectors. There was a suggestion that building on existing global learnings (e.g., from the Nile Delta) and working in partnership with engineering solution developers/providers would strengthen the program.

Does the Initiative Align with the Cohesion of the Portfolio as Described in the Companion Document?

There was some surprise that the potential for building on CGIAR's past and existing programs were not explicitly made, e.g., WorldFish, IWMI, ILRI, as well as AVRC, or appreciation of and stronger program with alignment with the partner country existing policies. Linkages to other Initiatives e.g., Resilient Aquatic Food Systems were mentioned but in not enough detail to appreciate any integration of effort or the specific opportunities for complementarities or synergies.

The reviewers scored each of the 17 QoR4D criterion individually. The review teams then built a consensus score. In the majority of reviews, the consensus score was the same as the mathematical average of the three individual reviewer scores. The criteria that received between a 0 or 1 are presented with a rationale below. Although not required, in order to keep review reports brief and readable, some teams opted to record in their reports their consensus-built rationale for scores above 1. Following the criteria is a figure presenting the individual reviewer scores and consensus. Please refer to p. 3 for the Likert scoring definitions.

Criteria	Proposal Sections	QoR4D Elements	Consensus Score
1. Clearly defined research problem that addresses Impact Areas, is a high priority in the targeted geographies, is well aligned to shared, multi-funder priorities, and is well informed by previous research findings and evaluations	Challenge statement 2.1, Learning from prior evaluations and Impact Assessments 2.3, Impact statements 5	Relevance, Effectiveness	2
The proposal would benefit from greater detail explaining what activities are likely to be pursued and in incorporating into the project the important existing programs, plans, and innovations by CGIAR & others working in the Deltas of the region. By creating synergies like this, it will accelerate the uptake of innovations and help amplify field impact.			
2. Evidence that the Initiative is demand driven through codesign with key stakeholders and partners (Investment Advisory Groups, governments, private sector, funders) and research collaborators within and outside CGIAR	Participatory design process 2.6, Challenge statement 2.1, Work Package ToCs 3.2	Relevance, Effectiveness	2
The listing of organizations and institutions contacted is impressive. It was not clear what was the level of engagement (e.g., whether this was a mapping of potential partners, who had been presented the Initiative's plans, who were active codesign partners). However, two reviewers were skeptical that a true co-creation consultation has taken place. Further, two critically important potential partner sectors appear to have been overlooked, senior government officials as well as agribusiness community.			
3. Research questions, objectives, outputs, and outcomes are aligned to the research problem, and are measurable with defined deliverables	Work Package ToCs 3.2, Measurable three-year (End of Initiative) outcomes 2.2, Priority-setting 2.4, Management plan 7.1	Relevance, Effectiveness	2
This criterion is satisfactorily addressed, but it is overly ambitious in terms of outcomes, and is fuzzy about exactly how the research questions will lead to the output, and some methods are vague. The proposal would benefit from more explanation of the approach (e.g., randomized controlled trial and behavioral analysis of what) and in attempting to better understand existing solutions and their potential for leveraging.			
4. Overall Theory of Change with intended outputs, outcomes, and impacts at scale clearly described. Assumptions are documented, causal linkages are clear, especially the role of partners in driving impact	Full Initiative ToC 3.1	Effectiveness, Relevance	2
The reviewers would like to see more insight, detail, and attention on how CGIAR plans to positively drive behavioral change. For example, will Work Package 1 will work on Boro rice cultivation in Bangladesh, or on salt tolerant crops, introduction new species, new practices, SRI approaches, guide breeding efforts, develop small			

Criteria	Proposal Sections	QoR4D Elements	Consensus Score
<p>scale mechanization. In Vietnam where the government policy in the Mekong is to move rice production elsewhere and concentrate on fisheries, how will the program respond.</p> <p>The causal linkages need more to be more action orientated.</p> <p>Working through multiple partners and across nationals requires a medley of important soft skills, ability to communicate with disparate sectors in the food system, as well as the tools and management information systems to successfully manage the process.</p>			
<p>4.a Individual Work Package Theories of Change (score individually)</p> <p>Work Package 1 Work Package 2 Work Package 3 Work Package 4</p>	<p>Work Package ToCs 3.2</p>	<p>Effectiveness, Relevance</p>	<p>WP1: 2 WP2: 2 WP3: 2 WP4: 3 WP5: 3</p>
<p>Work Package 1 would benefit from being more explicit on the likely areas of activity and could usefully build on the findings of Work Package 2 and the multiple existing programs underway in the region's deltas.</p> <p>There is genuine enthusiasm for the aims of Work Package 2. The focus seems to be on diagnostic research, which isn't itself a problem, but would appear to be disconnected from delivering field results by 2025. Work Package 2 will need to acknowledge and account for the complexity and challenges in achieving the hoped-for results. The proposal would benefit from building on what works.</p> <p>Work Package 3 is a potentially important tool. It is an arena of very rapid change, and disparity between private and public sectors in their skills and understanding of the technology, as well as the significant gap that exists between the capacity of digital technologists and digital skills of the proposed end users. With those caveats in mind, the type of information suggested has a high chance of creating a positive synergy between private applications and public goods by drawing in users.</p> <p>Work Package 4 is satisfactory, especially if the focus is on those women and young people who are adjudged to be able to drive a positive spill over in behavioral change.</p> <p>Work Package 5 is a valuable approach, which will require new CGIAR skills and ways of operating.</p>			
<p>5. Research methodology and methods (and supporting activities) are fit-for-purpose, feasible, and assumptions and risks are clearly stated</p>	<p>Work Package ToCs 3.2, Priority-setting 2.4, Innovation Packages and Scaling Readiness Plan 4.1</p>	<p>Credibility, Relevance, Effectiveness</p>	<p>2</p>
<p>This is an area of the CGIAR's core competence and is well demonstrated in the proposal. There is some disconnect between activities and objectives (especially Work Package 1). Questions were raised on needs and the planned research: (i) If the program too diagnostic, (ii) whether the project design has considered building on what the emerging successful innovations in the region's deltas are, and (iii) on the urgency of driving forward positive change.</p>			
<p>6. Analysis of trade-offs and synergies across the CGIAR Impact Areas; ex-ante assessment of project benefits provides logical rationale for scaling of impacts</p>	<p>Projection of benefits 2.7, Result framework 6.1, Impact statements 5, Innovation Packages and Scaling Readiness Plan 4.1</p>	<p>Effectiveness, Credibility</p>	<p>2</p>
<p>A permanent increase in income was viewed as ambitious, given climate change and the uncertainty about the future of the ecosystem. The proposal would benefit from more detail and ironing out inconsistencies. (e.g., will incomes go up by 25% as in text, or 50% in table).</p>			
<p>7. Evidence that the Initiative will likely lead to impact at scale through integrated systems approaches that drive innovation in research and partnerships, including linking to and leveraging of other Initiatives within and outside CGIAR</p>	<p>Projection of benefits 2.7, Work Package research plans and ToCs 3.2</p>	<p>Effectiveness, Credibility, Relevance,</p>	<p>2</p>

Criteria	Proposal Sections	QoR4D Elements	Consensus Score
There was appreciation of the interconnectedness of the interventions. Two reviewers felt that proposal did not give sufficient attention to the new ways CGIAR would need to operate to drive the practice of change, and in operating through/with multiple partners and public, private, and productive sectors.			
8. Ethics, including equitable partnerships, information disclosure, biases, and potential conflicts of interest are considered; proposal defines how formal research ethics approvals will be sought/granted	Policy compliance and oversight 8	Legitimacy, Credibility	NA
Only one reviewer commented, and then about the broader ethical issues that development projects can trigger, e.g., scale vs boutique, food prices versus farmer profits.			
9. Research design and proposed implementation demonstrates gender and social inclusion that can be tracked in outcomes	Gender equality, youth & social inclusion 5.3, Projection of benefits 2.7	Legitimacy, Effectiveness	2
There was broad support, but all reviewers wanted the proposal to provide some clear foresight of what are likely to be the actual action areas, and more insight into the how the program would function.			
10. A risk framework that details main project risks and mitigation actions across areas of science, funding, operations, partnerships, ethics, and environment	Risk assessment 7.3	Credibility, Legitimacy, Relevance	2
The proposed risk framework is broadly satisfactory. Two reviewers identified the risk of the Initiative not being able to drive large scale impact, and therefore the need to give greater attention to how that risk will be mitigated (e.g., policy alignment, local champions, management information systems, understanding incentives, communication strategy, effective community of practices, etc.) The inherent risks of the poor taking on the financial risks of new ways of operating was also raised, as well as the need to apply effective mitigation practices (e.g., resolved business models, voucher schemes, underwriting risks, etc.).			
11. CGIAR capacity and its comparative advantage and appropriateness to lead the work is justified. This includes the skills, diversity and multi-/trans-disciplinarity of the research team and approaches to meeting gender and diversity targets	Comparative advantage 2.5, Initiative team 9.1, Gender, diversity and inclusion in the workplace 9.2	Relevance, Legitimacy, Effectiveness	1
Two reviewers expressed serious concerns that the proposed team did not appear to have sufficient breadth of expertise and local experience (e.g., soft skills, diversity, trans disciplinarity, etc.). Further, it was not clear from the proposal what partner skills would be brought into the project team to address some key gaps and challenges in delivering outcomes and impact at scale to provide the Initiative with real comparative advantage.			
12. Capacity building within project teams, partners, and stakeholders captured in capacity development plan. This can include development of early career researchers and partner staff, support/empowerment for under-represented stakeholders, and building partner networks	Capacity development 9.3	Credibility, Legitimacy	2
The proposal is broadly satisfactory though it seems to rely on building capacity through a handful of trainings. Reviewers' suggestions included that (i) CGIAR staff should be involved in training and mentoring—which they will need training in themselves, (ii) leveraging the new media options that COVID has opened up, (iii) an aim at capacitating a broader range of partners than researchers and academics, and (iv) considering incentivizing participation through existing continuous professional development programs.			
13. Project management mechanisms and (if applicable) additional scientific oversight and governance measures effectively and efficiently support the Initiative objectives	Management plan and Risk assessment 7, Research governance 8.1	Legitimacy, Credibility	2
Broadly satisfactory, with one reviewer emphasizing that success will to a large extent depend on creating a strong and diverse management team, i.e., about 60% of success of a project are dependent on its management skills (this percentage excludes factors beyond a project team's control). Effort should therefore focus on building a project management team with the necessary skills, experience, and knowledge.			

Criteria	Proposal Sections	QoR4D Elements	Consensus Score
14. Justified and transparent costing explicitly linked to expected Research for Development results	Financial Resources 10	Legitimacy, Effectiveness	2
All reviewers considered the budget data inadequate to make sensible judgements. One reviewer considered that the program was underfunded, given the critical importance of the topic for regional food security. We were pleased to see a budget for cross cutting issues across the Work Packages (at 20%), or 25% for global. Too often funders do not appreciate the need and benefits of cross learnings, coordination, and country exchanges. Additionally, given the current situation, it is likely to be really difficult to work in and use the budget in Myanmar.			
15. Anticipated research outputs (knowledge, technical, or institutional advances, specific technologies or products, policy analyses) are described and knowledge/gaps they will fill are evident. Protocols for open-data and open-access compliance are evident in plan (including budget)	Work Package research plans and ToCs 3.2, Open and FAIR data assets 8.2	Credibility, Effectiveness	2
As emphasized earlier, two reviewers are very doubtful about whether the targeted results are achievable by 2025. This would not only appear to be a classic example of "Planning Fallacy," (i.e., overambitious despite experience to the contrary). But also, that the proposal, as yet, hasn't given sufficient thought about how large-scale behavioral change is likely to be achieved in short order. Reviewers' suggestions to strengthen the proposal included (i) policy alignment, (ii) co-creation in design, (iii) actively managing a loose coalition of partners across multiple sectors—including the private sector and very senior policymakers, (iv) creating a network of influential local champions, (v) building on what is already working, collecting the information needed to manage the network, (vi) actively building a skilled, experienced, and diverse project management team, (vii) creating regional buy-in, (viii) achieving early wins to gain the necessary credibility to motivate large behavioral change, (ix) creating and maintaining a diverse and enthusiastic community of practice, (x) giving and acknowledging credit to local partners, (xi) understating appropriate incentives and the real drivers of change, (xii) creating local networks to implement change and provide feedback loops, and (xiii) exhibiting strong interpersonal skills etc.			
16. Monitoring, evaluation & learning (MEL) plan for the Initiative is clearly defined, with flexibility to adapt. MEL plan supports effective management and learning, including baseline data collection, and evaluative and review processes corresponding to stage-gates and course-correction decisions. MEL occurs during the life of Initiative and is used proactively to reflect on and adapt the Theory of Change, where appropriate	MELIA plan 6.2, Planned MELIA studies and activities 6.3, Measurable three-year (End of Initiative) outcomes 2.2	Credibility, Effectiveness, Legitimacy	2
The MELIA plan was generally considered to be comprehensive, with one reviewer commenting that current MEL could prove to be unmanageable given the number of field data points required and the difficulty of their measurement. Further, more emphasis should be given to process/progress measure of the program. This will be important in facilitating the management of the program. It will also help demonstrate to donors that the program is engaging with a diverse coalition of partners and able to deliver results on the ground.			
17. Well-defined plan for Initiative-level evaluation and impact assessment based on expected end-of-Initiative outcomes and impact. Links between the impact assessment plan and indicators in the Theory of Change are clear	MELIA plan 6.2, Planned MELIA studies and activities 6.3, Full Initiative ToC 3.1, Work Package ToCs 3.2, Projection of benefits 2.7	Effectiveness, Relevance	3
There were diverse views from the reviewers. Two considered this to be satisfactory, but with caveats. There is concern whether randomized controlled trials would work, given how intensive they are, the planning required and the necessary community acceptance, as well as dilution issues. The question was also raised whether the program is sufficiently grounded in what is possible in the deltas, given the complexity and impact of the multiple other interventions underway already.			

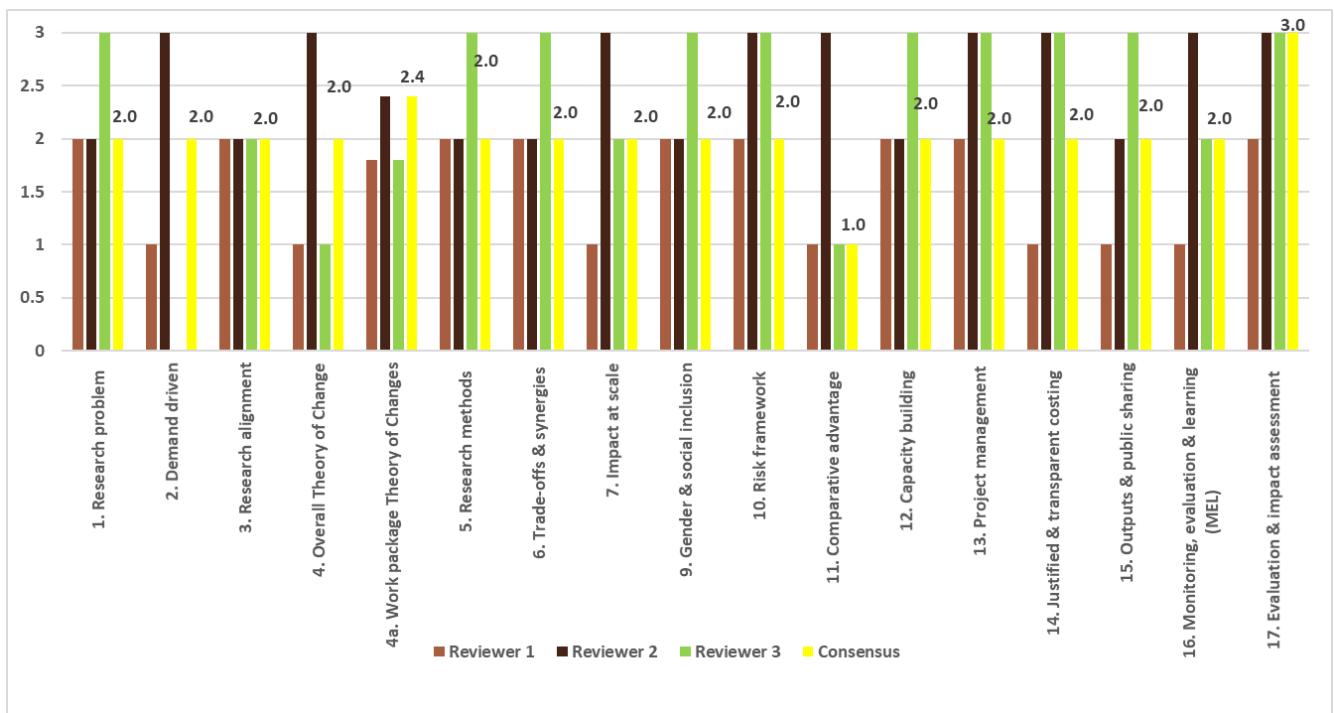
Additional Comments Not Presented Above

All the reviewers are acutely aware of the difficulties of proposal writing with its time pressures and the constraints of standard templates and having them peer reviewed. Consequently, some of our comments may be unfair or based on misunderstandings for which we apologize. However, we have taken the view that reviewing is more effective when the contrast is turned up, rather than turned down.

This, and other proposals, would benefit if a section was provided on the different interventions considered, and the reasons they were not pursued. Additionally, providing a stronger rationale for the importance of the program would be beneficial. This would help the reader better understand its importance, and why the specific strategy was chosen.

From the conversations with subject matter experts across a range of proposals, the points made about the need to demonstrate greater insight in operating a program to drive large-scale behavioral change would appear to be common across a range of the new Initiatives. As an example, some of the top agricultural engineering companies, when developing new technologies, form a core team of three. A researcher/innovator, a production engineer, and a sales/marketing person. This is to ensure that ease of manufacture, ease of use, product cost and demand are embedded into the process from the start.

The Figure below represents original scoring from each reviewer and consensus scoring for each criterion. *The consensus score across reviewers may not reflect the mathematical average.* For purposes of the Figure, the QoR4D criteria have been shortened. Please refer to p. 3 for full criteria definitions. Please note that criterion 8 on ethics was not scored.



*Two consensus scores were greater than 0.5 variance from the mathematical average. Please refer to criteria 10 and 13 above for rationale of these consensus scores (both with a variance of 0.7 from the mathematical average).

11. Sustainable Intensification of Mixed Farming Systems (SI-MFS)

Review Summary and Actionable Recommendation

Mixed Farming Systems (MFS) are the most important food production systems (milk, meat, and food crops) managed by smallholders in less-developed countries especially in SS Africa, South and SE Asia (including target countries Ghana, Ethiopia, Malawi, Bangladesh, Nepal; and Lao PR). Their Sustainable Intensification (SI) is therefore paramount for future food and nutritional security of billions of people in these regions. The Initiative is a three-year project within a 10-year framework and builds on previous research from relevant CGIAR CRPs that generated innovations for SI as well three bilateral projects CSISA, Africa Rising, and SIMLESA. It is a comprehensive and well-written proposal with notable strengths including the Work Packages and their integrated program logic and the MEL plan and the integration of both the biophysical and social dimensions of sustainable intensification of MFS. However, several weaknesses could affect the potential of SI-MFS to achieve the impacts that it aspires to. These include rethinking the impact statements especially with regard to more appropriate performance indicators; improving the impact assessment plans; further justification for the measurable three-year outcomes; and a more detailed budget breakdown and realistic budget for scaling readiness activities due to the complexity inherent in scaling innovation packages. Explicit recognition of the risks posed by the COVID pandemic and political instability in some of the target countries is also needed. It is therefore recommended that leadership group addresses these weaknesses.

Overall Strengths of Proposal

Section 3 Work Packages and associated Theories of Change: All Work Packages are comprehensive with useful Theories of Change. The organization into five Work Packages, each one addressing different aspects of SI of MFS, is logical. The narrative for each Work Package details research activities, the knowledge gaps that each Work Package will address, the ways the Work Packages are aligned/linked, the key partners/actors/beneficiaries, details of specific outputs and outcomes that will be produced, and the ways outputs, outcomes, and impacts are linked. Demand, scaling, and innovation partners are identified. Assumptions and risks are defined in each Work Package. The narrative shows the embedding of social justice and inclusion for fairer allocation of benefits to both women and men with each Work Package having a specific question on this topic. The high level of participation planned will also contribute to social inclusion. SI-MFS also shows a strong commitment to capacity building through Work Package 5 and the integration of capacity building throughout the entire initiative targeted at all levels: scientists, value chain participants, farmers, and policymakers and across public and private sectors and the initiative team itself. Appropriate tools and methods will be developed for each target and emphasis will be given to continuous reflection on processes and outcomes from systems approaches. Particular attention will be given to early career researchers especially women. Some specific adjustments in scientific methods in each Work Package are required as noted below.

Section 6.2 MEL Plan: The MEL plan is clearly defined, including a baseline study to be delivered in 2022 establishing baseline data for the End-of-Initiative outcome and within-Initiative outcome indicators. MEL data collection for tracking the progress of the indicators is planned on a semi-annual basis and the progress and evaluation of the initiative will be done annually through participatory approaches such as stakeholder consultations, focus group discussions, key informant interviews, or household surveys. This evaluative process ensures the flexibility to adapt and make course-correction decisions. The MEL plan is designed to be implemented during the three years of the initiative (2022-25).

Entire Initiative: SI needs to include clarity about principles and practices for priority setting, an inclusive understanding of the benefits, and awareness of trade-offs, based on scientifically acceptable, shared norms. It should be considered as a process of enquiry and analysis for navigating and sorting out issues and concerns. Sustainable intensification is about societal negotiation, institutional innovation, and adaptive management. SI-MFS addresses these issues by integrating both the biophysical and social dimensions of the planned activities.

Overall Weaknesses of Proposal

Section 5 Impact Statements: Despite some good work, this section suffers from major shortcomings. Similar and generic questions are repeated under the various sub-sections, although the proposal guideline asks for questions specific to each of the five Impact Areas. Some statements relating to indicators/metrics for measuring impacts are inappropriate. For example, on p. 37, the authors state "Food security, food availability, and household diet diversity score (HDDS) are standard indicators used to assess farming systems performance and are used in all five WPs." This statement is incorrect and demonstrates a lack of understanding of the associated concepts as well as of appropriate indicators for assessing the Impact Area of Nutrition, Health, and Food Security. For example, food security and food availability are different concepts, and the latter is a dimension of the former. Similarly, the HDDS primarily measures the access dimension of food security only. Moreover, HDDS is a household-level indicator and not a farming system-level indicator. For Impact Area 2, multiple indicators are

proposed together as one indicator (e.g., improved livelihoods of people through higher incomes). Here, the term “improved livelihoods” is vague, and income can be a component of improved livelihood. “Lift out of poverty” is difficult to comprehend as an indicator. For Impact Area 3, the suggestion of “empowerment” as an indicator is appropriate, but the way the concept has been defined is vague (e.g., objectives, aspirations, capacity). The specified dimension of empowerment (i.e., role in decision making) is quite narrow.⁹ More appropriate indicators however appear elsewhere in section 6.1. In Impact Area 4, the concept of resilience as an indicator is proposed with climate-induced productivity losses and large-scale adoption of GHG mitigation SI practices suggested. Neither of these indicators is broad enough to measure resilience (which is a complex, multidimensional construct) or captures the impacts of SI on climate change adaptation and mitigation.

The statements are also confusing, e.g., “large-scale adoption of SI innovations reduces GHG emission intensities (CO₂eq)” is suggested as a performance/impact indicator. However, this statement may suggest two indicators—adoption of SI innovations, and GHG emissions in CO₂ equivalent. Will both of these indicators be used? In addition, the role of livestock as GHG producers (methane) is ignored. Impact Area 5 (environmental health, biodiversity) should be a critical part of SI-MFS as it is a key element of sustainable intensification. However, relatively scant attention is given to measuring environmental outcomes (emissions, soil erosion, soil health, water quality, biodiversity, etc.). Indeed, there is no specific Outcome in the Results Framework (6.1) relating to the environment. Further, the statement regarding the performance indicators of Impact Area 5 is totally inappropriate and lacks correct indicators for assessing environmental health and biodiversity. How will this Initiative work with Nature+ and Transformational Agroecology since the Initiatives may have potentially conflicting objectives? How will the conflicting objectives be communicated across CGIAR and externally.

Section 6.2 MELIA Plan Impact Assessment: This part of Section 6 needs significant improvement. A plan for impact assessment (IA) research is provided. However, the proposed methods—including panel data methods, randomized controlled trials, and qualitative methods—are generic (i.e., each refers to a family of methods). Questions for causal IA research, as suggested in the proposal template, are not provided. The Annex provided through a hyperlink (p. 61) does not clarify the methods which will be used in evaluating program impacts and in establishing causality. For example, what methods will be used to assess the Initiative’s impacts on nutrition, food security, and health; poverty reduction, livelihoods, and jobs; gender equality, youth, and social inclusion; climate change mitigation and adaptation; environmental health and biodiversity? How will complex, multi-dimensional concepts like nutrition, food security, gender equality, poverty, adaptation, biodiversity, etc., be measured? What methods will be used to ensure that the observed impacts (e.g., variations in baseline and end line estimates of food security, nutrition, poverty, biodiversity, etc.) have occurred solely due to the Initiative’s interventions? Finally, SI-MFS dismisses previous IAs in SI as problematic and not of much use. At the least, they would contain lessons learnt to inform the proposal.

Section 2.2 Measurable three-year outcomes linked to Section 6.2: With the information presented in the proposal, it is not possible to judge with any certainty whether the planned outcomes and impacts are achievable. No basis to the *ex-ante* impact assessment is provided hence it is not possible to assess whether it provides a logical rationale for scaling benefits. As noted below, the numbers of people benefiting from the AR projects were small—thousands—not hundreds of thousands or millions. Given the end of three-year outcome is “improving the livelihoods of 1.5 million female and male actors,” more persuasive rationale is required to justify this optimistic outcome, especially when adoption curves suggest uptake rates of just over 1% by 2025.

Section 4.1 Scaling readiness plan and budget (also Section 10 Financial resources): This SR plan is modelled on the one developed by WUR and piloted successfully in CRP RTB Phase II. The main concern is that in RTB single innovations were tested while SI-MFS plans to test packages of innovations. This greatly increases the level of complexity and time and funds required to implement the scaling readiness process. The budget provided in Section 10 is not sufficient for the number of packages that are planned. For comparison, RTB subjected eight single innovations to the scaling readiness plan at a cost of USD 6 million.

A more detailed budget breakdown is needed for transparency and to judge whether the financial resources are adequate for the planned activities, though it is recognized this is a limitation of the proposal template.

Areas of Divergence among Review Team and ISDC Resolution

The consensus meeting resolved most areas of divergence among the three reviewers. Two reviewers were especially concerned about the lack of detail given in the budget (Score 1). One reviewer was less critical (Score 2) due to prior knowledge of CGIAR program/project budgets and noted the explicit inclusion of a sizable budget for cross-cutting activities.

⁹ Full guidance on developing the result framework content of the Initiative proposal template is available at <https://performance.cgiar.org>. Of special interest, collective global 2030 targets for transformation of food, land and water systems are available in the [CGIAR Performance and Results Management Framework 2022-2030](#). A support for the selection of appropriate indicators, by CGIAR Impact Areas, the Initiative will contribute to and will be able to provide data towards, is provided in the [MELIA Design Guidance](#).

Does the Initiative Align with the Cohesion of the Portfolio as Described in the Companion Document?

Cohesion of the CRP portfolio was an important aspiration during Phase II. However, the level of cohesion was limited due to lack of alignment among individual CRP priorities. Lessons must be learned from this to inform One CGIAR as portfolio integration is also very important. Figure 3 (Companion Document) shows how the planned Initiatives will link together for cohesion. Unfortunately, this only includes the first batch of Initiatives and does not include SI-MFS. The narrative of SI-MFS indicates that it will link with RIs, EiA, PHI, and HER+. It does not mention links with seed systems or plant breeding Initiatives—SeEdQual and ABI—or with climate mitigation strategies—MITIGATE. Such links are just as important as links with agronomy and pest management for sustainable intensification in farming systems. It should also be acknowledged that links will entail transaction costs and there is likely to be competition among initiatives. How will priorities be decided?

The management structure detailed in the Companion Document is very complex and most importantly, lacks clear hierarchy or lines of authority which could lead to delays in making decisions. There is a need for clear prioritization principles and guidelines as well as careful oversight on final decisions.

The reviewers scored each of the 17 QoR4D criterion individually. The review teams then built a consensus score. In the majority of reviews, the consensus score was the same as the mathematical average of the three individual reviewer scores. The criteria that received between a 0 or 1 are presented with a rationale below. Although not required, in order to keep review reports brief and readable, some teams opted to record in their reports their consensus-built rationale for scores above 1. Following the criteria is a figure presenting the individual reviewer scores and consensus. Please refer to p. 3 for the Likert scoring definitions.

Criteria	Proposal Sections	QoR4D Elements	Consensus Score
1. Clearly defined research problem that addresses Impact Areas, is a high priority in the targeted geographies, is well aligned to shared, multi-funder priorities, and is well informed by previous research findings and evaluations	Challenge statement 2.1, Learning from prior evaluations and Impact Assessments 2.3, Impact statements 5	Relevance, Effectiveness	2
The Initiative addresses SI of MFS. The Initiative is built on previous research from relevant CGIAR CRPs that generated innovations for SI as well three bilateral projects CSISA, Africa Rising, and SIMLESA. SI-MFS targets all Impact Areas to some extent however further clarity is needed (see Criterion 6 for more details). More thought is needed on how it will contribute to human health, food security and nutrition, employment, environmental health, and biodiversity. It is well-aligned with multi-funder priorities. The Challenge statement draws primarily on Herrerro et al. (2010)—a useful but dated publication. The proposal fails to take account of the successive theoretical and empirical advancements on MFS-SI, and related advances in innovation systems thinking, in the past 10 years. Stated knowledge gaps are broad and poorly articulated. Moreover, in several sections of the proposal, sweeping generalizations have been made without corresponding scientific evidence. Finally, SI-MFS dismisses previous impact assessments in SI as problematic and not of much use. At the least, they would contain lessons learnt to inform the proposal.			
2. Evidence that the Initiative is demand driven through co-design with key stakeholders and partners (Investment Advisory Groups, governments, private sector, funders) and research collaborators within and outside CGIAR	Participatory design process 2.6, Challenge statement 2.1, Work Package ToCs 3.2	Relevance, Effectiveness	2
The proposal builds on stakeholder consultations/surveys in the target countries. The stakeholders included both government and private-sector entities, within and outside of CGIAR. It provides some partner support statements; evidence of alignment with donor and target country priorities; and some evidence of consultations with global and regional initiatives. The proposal appears highly participatory. However, lack of details and non-functional hotlinks makes it impossible to assess the true extent of stakeholder consultations. A preliminary stakeholder analysis would have been useful. The proposed Initiative does not look like a codesign since the nature of stakeholder participation in the design process appears to be tokenistic. Evidence of community/farmer level consultations is also missing. Further clarity is needed.			
3. Research questions, objectives, outputs, and outcomes are aligned to the research problem, and are measurable with defined deliverables	Work Package ToCs 3.2, Measurable three-year (End of	Relevance, Effectiveness	2

Criteria	Proposal Sections	QoR4D Elements	Consensus Score
	Initiative) outcomes 2.2, Priority-setting 2.4, Management plan 7.1		
<p>The Initiative proposes five Work Packages that are aligned to the research problems and are logically organized to address SI-MFS problems at different scales of analysis. In each Work Package, the research questions are stated; outputs from each question are defined however objectives are missing. The rationale behind measurable outcomes is missing, which needs to be addressed given the ambitious three-year outcomes. Clarity is needed on who the policymakers are and the rationale behind Work Package 1 in the context of what is already known from previous projects on which SI-MFS is based. Without more details or impact assessments on how many farmers/stakeholders were reached by previous CGIAR research, we cannot judge the validity of the measurable three-year outcomes (2.2). Based on reach of some of the bilateral projects, 1.5 million is extremely ambitious. Finally for 7.1, the large size of the LG in the management plan needs more thought.</p>			
<p>4. Overall Theory of Change with intended outputs, outcomes, and impacts at scale clearly described. Assumptions are documented, causal linkages are clear, especially the role of partners in driving impact</p>	Full Initiative ToC 3.1	Effectiveness, Relevance	2
<p>For such a complex Initiative, the Theory of Change of SI-MFS is logical and clear. It builds on the best practices and outcomes from previous CGIAR research with five interlinked Work Packages, which are well-defined in the Theory of Change but less-well discriminated in Section 4a (see below). Outputs are not included. The outcome for WP5 is abstract and requires both quantitative and qualitative indicators. Causal links need more clarification: e.g., why would Work Package 1 outputs lead to a “transitioning” in the policies, priorities, and investment decisions of the IRIs and NRIs? Why and how would such transitioning, in turn, lead to farmers’ “use” of CSA technologies and practices? How will transitioning be measured and evaluated? Only one assumption is given; there is no reference to the pandemic or political risks. In Theory of Change narrative, the terms policies and institutions are used rather vaguely without clarifying what policies and institutions are important for SI-MFS.</p>			
<p>4.a Individual Work Package Theories of Change (score individually)</p> <p>Work package 1 Work package 2 Work package 3 Work package 4 Work package 5</p>	Work Package ToCs 3.2	Effectiveness, Relevance	WP1: 3 WP2: 2 WP3: 2 WP4: 2 WP5: 3
<p>For all Work Packages, the research objectives are relevant, and gender/inclusiveness is given priority. Partners are listed and assumptions and risks described. Synergies and dependencies with other Work Packages and Initiatives are outlined.</p> <p>Work Package 1: Status, trends, and future dynamics is a combination of understanding the best current targets and building foresight on the key drivers of change to inform changes in direction for SI-MFS. The proposal needs to clarify how much of the planned work has already been done. For example, how did the authors prioritize the target countries/regions without knowing their MFS status and trends first? Mention is made in the narrative of Work Package 1 about foresight and trade-off analyses, which is excellent, but there is nothing in the research questions or methods to indicate the approach or how it will be used.</p> <p>Work Packages 2, 3, & 4: These three Work Packages (tools and methods; MFS packages and scaling) are the “meat” of SI-MFS. Contributions to the scaling readiness plan are given for Work Packages 2 and 3 and Work Package 4 will implement the plan. The narrative places much emphasis on how these three Work Packages link together however this tends to blur the roles of the individual Work Packages and suggests some overlap. A figure could be included to show the individuality of each Work Package as well as their relationships more clearly. Wording of research questions in Work Package 3 suggests overlap with Work Package 1—further clarity needed.</p> <p>One area of activity missing from the Theory of Change for Work Packages 2 and 3 relates to labor. One of the challenges and areas of significant trade-off in most mixed farming systems is labor allocation between different components of the farming system and this also strongly links with gender issues. SI usually implies either a greater requirement for labor or at least greater attention paid to labor allocation, but SI-MFS is silent on this important socio-economic issue.</p> <p>Work Package 5: A comprehensive and well-crafted Work Package however proposed methods are broad/generic. How will the outcomes of “farming systems thinking” be measured?</p>			

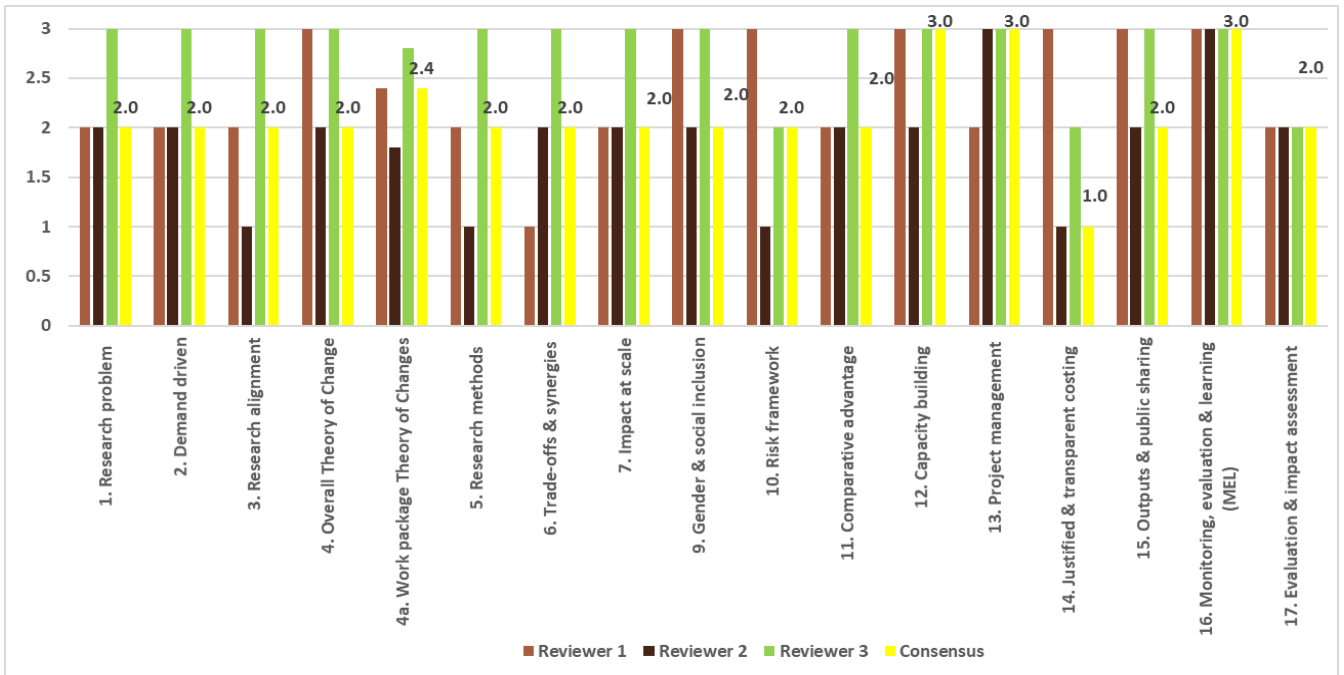
Criteria	Proposal Sections	QoR4D Elements	Consensus Score
<p>For all Work Packages: Assumptions are mostly reasonable in all the Work Packages, but risks are missing. One other concern is the apparent lack of linkages to key Initiatives on breeding improved varieties and seed systems (ABI, ACIPGT and SeEdQual)—these links are just as important as links to EiB and PHI in terms of developing packages for sustainable intensification in MFS.</p>			
<p>5. Research methodology and methods (and supporting activities) are fit-for-purpose, feasible, and assumptions and risks are clearly stated</p>	<p>Work Package ToCs 3.2, Priority-setting 2.4, Innovation Packages and Scaling Readiness Plan 4.1</p>	<p>Credibility, Relevance, Effectiveness</p>	<p>2</p>
<p>Research methodology and methods are broad, generic, and in some instances, inappropriate. A lack of methodological details about impact assessments is of particular concern. 2.4 Priority Setting is generic.</p> <p>Work Package 1. Methods—some specifications would have been useful, e.g., what would be the ‘cases’ in the proposed case studies? How are outputs 1.1 and 1.2 different? What method of literature review will be used? What ‘secondary data’ will be used? Key assumptions are correctly identified, but risks are missing.</p> <p>Work Package 2. The stated questions and methods need refinement, e.g., the question “what M&T allow...MFS for SI” is vague and does not match with the main aim of this Work Package (i.e., to develop, apply, and adapt existing and new M&Ts). We may use any M&T, but the question is whether an M&T is effective, reliable, and efficient. A recommendation is to revise the questions. Some stated methods appear to be vague (e.g., just stating qualitative, quantitative, and participatory isn’t informative enough) and inappropriate (e.g., farming system modelling is not a method for developing/adapting a new tool). Methods in the context of this Work Package should have been about methods of developing/adapting methods (e.g., what methods will be used to develop/adapt a toolkit for assessing the sustainability of an MFS?). Examples may be systematic reviews, Q-sort, Delphi, regression-type modelling (i.e., to validate an indicator), action research (e.g., create-apply-evaluate-revise a toolkit), etc. A suggestion is a “scientific methods” column in the respective table is re-written. Also, risks are missing in this Work Package as well.</p> <p>Work Package 3. Some proposed methods do not seem appropriate. For example, RQ2—how can this question concerning the uptake of SI innovations and their impacts in terms of MFSs’ sustainability, resilience, and equity be answered through stakeholder consultations and SWOT analyses? Assumptions are okay, but risks are not discussed, e.g., COVID-19 and conflicts (e.g., in Ethiopia) may affect participation. A lack of incentives for participation—a well-known problem—may itself be a risk. What measures will be adopted to mitigate this? In the absence of appropriate mitigation measures, participation may be difficult to achieve or just be tokenistic or box-ticking. There is a noticeable lack of attention given to measuring and monitoring environmental outcomes, which is a significant gap for an initiative on sustainable intensification.</p> <p>Work Package 4. The proposed method—use of the SIAF—for capacity building need assessment doesn’t seem appropriate. Assumptions are mostly okay but limited or no emphasis on partner resources (e.g., scaling needs money/resources). Work Package risks are missing.</p> <p>Work Package 5. The proposed methods (e.g., Bayesian networks, ag systems modelling) are generic (i.e., each represents a family of methods) and more specificity in this regard would be required. It’s not clear how agriculture systems modelling can answer RQ2. Work Package risks are missing.</p> <p>4.1 Innovation packages and Scaling Readiness Plan—12 innovation packages from previous CGIAR research potentially ready for scaling will be selected from a long list (provided) and will be subjected to the Scaling Readiness Plan developed by WUR (Sartas et al., 2020) which was successfully piloted with CRP RTB during Phase II for single innovations. The budget allocated to the plan is relatively small (USD 277K) compared to the funds used in RTB (USD 6 million) especially as SI-MFS is planning to scale packages and not just single innovations.</p>			
<p>6. Analysis of trade-offs and synergies across the CGIAR Impact Areas; ex-ante assessment of project benefits provides logical rationale for scaling of impacts</p>	<p>Projection of benefits 2.7, Result framework 6.1, Impact statements 5, Innovation Packages and Scaling Readiness Plan 4.1</p>	<p>Effectiveness, Credibility</p>	<p>2</p>

Criteria	Proposal Sections	QoR4D Elements	Consensus Score
<p>The proposal mentions trade-off analysis in various places but fails to clarify what those trade-off analyses would involve. It defines trade-off as conflict between sustainability and intensification. Such a definition does not suggest that the Initiative would undertake trade-off analyses between the five CGIAR Impact Areas. The proposal provides some quantitative estimates of potential/<i>ex-ante</i> impacts based on evidence from previous studies on the adoption of new agricultural practices at farm level. These projections seem to be logical for increased adoption rates, yield gains, income increase, and poverty reduction rate. Climate change adaptation and mitigation and environmental outcomes are missing. It is not possible to judge with any certainty whether the planned outcomes and impacts are achievable. No basis to the <i>ex-ante</i> impact assessment is provided hence it is not possible to assess whether it provides a logical rationale for scaling benefits. Clarity is needed on the scientific underpinning of the projected benefits regarding human health, biodiversity, and environmental health (e.g., livestock and methane production) and gender, youth, and social inclusion.</p>			
<p>7. Evidence that the Initiative will likely lead to impact at scale through integrated systems approaches that drive innovation in research and partnerships, including linking to and leveraging of other Initiatives within and outside CGIAR</p>	<p>Projection of benefits 2.7, Work Package research plans and ToCs 3.2</p>	<p>Effectiveness, Credibility, Relevance,</p>	<p>2</p>
<p>The projected benefits of the Initiative have the potential to create impacts at scale through the use of innovative systems approaches. It links explicitly with other Initiatives within and outside CGIAR. The Initiative includes an analysis of different scenarios of technology adoption for the different MFS, indicating the impact of the different indicators proposed in the project of the benefits section. Demand, scaling, and innovation partners are clearly defined in each Work Package. However, it is not very clear why and how those projected benefits/impacts will occur and be evaluated. Further clarity is needed on links with key Initiatives. For example, for developing improved varieties and demand-led, quality seed systems. An argument could be made that these are just as important for SI-MFS as agronomy and plant health. Transaction costs in linking with other Initiatives also need to be considered.</p>			
<p>9. Research design and proposed implementation demonstrates gender and social inclusion that can be tracked in outcomes</p>	<p>Gender equality, youth & social inclusion 5.3, Projection of benefits 2.7</p>	<p>Legitimacy, Effectiveness</p>	<p>2</p>
<p>Gender and social inclusion have been explicitly addressed in the proposal design and implementation. A number of outcomes and outputs relating to these are also proposed (e.g., as in section 6.1). All of these however relate primarily to the gender issue. Social inclusion and equality of youth and other marginalized groups (e.g., ethnic minorities) involved in agriculture are not adequately addressed. The success (?) of programs such as Youth Entrepreneurs and TAAT in providing sustainable opportunities for youth needs to be clarified if these are to be used as models. The capacity of a team with only two gender scientists needs to be re-considered.</p>			
<p>10. A risk framework that details main project risks and mitigation actions across areas of science, funding, operations, partnerships, ethics, and environment</p>	<p>Risk assessment 7.3</p>	<p>Credibility, Legitimacy, Relevance</p>	<p>2</p>
<p>The risk assessment was done to highlight potential areas of concern related to science and scientist skills, cohesion, legacy work, partnerships, operational, ethical, and legal among others. Potential mitigation strategies are noted as well. Some risks seem to be assumptions (e.g., partners do not have sufficient skills) rather than risks. No mention was made of the likely effects of the COVID pandemic. Also, no risks from conflicts (e.g., Ethiopia) and political changes in the countries that are participating were mentioned. These may well threaten the accomplishment of the initiative in time.</p>			
<p>11. CGIAR capacity and its comparative advantage and appropriateness to lead the work is justified. This includes the skills, diversity and multi-/trans-disciplinarity of the research team and approaches to meeting gender and diversity targets</p>	<p>Comparative advantage 2.5, Initiative team 9.1, Gender, diversity and inclusion in the workplace 9.2</p>	<p>Relevance, Legitimacy, Effectiveness</p>	<p>2</p>
<p>The comparative advantage of the CGIAR and its partners to lead and implement the proposed workplan is supported and justified by 10 key points and 30 citations proving considerable relevant research has already been done and experience gained. Of note are the formulation and implementation of interdisciplinary research activities, capacity in systems research ensuring that SI innovations are inclusive, implementation of research into practice for specific cases including partnership building, capacity development, and support for scaling, leadership on SI and the partnerships with NARES. However, Section 9.1 lacks team members' identities (profiles, CVs etc.) and more importantly, locations. This information is needed to fully answer this question. This</p>			

Criteria	Proposal Sections	QoR4D Elements	Consensus Score
<p>section should also disclose who the Work Package leaders are and whether they are CGIAR or partner scientists to be able to fully assess legitimacy. And it may be better if the Initiative's coleader is selected from the global South.</p>			
<p>12. Capacity building within project teams, partners, and stakeholders captured in capacity development plan. This can include development of early career researchers and partner staff, support/empowerment for under-represented stakeholders, and building partner networks</p>	Capacity development 9.3	Credibility, Legitimacy	3
<p>SI-MFS shows a strong commitment to capacity building through Work Package 5 and the integration of capacity building throughout the entire initiative. Capacity building will be targeted at all levels: scientists, value chain participants, farmers, and policymakers and across public and private sectors and the Initiative team itself. Appropriate tools and methods will be developed for each target and emphasis will be given to continuous reflection on processes and outcomes from systems approaches. How positive changes to this process will be measured needs to be identified early on in the implementation of the initiative. Particular attention will be given to early career researchers especially women. Further clarity is needed on the degree to which actors lack capacity. A pre-proposal capacity needs assessment would have been ideal.</p>			
<p>13. Project management mechanisms and (if applicable) additional scientific oversight and governance measures effectively and efficiently support the Initiative objectives</p>	Management plan and Risk assessment 7, Research governance 8.1	Legitimacy, Credibility	3
<p>The proposed management plan seems appropriate and feasible. The Initiative will be managed by a Leadership Group (LG) consisting of the Initiative leader and coleader, the leaders of each Work Packages s and MELIA, and also country representatives. Members from other Initiatives, which are related to SI-MFS are included. With already 14 members, there is a danger that the LG will be too large to reach consensus on critical issues if it is expanded to include leaders of other Initiatives this could facilitate linkages among Initiatives. For this purpose, the planned External Advisory Board (EAB) may be the better option in its role in advising on strengthening synergies between SI-MFS and other Initiatives. It would be useful to include in the EAB some members from the global South and to maintain a gender balance. Just one communication specialist for such a complex and multi-country Initiative seems a bit unrealistic.</p> <p>IS-MFS is already discussing possible linkages with EiB and PHI and this will help to align priorities. At this stage, it should also consider dialogue with other key initiatives such as the breeding and seed initiatives and MITIGATE. Once these initiatives set priorities, if these do not align with those of SI-MFS it will be difficult to capture synergies (lessons learned from CRP Phase II).</p> <p>Research governance arrangements have been agreed and standardized across the Initiative portfolio.</p>			
<p>14. Justified and transparent costing explicitly linked to expected Research for Development results</p>	Financial Resources 10	Legitimacy, Effectiveness	1
<p>Whether or not the budgets are justified and transparent is difficult to assess without corresponding budget narratives (justifications) and further breakdowns according to personnel (salaries & wages), equipment, materials and supplies, services, travel costs, overheads, proportional allocations between developing and developed country partners. However, accepting that the template only allowed for a summary budget, the allocation of an equal share to both SS Africa and Asia is noted as well as the sizable budget for cross-cutting activities which will foster linkages between Work Packages. Allocations of the same funds across years for some Work Packages needs clarification. Nearly \$4.5 million for Work Package may seem unreasonable as this Work Package primarily involves desk research with some interviews/focus groups and GIS mapping. Also, it is likely that mush of this information is already available. The terms global and crosscutting need clarification. The budget allocated to innovation packages and scaling readiness plan is not sufficient for the planned activities—12 packages of innovations—based on my experience with the SRP in RTB (as noted above).</p>			
<p>15. Anticipated research outputs (knowledge, technical, or institutional advances, specific technologies or products, policy analyses) are described and knowledge/gaps they will fill are evident. Protocols for open-data and open-access compliance are evident in plan (including budget)</p>	Work Package research plans and ToCs 3.2, Open and FAIR data assets 8.2	Credibility, Effectiveness	2

Criteria	Proposal Sections	QoR4D Elements	Consensus Score
Anticipated research outputs are logical, important, and adequately described. The section on Open and FAIR data assets notes that all relevant guidelines and protocols will be followed however further clarity is needed as the narrative only refers to some Work Package 3 outputs only. The proposal does not include any separate budget for open access compliance.			
16. Monitoring, evaluation & learning (MEL) plan for the Initiative is clearly defined, with flexibility to adapt. MEL plan supports effective management and learning, including baseline data collection, and evaluative and review processes corresponding to stage-gates and course-correction decisions. MEL occurs during the life of Initiative and is used proactively to reflect on and adapt the Theory of Change, where appropriate	MELIA plan 6.2, Planned MELIA studies and activities 6.3, Measurable three-year (End of Initiative) outcomes 2.2	Credibility, Effectiveness, Legitimacy	3
The MEL plan is clearly defined, including baseline study with deliveries in 2022 establishing a baseline data for the end-of-initiative outcome and within-initiative outcome indicators. MEL data collection for tracking the progress of the indicators is planned on a semi-annual basis and the progress and evaluation of the initiative will be done annually through participatory approaches like stakeholder consultations, focus group discussions, key informant interviews, or household surveys. This evaluative process ensures the flexibility to adapt and make course-correction decisions. The MEL plan is designed to occur during the three years of the Initiative (2022-25).			
17. Well-defined plan for Initiative-level evaluation and impact assessment based on expected end-of-Initiative outcomes and impact. Links between the impact assessment plan and indicators in the Theory of Change are clear	MELIA plan 6.2, Planned MELIA studies and activities 6.3, Full Initiative ToC 3.1, Work Package ToCs 3.2, Projection of benefits 2.7	Effectiveness, Relevance	2
<p>The MELIA plan describes the End-of-Initiative outcomes and outputs in each Work Package. Data source, collection method, and frequency of collection are described. The target year is the end of the Initiative. A concern is that no baseline value and year are included for each outcome. The projection of benefits describes well indicators, breadth, depth, and the probability to achieve impacts in 2030 with the Initiative. These impact indicators are resumed in the MELIA Plan and there are five Action Area outcomes that are relevant for the Initiative. However, GI 7 (farmers have access to and use climate-resilient, nutritious, market-demanded crop varieties), which is part of the innovations expected from Work Package 3 should be included.</p> <p>The proposal also provides a plan for IA research. However, questions for causal IA research, as suggested in the proposal template, are not provided. The Annex provided through a hyperlink (p. 61) does not clarify the methods which will be used in evaluating program impacts and in establishing causality (see further details under main weaknesses).</p>			
Additional Comments Not Presented Above			
The lack of space in the proposal template (word limits etc.) significantly reduces the amount of detail required to fully answer many of the criteria questions. We suspect that the Initiative team has much of the information that is missing. It is likely that some of the scores would have been higher if this information had been included in the template.			

The Figure below represents original scoring from each reviewer and consensus scoring for each criterion. *The consensus score across reviewers may not reflect the mathematical average.* For purposes of the Figure, the QoR4D criteria have been shortened. Please refer to p. 3 for full criteria definitions. Please note that criterion 8 on ethics was not scored.



*Four consensus scores were greater than 0.5 variance from the mathematical average. Please refer to criteria 4, 9, and 15 for rationale of these consensus scores (all with a variance of 0.7 from the mathematical average). Criterion 14 was facing a consensus variance of 1.0, which is described in rationale for that criterion above.

12. Transforming AgriFood Systems in West and Central Africa (TAFS-WCA)

Review Summary and Actionable Recommendation

This proposal is relevant. It addresses sustainable ways for quality food production using climate-smart solutions and it also deals with the net effect of food import while addressing domestic needs for quality food production and consumption. Several dimensions come with the Initiative that are articulated around dealing with climate change impact, adopting sustainable food production practices, accessing the right crop varieties to boost production, improving land management, and raising financial resource for improved investment in sustainable food production and nutrition security. However, there is a disjuncture in the Initiative in the sense that a lot of very useful information and analysis is provided to make the case for a relevant project, while at the same time the argument rests at least partly on some contended and pessimistic interpretations about what is actually going on in the countries targeted. As a result, is not clear what is the target for the various interventions.

Some considerations for improvement:

The baseline studies planned should be clarified. For example, drivers of food choices in target locations, evaluation of post-harvest technology options, political economy, etc. Some components within Work Packages appear to be starting from a position of unknown and the references to the learning material do not sufficiently demonstrate context. Additionally, the proposal should define the causal steps concerning key stakeholders, norms, and rules that need to change in order to achieve desired goals. We recommend to take a more integrated approach to strengthening food systems as articulated in the PIM CRP Review 2020 Brief (integration of ways the project implementation leads to the expected food security). The proposal could also be explicit about the strong engagement expected from actors and governments by making it clear the way they will have not only their buy-in but also the financial commitment that could be expected from governments as a way to fulfil their specific food systems goals referring, for example, to the Maputo Agreement.

Considering that food safety is part of food security, it is recommended that the proposal considers and addresses food safety in addition to providing nutritious food. "Food is not nutritious if it is not safe." There is a need to embed food safety and quality into nutritious offerings. This could start by "safe and nutritious" as opposed to "nutritious" narrative throughout the proposal.

It is not clear what the capacity building targets in-country mean in terms of current deficit and bridging gaps.

A detailed budget and level of flexibility in responding to dynamic changes/opportunities/constraints during program implementation-to assess if targets are realistic.

Overall Strengths of Proposal

The Initiative has strong alignment of Work Packages with previous impact assessment, Theory of Change, and integration of One Health concept. There is a strong need to address nutritious food for poor communities. The integrated approach for sustainable food systems integrates these aspects.

The gender transformative agri-business hubs are a great approach, and likely to facilitate access to finance, a key constraint. The Initiative will also include a technical hub to provide sufficient evidence to conclude that the Initiative will result in impact at scale by leveraging CGIAR and external resources.

Arguments around the comparative advantage of CGIAR, coupled with the plans for capacity building, represent a real strength of the Initiative. While this may sound like a contradiction, one does not want to see complacency in the sense that CGIAR, recognizing its strengths, wants to rest on its laurels.

Overall Weaknesses of Proposal

The research problem can be misleading and appears to be based on a defective Challenge Statement. Partially this is related to the disaggregation of context by target countries, rather than the generalized opportunities/threats/risks approach within the Work Packages linked to the country targets, budgets, and assumptions. For instance, how the proposal will improve yield of selected food products or trade-offs and the mismatch between the Research Questions as stated in Section 3 and those discussed in Section 5. For instance, a better alignment between RQ5.1 about management systems and the impact statement 5.2 about policy and institutional arrangements. Another alignment needed is for RQ5.2 about advocacy and media for community engagement, and the impact statement 5.2 about the vulnerability of rural poor to food and nutrition insecurity.

The proposal needs to better articulate the food safety and post-harvest interventions that might address foodborne illnesses across the different food systems (include reference to specific structural constraints) upstream-downstream and embed food safety and quality into the nutrition agenda.

Market incentives and entrepreneurship opportunities will be developed for young people and women. This brings the concerns for instance about how some components within Work Packages appear to be starting from a position of unknown and the references to the learning material do not sufficiently demonstrate context.

Areas of Divergence among Review Team and ISDC Resolution

The approach seems to repeat what CGIAR has been doing for decades, which raises the question: why will it work this time? Much of what is discussed in the proposal does address this latter question but would be more plausible if the overall research question were more clearly enunciated.

Does the Initiative Align with the Cohesion of the Portfolio as Described in the Companion Document?

No doubt that the Initiative includes acceptable level of coherence expected in a quality document. The content aligns with requirements for priority setting both at a national and community level, despite the blanket and generalized arguments in various sections of the document. The partnership in the consortium and targeted countries are relevant, although some countries seem to be already spoiled with international funding. The integrated approaches could be improved but there are ingredients of a good portfolio platform reflecting the Common CGIAR policies.

The reviewers scored each of the 17 QoR4D criterion individually. The review teams then built a consensus score. In the majority of reviews, the consensus score was the same as the mathematical average of the three individual reviewer scores. The criteria that received between a 0 or 1 are presented with a rationale below. Although not required, in order to keep review reports brief and readable, some teams opted to record in their reports their consensus-built rationale for scores above 1. Following the criteria is a figure presenting the individual reviewer scores and consensus. Please refer to p. 5 for the Likert scoring definitions.

Criteria	Proposal Sections	QoR4D Elements	Consensus Score
1. Clearly defined research problem that addresses Impact Areas, is a high priority in the targeted geographies, is well aligned to shared, multi-funder priorities, and is well informed by previous research findings and evaluations	Challenge statement 2.1, Learning from prior evaluations and Impact Assessments 2.3, Impact statements 5	Relevance, Effectiveness	1
<p>It will be beneficial to provide some references to previous research findings on opportunities concerning structural deficits and political economies in target locations. In this proposal, it appears that the white papers developed under AFDB for the Feeding Africa Initiative could be very relevant. There are a number of white papers that were developed for the Dakar Forum in 2015 to initiate the AFDB Feeding Africa Program. TAAT was one of the projects that came out of that Forum. Most CGIAR entities in Africa contributed to that study. They captured essential elements but before the Initiative starts, it will be useful to carry further scoping to capture the breadth of information available under development agencies.</p> <p>The Challenge Statement starts with some worrying and inaccurate statements. Furthermore, it highlights that if agrifood systems become more climate adapted, good things will happen. "Agrifood systems" is then defined as a) quality, nutrient-dense seed; b) climate-smart good agricultural practices (GAP); and c) reduced post-harvest losses. "Good things" are defined as "food and nutrition and health security." The question is how will this be done, and does this constitute a research question, in the sense that it is comprehensive and breaks new ground? Is this not what CGIAR has been doing for decades? Why will it work this time? For example, what is the unique research question to be addressed in this research?</p> <p>Nonetheless, the focus in the humid zone is appropriate to foster nutritious and climate-adapted food production systems in West and Central Africa (WCA). Furthermore, the targeted crops and species are well suited to the humid and transition zones of WCA. The inclusion of tilapia farming and crops, trees, and plants brings good diversity to the Initiative which aims to expand to the Sahel in subsequent phases.</p>			
2. Evidence that the Initiative is demand driven through codesign with key stakeholders and partners (Investment Advisory Groups, governments, private	Participatory design process 2.6,	Relevance, Effectiveness	2

Criteria	Proposal Sections	QoR4D Elements	Consensus Score
sector, funders) and research collaborators within and outside CGIAR	Challenge statement 2.1, Work Package ToCs 3.2		
<p>The project is timely and highly relevant. The rationale is quite strong and the links to policy priorities and location-specific relevance was demonstrated. Reviewers noted, however, an insufficient articulation of the root causes of the problems, thereby lacking contextualization. Rather, it is the symptoms that the proposed work targets for action. This shortcoming appears as early as in the Challenge Statement and in the research questions, even though there is much evidence from the wide range of project participants and their work that could point to root causes.</p>			
<p>3. Research questions, objectives, outputs, and outcomes are aligned to the research problem, and are measurable with defined deliverables</p>	<p>Work Package ToCs 3.2, Measurable three-year (End of Initiative) outcomes 2.2, Priority-setting 2.4, Management plan 7.1</p>	<p>Relevance, Effectiveness</p>	<p>2</p>
<p>All research questions are relevant. It will be good during the implementation phase to break them down to sub-questions that bring partial responses that could be aggregated to address the overall impact-oriented questions. The level of complexity of this project will not make it easy to have bundled questions to address a wicked problem.</p> <p>The review suggests additional research questions associated with:</p> <ul style="list-style-type: none"> • Root causes of problem statements (which are symptoms of systemic issues), that support better definition of the scope of interest/sphere of research within this program • Safety and quality (across the pre-post-harvest continuum) for target crops in traditional value chains-with potential to undermine accomplishment of targets for 'biofortified' value chains <p>The section on Priority setting (2.4) does not really address the setting of priorities.</p>			
<p>4. Overall Theory of Change with intended outputs, outcomes, and impacts at scale clearly described. Assumptions are documented, causal linkages are clear, especially the role of partners in driving impact</p>	<p>Full Initiative ToC 3.1</p>	<p>Effectiveness, Relevance</p>	<p>2</p>
<p>The overall Theory of Change is excellent. However, areas for strengthening this Theory of Change are linked to underlying assumptions with section 3 and 4 (Work Packages 1, 3, and 4). The general Work Package is well reasoned, but under the full Initiative Theory of Change narrative, more attention should be given to the link between Work Package 1 and 5.</p> <p>There is a lack of information about the expected increase in quality food production. The proponents went straight to the improvement of diet and income. Regarding digital tools and platforms to be used, the counterparts/users addressed by this project may not have access to digital tools, unless the project itself supplies needed soft- and hardware for such platforms. If such access was foreseen (e.g., through alternatives/partners), it needs to be described.</p>			
<p>4.a Individual Work Package Theories of Change (score individually)</p> <p>Work Package 1 Work Package 2 Work Package 3 Work Package 4</p>	<p>Work Package ToCs 3.2</p>	<p>Effectiveness, Relevance</p>	<p>WP1: 3 WP2: 2 WP3: 2 WP4: 2 WP5: 2</p>
<p>Work Package 1: Generally sound, the reviewers note that the targets may be overly ambitious for the seed systems. NARIs need considerable time to evolve as evidenced by Vitamin A cassava and OFSP, for example (see GAIN's Formative Research on Advancing Staple Crop Value Chains, Nigeria 2021 and others). Overall, the Work Package Theory of Change is good, but a legend is needed to understand the different colors and shapes. Smallholders may already implement good CSA practices—the Initiative might have taken the approach to identify those (current practices) and then amplify their use. The proposal reviewers were left questioning how new are the proposed CSA practices for target communities? New practices would take a lot of time to be adopted. About the key assumptions (p. 17), the causal linkage 1–10 is debatable. Past experiences show that</p>			

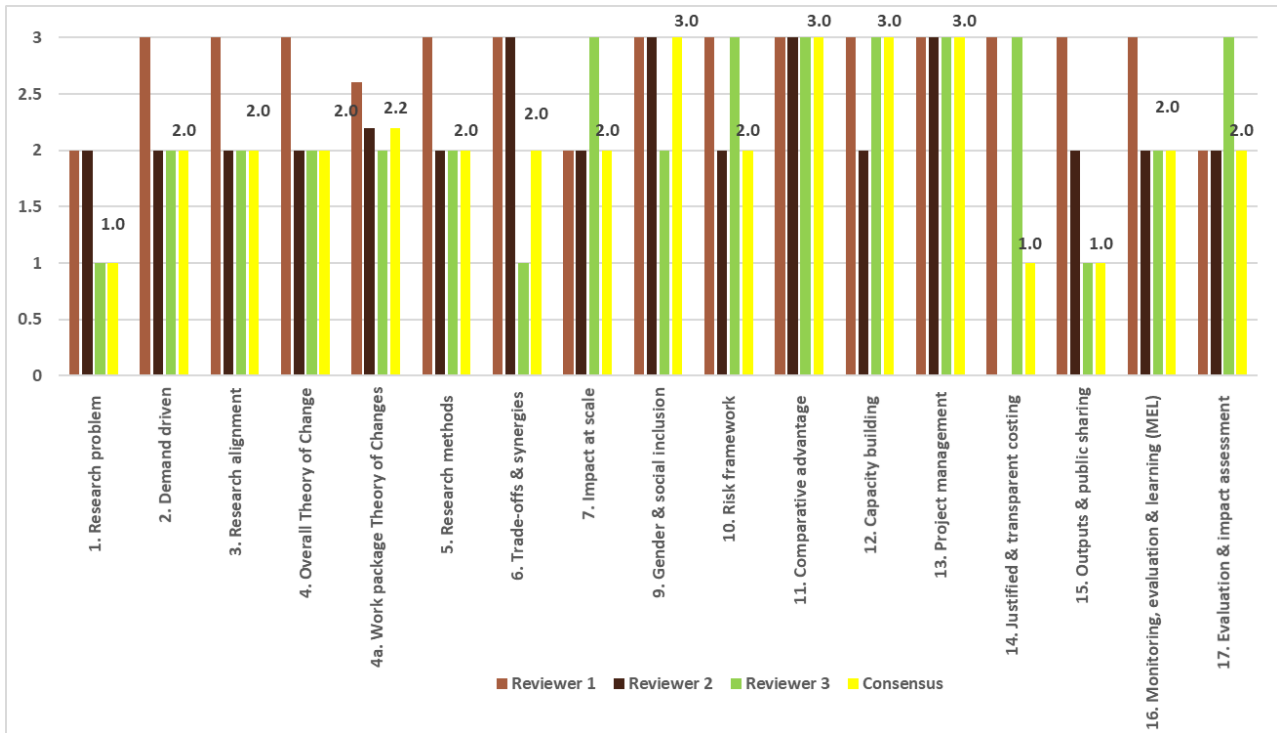
Criteria	Proposal Sections	QoR4D Elements	Consensus Score
<p>new products introduced within established production systems require trade-offs. The review doubts if the pilot stage would allow sufficient time for the Initiative to pilot and deploy innovations and answer the posited research questions.</p> <p>Work Package 2: Potentially transformative. The key will be business incentive to engage. The scientific methods are not provided for RQ2.2. Methods should address the "how?" issue. OP2.2.1. not clear. WP2 Theory of Change is simplistic. It is unclear how the various aspects of OP 2.1.1/2/3 link to OP4.3.1. and OP 2.2.1. is the reviewers struggle to understand the causal linkages in this Theory of Change. For RQ 2.2., using digital models require some pre-requisite in terms of IT access; the reviewers questioned if and how such access has been secured.</p> <p>Work Package 3: Further systems analysis of what has been/is preventing equal access to water and land in target locations, and why the proposed approach(es) will be successful in overcoming these, would make WP3 to be more credible. The scientific methods are not provided for all the research questions. Methods should address/explain the "how?" issue. The causal links are a bit shallow; meaning, the WP3 presents an aggregate of intended actions for scaling land and water innovations. What sustains these actions and how they are related to other Work Packages is not explicit. Also, the targeting is unclear: are the 100 rural communities equal to the 30 million persons using the innovations as stated earlier?</p> <p>Work Package 4: RQ4.4 should recognize that the safety issues and post-harvest technologies to reduce losses in biofortified crops are, in many cases, identical to those in the traditional dysfunctional supply chains. Planning to address these issues will go some way into addressing the constraints encountered with biofortified chains. There is little recognition in target outcomes that these are overlapping groups (e.g., young women). This Work Package seems to rely on two impact pathways: the financial/value chain pathway and the technology/innovations pathway. These could be much clearer and organized in a way that better showcases the logic for reaching the outcomes.</p> <p>Work Package 5: Innovation packages and Scaling Readiness Plan (SRP) are relevant and expected to be effective. This seems to link with Work Package 1 (especially via RQ5.2). However, some of the scientific methods are vague, e.g., "household survey instruments." Furthermore, there are inconsistencies in descriptions of WP 5 among a) the full Initiative Theory of Change narrative b) the Work Package main focus and prioritization, and c) the causal processes). A more uniform description should be decided upon. The reviewers don't understand what the authors intend to do to optimize scaling readiness (OP 5.1.1.). This Work Package combines several aspects that lead to different outcomes. Example: innovation and stakeholder management are not leading to the same outcomes. Consistency in bringing together all these aspects is required to make sure that what the proponents deliver in this Work Package leads to expected outcomes.</p>			
<p>5. Research methodology and methods (and supporting activities) are fit-for-purpose, feasible, and assumptions and risks are clearly stated</p>	<p>Work Package ToCs 3.2, Priority-setting 2.4, Innovation Packages and Scaling Readiness Plan 4.1</p>	<p>Credibility, Relevance, Effectiveness</p>	<p>2</p>
<p>This Initiative is believed to align with country, sector development strategies, and other agency priorities. However, there is insufficient analysis of the political economy for each target country, and trends demonstrating commitment to the sector, so that the assumptions and risks can be fully considered.</p> <p>The scientific method is not provided for some of the research questions. Also, the methods proposed are all off-the-shelf. The review team has encouraged to consider more innovation in this respect.</p> <p>The reviewers note that contribution is being conflated with attribution. There is doubt that the 3.93 million ha will exclusively improve under this project. It will be hard to say that this enormous area will change because of the project alone. Rather the project will contribute to influencing the adoption of innovation that may affect a large swathe of farming land (p. 13). Furthermore, there is an "old methods concern" that questions the uniqueness of the project approach and throws into question ability to reach impact.</p>			
<p>6. Analysis of trade-offs and synergies across the CGIAR Impact Areas; ex-ante assessment of project benefits provides logical rationale for scaling of impacts</p>	<p>Projection of benefits 2.7, Result framework 6.1, Impact statements 5, Innovation Packages and Scaling</p>	<p>Effectiveness, Credibility</p>	<p>2</p>

Criteria	Proposal Sections	QoR4D Elements	Consensus Score
	Readiness Plan 4.1		
<p>Rationale and priorities have been clearly articulated across activities and Impact Areas. However, the trade-offs do not receive enough attention in the proposal. There are a number of issues under section 5. They include incorrect data, failure to clearly identify impacts and a mismatch between the research questions as stated in Section 3 and those discussed in Section 5.</p> <p>Thought the arguments are slim, what the proposal presents regarding trade-offs is considered mostly as acceptable argumentation. However, the permanent increase of income from food production was shown with high certainty. The analysis of risks related to the increased production particularly from climate change was not clearly mentioned. Additionally, it is important to mention that income is related to markets. The weakness of markets under the poverty contexts of the study areas were not very apparent. These are factors that could reduce the certainty [of reaching impact]. The productivity dimension has been missing in the design of the project (p. 35).</p>			
7. Evidence that the Initiative will likely lead to impact at scale through integrated systems approaches that drive innovation in research and partnerships, including linking to and leveraging of other Initiatives within and outside CGIAR	Projection of benefits 2.7, Work Package research plans and ToCs 3.2	Effectiveness, Credibility, Relevance,	2
<p>Reference or further scoping for synergies with other targeted/long-term initiatives in-country would be beneficial (e.g., COLEACP and others). Capacity of partners should be objectively evaluated for both scope and quality. Agencies such as WFP, for example, may experience fragmented resources along with frequent staff changeovers that potentially negatively impact implementation continuity.</p> <p>The arguments in Section 2.7 are plausible and reveal how the numbers impacted were estimated.</p> <p>Partnerships point to a wide selection of CGIAR centers. Nonetheless, a sensitive question is to explore how previous CGIAR centers that have not joined One CGIAR could be included in the partnership mix (such as ICRISAT and ICRAF/CIFOR). While the proposal is forward looking regarding the fusion of center capacities within the Initiative design, the proposal begs the question on how best way to engage with those previous CGIAR centers in which a great deal of data and information for the subject matter remains valuable for the implementation of the Initiative?</p> <p>The partnership includes all layers needed from strategic orientation of the international development team to the local agencies and organizations for implementation. The partnership is set in a way that emphasizes complementarity and institutional capacity development to ensure success. The reviewers mention also that FARA could be a good partner for the policy and advocacy aspects.</p>			
9. Research design and proposed implementation demonstrates gender and social inclusion that can be tracked in outcomes	Gender equality, youth & social inclusion 5.3, Projection of benefits 2.7	Legitimacy, Effectiveness	3
<p>Gender and social inclusion, along with appropriate indicators, feature strongly throughout the proposal. However, in light of cultural and social norms, gender inclusion and equality for the most disadvantaged groups across some parts of the target countries may take longer than expected in the proposal. Policy advocacy tools will need to be strong.</p> <p>The reviewers question sticking with certain approaches that may have shown in the past to not be totally reliable. If one has between a 30-50% certainty that something will work, based on past experience (see section 2.7, subsection 3), then surely existing approaches/processes/mechanisms need to be changed, expanded or improved upon, not repeated. Probabilities this low ought to be seen as a challenge to develop new procedures and support the more efficient use of the project budget by investing in methods that produce greater ROI. The same remark is repeated for the "environmental health & biodiversity" function, where the certainty level is also 30-50%.</p> <p>In the various countries, there are many initiatives that promote youth entrepreneurship and agri-business startups. The reviewers suggest this Initiative work on their mapping and support them (p. 12).</p>			
10. A risk framework that details main project risks and mitigation actions across areas of science, funding, operations, partnerships, ethics, and environment	Risk assessment 7.3	Credibility, Legitimacy, Relevance	2

Criteria	Proposal Sections	QoR4D Elements	Consensus Score
<p>Include unfavorable developments in the political economy. Risks related to the market, and risks related to technologies used are lacking. Both market and technologies act as potential places of risk.</p>			
<p>11. CGIAR capacity and its comparative advantage and appropriateness to lead the work is justified. This includes the skills, diversity and multi-/trans-disciplinarity of the research team and approaches to meeting gender and diversity targets</p>	<p>Comparative advantage 2.5, Initiative team 9.1, Gender, diversity and inclusion in the workplace 9.2</p>	<p>Relevance, Legitimacy, Effectiveness</p>	<p>3</p>
<p>CGIAR has significant strategic partnerships. The acumen exists and the skilled scientists and partnership are strong assets for this project. Section 2.5 contains a solid statement of CGIAR's comparative advantage. However, the capacity of NARES should be assessed as an important resource. There is competition from many ongoing (and future) initiatives in-country for NARES staff. Capacity deficits should be evaluated and addressed prior to the implementation stage.</p> <p>One comment on Section 9.2: The aspect of "gender, diversity and inclusion in the workplace" (section 9) is comprehensively addressed. Nevertheless, it is disconcerting to read that gender balance in the implementing team is unlikely to be achieved (section 9.2).</p>			
<p>12. Capacity building within project teams, partners, and stakeholders captured in capacity development plan. This can include development of early career researchers and partner staff, support/empowerment for under-represented stakeholders, and building partner networks</p>	<p>Capacity development 9.3</p>	<p>Credibility, Legitimacy</p>	<p>3</p>
<p>The plans are good, but their implementation must be monitored. It would be useful to understand how the targets fit into an overall capacity deficit, in order to be able to assess potential impact. Also, the exit strategy of the project could include stronger engagement of local champions and leaders.</p>			
<p>13. Project management mechanisms and (if applicable) additional scientific oversight and governance measures effectively and efficiently support the Initiative objectives</p>	<p>Management plan and Risk assessment 7, Research governance 8.1</p>	<p>Legitimacy, Credibility</p>	<p>3</p>
<p>The project management approach seems well articulated with an inclusive management approach. There may be some additional focus on country initiatives and their custodians needed (see notes on MELIA resourcing in criterion 16).</p>			
<p>14. Justified and transparent costing explicitly linked to expected Research for Development results</p>	<p>Financial Resources 10</p>	<p>Legitimacy, Effectiveness</p>	<p>1</p>
<p>Although the allocation of the budget to the different Work Packages is rational given the descriptions of work to be done in each and has been wisely distilled across activities and countries, more details and justification needs to be provided for completeness (insufficient details).</p>			
<p>15. Anticipated research outputs (knowledge, technical, or institutional advances, specific technologies or products, policy analyses) are described and knowledge/gaps they will fill are evident. Protocols for open-data and open-access compliance are evident in plan (including budget)</p>	<p>Work Package research plans and ToCs 3.2, Open and FAIR data assets 8.2</p>	<p>Credibility, Effectiveness</p>	<p>1</p>
<p>The Initiative lacks inclusion of political economy and impact of structural constraints.</p> <p>There's not much evidence of innovative research approaches: methods, etc. seem to be off-the-shelf and standard. This could be because there is no hard, recognizable identification of knowledge gaps in the proposal, which could in turn be the product of a vague overall Theory of Change.</p>			

Criteria	Proposal Sections	QoR4D Elements	Consensus Score
<p>16. Monitoring, evaluation & learning (MEL) plan for the Initiative is clearly defined, with flexibility to adapt. MEL plan supports effective management and learning, including baseline data collection, and evaluative and review processes corresponding to stage-gates and course-correction decisions. MEL occurs during the life of Initiative and is used proactively to reflect on and adapt the Theory of Change, where appropriate</p>	<p>MELIA plan 6.2, Planned MELIA studies and activities 6.3, Measurable three-year (End of Initiative) outcomes 2.2</p>	<p>Credibility, Effectiveness, Legitimacy</p>	<p>2</p>
<p>There are strong links (with what?) and the Initiative is well aligned with Performance and Results Management System (PRMS). It is suggested that the MELIA specialist is full time, not part time, as this is a critical component. Previous internal learning material suggested insufficient resources allocated to MELIA. This will therefore also facilitate impact assessment with the additional resource, if adopted.</p> <p>Create provision for post-Initiative evaluation and handing over, in order to facilitate impact assessment, unforeseen externalities, greater sustainability, and reduce cost of future baseline studies. This implies sufficient capacity at stakeholder-level, appropriate stakeholder-indicators, which are also easy to capture during routine operations.</p> <p>An aspect that is not described are consequences if planned targets and End of Initiative outcomes go unmet. One recommendation is to work with partners on the indicators before the start of the project to segment various indicators in various categories of impacts pathways.</p>			
<p>17. Well-defined plan for Initiative-level evaluation and impact assessment based on expected end-of-Initiative outcomes and impact. Links between the impact assessment plan and indicators in the Theory of Change are clear</p>	<p>MELIA plan 6.2, Planned MELIA studies and activities 6.3, Full Initiative ToC 3.1, Work Package ToCs 3.2, Projection of benefits 2.7</p>	<p>Effectiveness, Relevance</p>	<p>2</p>
<p>The budget needs to be linked to expected end-of-Initiative outcomes and impact, otherwise the Initiative is impossible to assess. Planned innovations may have been developed in the appraisal aspects solely concentrating on CGIAR-based innovations (p. 9); this review suggests that non-CGIAR innovations appear to be a suitable complement to the Initiative outcomes. Additionally, to assess impact more robustly, nutrition, health, and food security might consider qualitative indicators (p. 10). Currently, as presented, a relative improvement of 10% on income is not adequate for the Impact Area of nutrition, health, and food security. Furthermore, the measure for climate adaptation may not be solely about increase in income but better assessed through the understanding of adoption of climate sensitive innovations.</p>			

The Figure below represents original scoring from each reviewer and consensus scoring for each criterion. *The consensus score across reviewers may not reflect the mathematical average.* For purposes of the Figure, the QoR4D criteria have been shortened. Please refer to p. 3 for full criteria definitions. Please note that criterion 8 on ethics was not scored.



**Four consensus scores were greater than 0.5 variance from the mathematical average. Please refer to criteria 1 and 10 (both with a variance of 0.7 from the mathematical average) and criteria 14 and 15 (both with a variance of 1.0 from the mathematical average) above for rationale of these consensus scores.*

Appendix A: Eschborn Principles

Inclusion of Eschborn Principles in QoR4D Criteria

Each QoR4D criterion was mapped against the Eschborn Principles that were adapted by Transition Consultation Forum in April 2020.

1. Major multi-funder, strategically aligned, fully funded CGIAR Initiatives, laid out in multi-year investment plan. This definition explicitly rules out “buckets” or “gluing” together of bilaterally funded projects. Together, these CGIAR Initiatives constitute the CGIAR shared agenda funded by pooled funding.
2. Different disciplinary knowledge and research is used to address food, land and water system issues identified with the stakeholders in any specific region/country, drawing on the global agenda of work.
3. Compelling Theory of Change to achieve impact at scale on SDG2 and other Sustainable Development Goals (as framed by CGIAR’s five Impact Areas).
4. A clear problem statement, rigorous priority-setting, purpose-driven solutions and a focused set of metrics for success.
5. Generate diverse approaches designed to address the stated problem as effectively as possible using an integrated systems-based approach, rather than relying on supply-driven solutions.
6. Apply operational and geographic focus in areas of recognized CGIAR competencies and achieve impact by working strategically with partners that have complementary competencies, at all stages of research-for-development.
7. Manage the research-to-development process via a sequence of stage-gated decision points at which there is a review progress along the theory of change and a resulting reallocation of resources, to support an ongoing funnel of best-bet innovations from early stage through to scaling.
8. Realistic and transparent costing explicitly linked to expected results.
9. Inspired by the future (where we want to get to, but also unforeseen events) not only by where we come from; some innovations might not be demanded at the present, but their importance will emerge (in often unpredictable ways).
10. Use appropriate and innovative metrics of success, considering time lags from research to large-scale impacts, and making the most of modern tools such as genetic markers.
11. Integrate strongly with emerging work on country-collaboration, financial modalities, resource mobilization, governance and shared services (through smart interactions with other TAGs).

Appendix B: Companion Document Review Responses

Initiative	Does the Initiative Align with the Cohesion of the Portfolio as Described in the Companion Document? ¹⁰
AgriLAC Resiliente: Resilient Agrifood Innovation Systems Driving Food Security, Inclusive Growth, and Reduced Out-Migration in Latin America and the Caribbean (LAC)	<p>The outputs, outcomes, and impacts of the Initiative are well-aligned with the CGIAR Impact Areas and the SDGs. The proposal states that the work will be conducted in collaboration with the regional and country offices. This alignment to the CGIAR Performance and Results Management Framework will allow for a coherent and cohesive integration of the portfolio. The proposal demonstrates cohesion through synergies with local partners. The different elements of the CGIAR Performance and Results Management Framework have been taken into consideration and are explicitly mentioned in the proposal. Regional and activity integration, co-creation with local partners, and country offices are an important part of the proposal and addressed in the Work Packages.</p>
Foresight and Metrics to Accelerate Food, Land, and Water Systems Transformation	<p>One of the great challenges posed by the UNFSS was the need to de-silo the agriculture sector, and enable informed decision making about food systems transformation, through analyzing synergies and trade-offs inherent in alternative policy and investment pathways across food, land, water and health sectors. The potential contributions of foresight modeling (and related capacity strengthening at regional and national levels) to meet these challenges would seem to be immense. Yet this proposal falls disappointingly short, especially given the CGIAR’s considerable foresight expertise and track record to date.</p> <p>It is not clear that the vision and planning for the Initiative stretches more than incrementally beyond the usual agriculture sector focus. If it cannot adapt to meet the broader, cross-sector needs—and identify more specifically how existing models will change to do so—it is unclear how the Foresight and Metrics Initiative can serve the linchpin role anticipated for it in One CGIAR and that is so badly needed at country, regional, and global levels. In addition, although the Foresight and Metrics Initiative mentions planned linkages with several other initiatives throughout the proposal, it is not at all clear how these will be effected/coordinated. And, as mentioned elsewhere in this review, there is little/no discussion of how CGIAR expertise will be used to inform the Foresight and Metrics Initiative model development and analyses.</p>
From Fragility to Resilience in Central and West Asia and North Africa (F2R-CWANA)	<p>At one level, F2R CWANA’s Work Package-level theory of change diagrams are quite effective in depicting how elements of this Initiative interact with other CGIAR Initiatives. On the other hand, like so much of this proposal, the text is not really clear about how these are articulated with the broader CGIAR portfolio. More important than all those arrows on the diagrams, and as emphasized in the commentary on specific review criteria, the absence of a coherent theory of change for this Initiative. This raises questions about synergistic links with the entire One CGIAR portfolio. One can imagine how some of the relevant Genetic Innovation Initiatives might contribute germplasm to F2R CWANA (though this only is touched upon in the proposal and without specifics), but it is difficult to see the potential for flows in the opposite direction or even potential for the generation of regional or international public goods. One exception here might be the interactions described with NEXUS Gains, though that seems confined to Work Package 4. The</p>

¹⁰ The responses received light, technical editing for understanding and clarity.

Initiative	Does the Initiative Align with the Cohesion of the Portfolio as Described in the Companion Document? ¹⁰
	<p>Companion Document stresses (on p. 12) the “Impact Area Platforms” in generating “impact-oriented cohesion” (which would be great!), but the very weak treatment of trade-offs/synergies across Impact Areas and absence of ex-ante assessment of scaling of benefits to impacts (discussed under criterion 6) gives neither a logical rationale nor a basis for confidence.</p>
<p>Fruit and Vegetables for Sustainable Healthy Diets (FRESH)</p>	<p>Within the proposal, there is some evidence of internal and external coherence, interdependencies, and management of uncertainties, though they need to be strengthened in parts. These are highlighted throughout the review. Resilience on the ground is highly dependent on local skills and knowledge of an integrated set of stakeholders to adapt to changing food systems and their environments.</p> <p>This implies capacity building must be strategic and well targeted, along with adequate and sustained contribution of financial resources (commitment) provided by national governments.</p> <p>Internal coherence and interconnectedness with other Initiatives: There is some evidence presented regarding various Initiatives FRESH would work with, but the proposal does not articulate how. The recognition of interdependencies across Initiatives is currently weak. As commented above, it must be deliberate and regularly reviewed before rolling-out all the relevant Initiatives. The risk of overlap, duplication, and lack of joined-up thinking is high, but this risk could be mitigated if such efforts are made before the roll-out of FRESH and the relevant Initiatives.</p> <p>External coherence: While there is some evidence of stakeholder consultations, the proposal is weak on how it is positioned vis-à-vis several other strong global and regional research programs, especially related Work Package 1 and 5 and what its comparative advantage is.</p>
<p>Harnessing Digital Technologies for Timely Decision-Making across Food, Water, and Land Systems</p>	<p>External coherence (country, regional, global levels): This aspect is overall very good, although some reviewers questioned the choice of criteria used to select the target countries, as well as the problem definition through consultation with partners in the United States, Europe, and Australia (i.e., non-target countries). Greater emphasis on end-user demand and needs rather than top-down suppliers (such as agtech companies) will be critical for success of DX1.</p> <p>Internal coherence including Impact Area Platforms: While the Work Packages were clearly defined, the inter-linkages between them could be more announced. The linkages between the research plan and the proposed methods also seem vague: this sometimes leads to a lack of continuity between the research plan, scientific methods, and outputs. Internal coherence could be improved by stating how methods and outputs from each Work Package will feed into the other. In terms of management, there is a need for a person in the Integrated Management Team to have oversight of all Work Packages to provide coordination and consistency (assuming this role does not already exist).</p> <p>Interdependencies between other thematic/regional Initiatives: By definition, DX1 is more focused on methods (i.e., digital technology) rather than outcomes (generation of positive economic, environmental, social outcomes e.g., see Outcome boxes in each Work Package). As such, there is a tension between the creation of outcomes and the creation of frameworks (technology development and adoption) used to catalyze positive outcomes. To clarify what outcomes will be led by DX1 relative to those generated in other Initiatives (e.g., ClimBER, NEXUS Grains, MITIGATE+), the authors could add a column to the End of Initiatives outcomes in the tables on p. 46 that clearly delineate which outcomes are led by DX1, and which come from other Initiatives. This</p>

Initiative	Does the Initiative Align with the Cohesion of the Portfolio as Described in the Companion Document? ¹⁰
	<p>will help proponents better shape their intentions towards creation of beneficial outcomes for society, whether it be through alliances or from research within DX1 per se.</p> <p>Management of funding uncertainties: This aspect was completely absent from the proposal. Perhaps this deficit is a short coming in the proposal template design.</p> <p>Integrated results framework at Initiative, Action Area and CGIAR Levels: In general, this is well covered.</p> <p>Measurement and reporting at multiple levels and timeframes: Ex-post impact assessments of Work Packages are covered, but ex-post assessments of the monitoring and evaluation (and how this assessment will be used to refine management going forwards) could be more transparent. Linkage of management to the Work Packages could be more detailed (metrics, aim, achievability and timing). The MELIA plan is reasonable overall, although it is difficult to judge how and where baselines will be collected (e.g., p. 54). "Causal impact evaluation" is relegated to the last year of the Initiative (e.g., p. 54); perhaps this would be better incorporated throughout the implementation process.</p> <p>The Measurable End of Initiative outcomes for the "digital divide" challenge area seem to focus more on strengthening digital agrifood systems, rather than the digital divide per se, i.e., the people who have limited or nil access to mobile internet, digital services etc.</p>
<p>HER+: Harnessing Gender and Social Equality for Resilience in Agrifood Systems</p>	<p>From what is presented, this important Initiative is well aligned and integrated with the CGIAR Impact Areas, the SDGs and other CGIAR Initiatives.</p> <p>External cohesion: While HER+ has demonstrated some consultative processes in designing this Initiative, the depth and nature of these is unclear. External coherence could be addressed by articulating how HER+ is positioned and complements other Initiatives at global and regional levels. It also does not articulate its comparative advantage vis-a-vis other global players on gender or climate change research and in STIB development and delivery.</p> <p>Interconnectedness of thematic and regional integrated Initiatives: There is some evidence presented regarding various initiatives HER+ would work with. The recognition of interdependencies across initiatives is currently weak. As commented above, it must be deliberate and regularly reviewed during the implementation.</p>
<p>NATURE+: Nature-positive Solutions for Shifting Agrifood Systems to More Resilient and Sustainable Pathways</p>	<p><i>Lateral cohesion:</i> This Initiative is part of the overarching results framework that comprises the 32 Initiatives that will be deployed by the CGIAR. Many of the scientists participating in this Initiative will also participate in other CGIAR Initiatives, and several Initiatives will be implemented in exactly the same target regions (and likely with the same households and communities). <u>However, the messages conveyed by these different Initiatives are often contradictory.</u> This will create confusion among the partners and beneficiaries on the ground. If CGIAR wants to take up NPSs as its main approach to agricultural innovation and rural development, then the entire portfolio of 32 Initiatives should follow a nature-positive approach. It is not advisable to promote different practices that may contradict each other (at all, but specially not) in the same locations.</p> <p><i>Vertical cohesion:</i> The Initiative is structured on the basis of the Theory of Change framework and hence it aligns well with the CGIAR result framework (cf. Fig. 6 and Annex 1 in companion document) which is basically an application of the classical Theory of Change. However, this model works well when Initiatives are top-down and/or unidirectional, moving from activities, to outputs, outcomes, and impacts, assuming causalities between these steps and an if-then logic. This model is however not always applicable—almost never—to co-innovation approaches, as proposed by this Initiative, which are bottom-up, adaptive, and where the role of researchers</p>

Initiative	Does the Initiative Align with the Cohesion of the Portfolio as Described in the Companion Document? ¹⁰
	<p>and development agents is to broker knowledge, to create a dialogue to support self-investment, a sense of ownership and a risk-taking attitude by local actors. Problems are addressed as they emerge, and solutions are developed through trial and error, experimenting together between different actors of a platform. This adaptive dynamic, inherent to managing projects as complex adaptive systems, cannot be captured in the classical Theory of Change structure.</p>
<p>Resilient Aquatic Food Systems for Healthy People and Planet</p>	<p>External coherence: The Initiative is clearly demand driven. The proposal demonstrates a strong commitment to engagement of partners, many of which are identified. But their collaboration is frequently treated as an assumption and there is limited information on how this will be practically achieved. RAqFS has dedicated Work Packages on partnerships and AquaLabs (national innovation platforms) that provide useful platforms for developing and exploring mechanisms for coherence and sustainability.</p> <p>Internal coherence: This is apparent in the overall Theory of Change, which reflects the complementarity of the five programs as well as the anticipated collaboration amongst Work Packages (which is stated rather than described). Two of the Work Packages (3 and 5) offer strong integrative potential at a landscape or national level. It is not clear to what extent this potential will be realized in view of limited overlap in the choice of focal countries.</p> <p>Interdependencies: RAqFS is a partnership between five CGIAR entities. RAqFS anticipates collaboration with seven System Transformation Initiatives (AE-I; SHiFT; NPS; HER+; MITIGATE+; Markets and Value Chains; Foresight; Digital Technologies), two Resilient Agrifood Systems (Resilient cities; One Health) and four RII (AMD; ESA; TAFS-WCA; TAFSSA). Several of the individual work programs specify the intended collaboration but do not describe this in detail. National consultations in some target countries were coordinated with other Initiative Design Teams. The Initiative clearly supports the CGIAR Performance and Results Management Framework and includes cross-cutting budgets and human resources organized around each of the Impact Areas.</p>
<p>Resilient Cities Through Sustainable Urban and Peri-urban Agrifood Systems</p>	<p>Yes, to some extent. Certainly, urban food systems resilience is a key component of the Resilient AgriFood Systems Action Area, and this proposal has the potential to be synergistic with a range of Initiatives, ranging from those relating to livestock, fish, and vegetables, to the markets and value chains and sustainable healthy diets programs, on gender, as well as the regional ones. That said, these potential synergies are not described in any detail, and are mostly restricted to scattered mention in the Work Package description tables. There are many areas where strong partnerships with related Initiatives from the outset could exploit complementarities and result in more efficient use of resources.</p> <p>One example relates to Work Package 4 as described elsewhere in this report—a close collaboration on diets and food environments with other Initiatives would ensure resources are used efficiently across the portfolio. Partly this reflects excessive focus on UPU primary agricultural production, although UPU agrifood systems feature primarily consumers and post-harvest workers in the value chain. Even in lower-middle income countries, a sizeable majority of consumer food expenditures accrues to actors' post-harvest and that share is surely highest in UPU. Paying appropriate attention to the broader entry points and impacts of UPU agrifood systems will require careful coordination with country and regional offices, with the markets and gender global initiatives (and others), to devote sufficient attention to these crucial components of the systems. This is not historically an area of CGIAR strength, so likely needs more creative and far-reaching partnerships but upstream (to ARIs) and downstream, and especially with private sector actors, including large firms that heavily influence UPU agrifood systems.</p>

Initiative	Does the Initiative Align with the Cohesion of the Portfolio as Described in the Companion Document? ¹⁰
Securing the food systems of Asian Mega-Deltas for climate and livelihood resilience (AMD)	<p>There was some surprise that the potential for building on CGIAR’s past and existing programs were not explicitly made, e.g., WorldFish, IWMI, ILRI, as well as AVRC, or appreciation of and stronger program with alignment with the partner country existing policies. Linkages to other Initiatives e.g., Resilient Aquatic Food Systems were mentioned but in not enough detail to appreciate any integration of effort or the specific opportunities for complementarities or synergies.</p>
Sustainable Intensification of Mixed Farming Systems (SI-MFS)	<p>Cohesion of the CRP portfolio was an important aspiration during Phase II. However, the level of cohesion was limited due to lack of alignment among individual CRP priorities. Lessons must be learned from this to inform One CGIAR as portfolio integration is also very important. Figure 3 (Companion Document) shows how the planned Initiatives will link together for cohesion. Unfortunately, this only includes the first batch of Initiatives and does not include SI-MFS. The narrative of SI-MFS indicates that it will link with RIs, EiA, PHI, and HER+. It does not mention links with seed systems or plant breeding initiatives—SeEdQual and ABI—or with climate mitigation strategies—MITIGATE. Such links are just as important as links with agronomy and pest management for sustainable intensification in farming systems. It should also be acknowledged that links will entail transaction costs and there is likely to be competition among initiatives. How will priorities be decided?</p> <p>The management structure detailed in the Companion Document is very complex and most importantly, lacks clear hierarchy or lines of authority which could lead to delays in making decisions. There is a need for clear prioritization principles and guidelines as well as careful oversight on final decisions.</p>
Transforming AgriFood Systems in West and Central Africa (TAFS-WCA)	<p>No doubt that the Initiative includes acceptable level of coherence expected in a quality document. The content aligns with requirements for priority setting both at a national and community level, despite the blanket and generalized arguments in various sections of the document. The partnership in the consortium and targeted countries are relevant, although some countries seem to be already spoiled with international funding. The integrated approaches could be improved but there are ingredients of a good portfolio platform reflecting the Common CGIAR policies.</p>



Advisory
Services

CGIAR Advisory Services – ISDC

Via di San Domenico 1, 00153 Rome, Italy

Email: isdc@cgiar.org

URL: <https://cas.cgiar.org/isdc>