Advisory Services

CGIAR

CGIAR Research Program 2020 Reviews: WHEAT

Authors: Victor Sadras, Donna Podems

Date: 3 July 2020

This evaluation has been commissioned by the CGIAR Advisory Services Shared Secretariat (CAS Secretariat) Evaluation Function (CAS/Evaluation). CAS/Evaluation encourages fair use of this material provided proper citation is made.

Recommended citation: CAS Secretariat (CGIAR Advisory Services Shared Secretariat). (2020). *CGIAR Research Program 2020 reviews: WHEAT*. Rome: CAS Secretariat Evaluation Function https://cas.cgiar.org/

Cover image: Farmer in wheat field, Mexico (Cimmyt, X. Fonseca)

Design and layout: Luca Pierotti and Macaroni Bros

CGIAR Research Program 2020 Reviews: WHEAT

Authors: Victor Sadras, Donna Podems

Date: 3 July 2020

Acknowledgments

The authors thank Hans-Joachim Braun, Victor Kommerell, Olaf Erenstein, Sarah Hearne, Kevin Pixley, Ravi Singh, Michael Baum, Bruno Gérard, Barbara Rischkowsky, Vinay Nangia, Mina Devkota, Rachid Moussadek, and Jacques Wery for their time, insights, and fact checking; Tony Fischer, John Porter, Richard Trethowan, Gustavo Slafer, Marta Monjardino, Mariano Cossani, Daniel Calderini, and Eric Huttner for sharing their perspectives; Allison Grove Smith, Ravi Ram, Paolo Sarfatti, and Ángela Giménez Barrera for coordination and management of the review; Gaia Gullotta, Max Runzel, Julien Colomer, and Enrico Bonaiuti for analytical support; and Shaylyn Gaffney, David Lara González, and Hamish Dunsford for logistics.

Disclaimer

Contents

Annexes are provided at the following link: https://bit.ly/WHEAT-CRP2020-ANNEX

List of Tables

List of Figures

Abbreviations

agR&D	Research and development in agriculture
AR	Annual report
BPAT	Breeding Program Assessment Tool
CA	Conservation agriculture
CAIGE	CIMMYT Australia ICARDA Germplasm Evaluation
CAS	Secretariat CGIAR Advisory Services Shared Secretariat
CoA	Cluster of activities
CSAP	Climate-smart agricultural practice
CCAFS	CGIAR Research Program on Climate Change, Agriculture and Food Security
CIMMYT	International Maize and Wheat Improvement Center
CRP	CGIAR Research Program
DGGW	Delivering Genetic Gains in Wheat
EiB	Excellence in Breeding Platform
FP	Flagship program
FTE	Full-time equivalent
GENEBANK	Genebank Platform
GENNOVATE	Global comparative gender norms research initiative
GWAS	Genome-wide association study
h	Hirsh index
HeDWIC	Heat and Drought Wheat Improvement Consortium
ICARDA	International Centre for Agricultural Research in the Dry Areas
IDO	Intermediate Development Outcome
IF5	5-year impact factor
IWIN	International Wheat Improvement Network
LMS	Learning Management System
MARLO	Managing Agricultural Research for Learning and Outcomes Program
MC	Management Committee
M&E	Monitoring and evaluation
MEL	Monitoring, evaluation, and learning platform
MELIA	Monitoring, evaluation, learning, and impact assessment
MELIAG	Monitoring, evaluation, learning, impact assessment, and gender workshop
MENA	Middle East and North Africa
NARS	National agricultural research system
OIC	Outcome Impact Case
OICR	Outcome Impact Case Report

OWP	One CGIAR Global Wheat Program
POWB	Plan of work and budget
Q	Journal quartile
QoR4D	Quality of Research for Development
R&D	Research and development
RBM	Results-based management
SDG	Sustainable Development Goal
SJCR	SCImago Journal & Country Rank
SJR	Scientific journal ranking
SLO	CGIAR System-Level Outcome
SMO	CGIAR System Management Office
SPIA	Standing Panel on Impact Assessment
ToC	Theory of change
ToR	Terms of reference of the review
USAID	United States Agency for International Development
W1, W2, W3	Funding Windows 1, 2, and 3
WHEAT	CGIAR Research Program on Wheat
W-ISC	WHEAT Independent Steering Committee
W-MC	WHEAT Management Committee

Executive Summary

Background and Context

Globally, wheat contributes 20% of the calories and protein to the human diet. The projected increase in global demand between 2005-07 and 2050 is 44%, equivalent to a compound growth rate of 0.83% p.a. WHEAT targets a sustained increase in yield of 1.4% p.a. with equal contributions from breeding and agronomy, the two pillars of the program further supported by R&D efforts to enable adoption and subsequent scaling, of which policy is a significant part.

Purpose and Scope of the CRP 2020 Review

The review's purposes are to assess to what extent WHEAT is (1) delivering quality of science, and (2) demonstrating effectiveness in relation to its own Theories of Change (ToC). The third purpose is to provide insights and lessons to inform the program's future. The review was designed to deliver top-level findings proportional to the desk-top work of two reviewers for 11 weeks. The primary audience is the CGIAR System Council. Other audiences may find the review useful, such as WHEAT managers and scientists, their partners and other CGIAR research centers and CRPs.

Review Questions

The review asks (1) To what extent does the CRP deliver quality of science, based on its work from 2017 through 2019? (2) What outputs and outcomes have been achieved and what is their importance? (3) What is the evidence for future effectiveness within the life of the program (through 2021) considering the comparative advantages of the CRP and its FPs and drawing on the CRP and FPs progression according to their ToC?

Methods for the Review

Framework and data sources. The review used CGIAR's Quality of Research for Development (QoR4D) framework that defines and provides metrics for relevance, scientific credibility, legitimacy, and effectiveness. Data sources included documents and records provided by CAS and WHEAT, and surveys and interviews with managers, scientists and other stakeholders. To solve inconsistencies between data sources and to ensure robust findings, secondary sources and highly aggregated data were not taken at face value but used primarily as pointers to primary sources, which were then critically analyzed. CAS and WHEAT checked the review for factual accuracy.

Quality of science. We used Web of Science and SCImago Journal & Country Rank to assign journals to quartiles and to retrieve bibliometric indices. A subsample of publications was read and analyzed in detail to complement the quantitative analysis. We analyzed data from surveys, interviews and documents focusing on technological outputs, policy and capacity building.

Effectiveness. Data were gathered using desk review and interviews. The desk review used data sources that (1) described the program and what it is addressing, for whom, where, and why, (2) answered to the extent possible, the evaluation questions, and (3) revealed data gaps and data that needed triangulation, thus informing a focused research and data gathering pathway. Semi-structured interviews with CGIAR and WHEAT staff and other stakeholders complemented the desk review. Individual and group interviews were implemented depending on (1) type of information being asked, (2) timing, and (3) cultural appropriateness. Data were analyzed thematically against the evaluation questions.

Findings and Recommendations

Quality of Science

Key finding: With approximately 80 scientists, WHEAT is subcritical[1]. Over decades, WHEAT has catalyzed a global network of R&D that has delivered and continues to deliver a disproportionate wealth of outputs in relation to investment. Partnerships, and WHEAT reputation as a reliable partner, are vulnerable to funding volatility.

Recommendation: Support strategic investment in network development and maintenance.

Key finding: The network of partners might be scale free – i.e., driven by preferential attachment or the rich-get-richer, with implications for (1) inclusiveness – it may be hard for outsiders to enter the network, and (2) resilience – the system is resilient to random perturbation, but vulnerable to disruption of large nodes such as the centralized breeding system.

Recommendation: Investigate the nature of the network using larger samples and complementary metrics beyond authorship. Use this information to ensure network resilience in response to both random and targeted disruptions. Consider opportunities for expanding networks beyond current nodes.

Key finding: 79% of the scientific publications were in top 50% of journals; 21% of the publications in the bottom half is a symptom that needs attention. Publications in Quartile 1 journals (Q1, top 25%) were world-class and some pushed scientific boundaries; publications in Q2 journals were sound but often routine work; publications in Q3Q4 journals commonly featured flaws.

Recommendation: Set targets (time frame, rates) to shift a proportion of Q2 papers to Q1. Set targets (time frame, rates) to phase out Q3 and Q4 papers. Set up mentoring systems to avoid work that leads to lower quality papers in the first place. Revise evaluation and reward system to improve the quality-to-volume ratio of scientific output, i.e. fewer but better papers in the context of different stages in the scientist's career.

Effectiveness

Key findings: In 2017- 2019, WHEAT mainly achieved its planned outputs and outcomes, and in addition had unplanned outcomes. For the three years reviewed, WHEAT did not drop any research line. Slight changes included WHEAT adding research priorities on mechanization (FP4) and soil-borne diseases (FP3) in 2017.

Recommendation: Establish how WHEAT, or any CRP, will be assessed for effectiveness, when the proposal is submitted and approved, and set clear criteria for judging that effectiveness. Alternative suggestions for assessing WHEAT's effectiveness, which can also be used for improving the intervention, management decisions, and judgement are provided in the Future Orientation section.

Key findings: WHEAT's roles, responsibilities and accountabilities are clearly defined and exercised, and governance is sufficiently independent. Some slight challenges related to funding exist. Monitoring and evaluation had some challenges yet appear to adequately support the CRP. Redundant reporting and frequently changing requirements are burdensome.

Recommendation: Indicators and other approaches that use quantitative data need to continue, as they are expected and considered necessary by most donors. Reducing the focus on these, and increasing the use of case studies (which include numbers and words, or statistics and narratives) that focus on countries or themes (e.g. Ethiopia or mechanization) will likely bring broader perspectives and a more informed (deeper) understanding of WHEAT's outcomes, and its effectiveness. See Future Orientation.

Key Finding: "Gender" is used as a research focus, and "gender" is used to identify a research or capacity development beneficiary; gender achievements in research, while often notable, were often siloed.

Recommendation: Select a few key gender findings that are useful to specific FP interventions or research, integrate these findings, and ensure these gender aspects are included in effectiveness assessments. Require that gender statistics are collected and reported for all training, workshops, and

[1] Technological innovation is about finding new combinations of existing stuff – a current cell phone is a phone + GPS + camera + computer etc. The richness of the innovation system depends on two things: the diversity of resources to be combined, and the diversity in the rules for combinations. These two dimensions define subcritical and supractitical systems. WHEAT is subcritical but taps on the supracritical global technological system.

conferences; use statistics to better understand and improve participation levels where appropriate, not as a box-ticking exercise.

Key Finding: the CGIAR approved WHEAT's approach to Youth (along with gender) and results on youth reflect that agreement, with a notable achievement in 2017 that informs youth oriented R4D.

Recommendation: Use the 2017 research paper to consider how to bundle youth-relevant R4D across CRPs and provide a core budget.

Key finding: The ToC provides an explicit shared thinking about how change comes about in a larger context, and is useful for (1) priority setting, (2) assessing contribution of scientific outputs, (3) seeking and justifying funding, (4) mapping trajectory to impact and (5) reporting, but is unsuitable for (6) assessing WHEAT or its Flagships effectiveness by judging progress towards the SLOs.

Recommendation: Continue to refine the ToC as needed, and recognize its five uses, in its current form. Do not assess the CRP's progress towards the identified SLOs to judge WHEAT's effectiveness. Stated another way, it is not useful to assess how far "up the chain" WHEAT has identified accomplishments, rather explore what WHEAT accomplished within their selected areas of the ToC, how these areas were selected, how results contributed to the global effort, and why it's important.

Future Orientation

Key finding: Wheat as a crop is bound to be central to global food security in the foreseeable future. WHEAT as a R&D agent has a track record of delivering local solutions with a global perspective and is well positioned to continue this trajectory in the next decade. There are opportunities and challenges for the way ahead, including the risk of fragmenting the global breeding program; restrictions to exchange germplasm and ideas; the opportunity to integrate R&D in agronomy; misguided emphasis on minor crops; and CGIAR's focus on process at the expense of results.

Recommendation: Ensure support to both modernization of breeding process and integrated approaches to sustainable intensification including mechanization. For management, monitoring and evaluation with purpose, consider integrating three elements based on well-known social and scientific theories: (1) Principle-focused Evaluation to examine the extent to which statements of principles provide meaningful guidance, are useful in decision-making, are inspirational, support adaptation and development, and are evaluable; (2) Developmental Evaluation as a strategic learning tool that supports innovation and social change in complex or uncertain environments; (3) Alston's equation – evaluations before, during and after investments are critical but must not suffocate scientists.

CGIAR System-level Recommendations

Our core recommendations above targeted not only CRP but also donors and CGIAR higher-level management. Further, given that the CRP phase will end in 2021, little time remains for any shifts in CRP management, governance or resources/budgets. As one respondent reflected.

> *Everyone is waiting for One CGIAR, and we will then make shifts accordingly.*

1 Background to the CRP 2020 Review

1.1 Purpose and Audience of the Review

The terms of reference define the purpose and primary audience of the review (Annex 1). The review's purposes are to assess to what extent WHEAT is (1) delivering quality of science, and (2) demonstrating effectiveness in relation to its own Theories of Change (ToC). The third purpose is to provide insights and lessons to inform the program's future. The review was designed to deliver top-level findings proportional to the desk-top work of two reviewers for 11 weeks.

The review has three objectives: (1) To fulfil CGIAR's accountability obligations regarding the use of public funds and donor support for international agricultural research. (2) To assess the effectiveness and evolution of work in research programs under CRP 2017-2021. (3) To provide an opportunity for programs under review to generate insights about their research contexts and programs of work, including lessons for future CGIAR research modalities.

The primary audience is the CGIAR System Council. Other audiences may find the review useful, such as WHEAT managers and scientists, their partners and other CGIAR research centers and CRPs.

1.2 Overview of the CRP and Its Context in Research for Development

Section 1.2 provides an overview of (1) the global wheat scenario, (2) WHEAT, and (3) the management of research and development (R&D) resources.

1.2.1 Global Wheat Scenario

Globally, wheat contributes 20% of the calories and protein to the human diet (FAO 2016). Population growth, increasing income per capita, alternative uses of grain and commodity prices are driving an increase in demand of grain. The projected increase in global demand between 2005-07 and 2050 is 44%, equivalent to a compound growth rate of 0.83% per annum (p.a.) (Fischer et al. 2014). Increased wheat production to match increasing demand will remain critical for food security in the foreseeable future. For wheat, current rates of improvement are between 0.3% p.a. and 1.7% p.a. for farm yield, and between 0.3% p.a. and 1.1% p.a. for potential yield, with a yield gap between 26% and 69% (Fischer et al. 2014). Raising potential yield through breeding and closing the yield gap through sustainable intensification are critical to future increase farm productivity and profitability, and form WHEAT's two pillars.

1.2.2 WHEAT

The CGIAR portfolio includes two groups of research programs (CRP) and three Research Support Platforms launched in 2012 (Annex 2, Annex 3). is part of the Agri-Food Systems Programs led by CIMMYT, with ICARDA as a primary research partner. Other key partners and funders include the Australian Centre for International Agricultural Research (ACIAR), the British Biotechnology and Biological Sciences Research Council (BBSRC), the Indian Council of Agricultural Research (ICAR), and a community of more than 200 public and private organizations worldwide, among them national governments, companies, international centers, and regional and local agencies and farmers.

WHEAT targets six mega-environments where 90% of poor people earning less than $2 a day in wheat growing environments live and 85% of the wheat in developing countries is grown. WHEAT seeks to improve the livelihood of smallholders in wheat Agri-Food Systems against the backdrop of increasingly virulent biotic stresses, less water, more erratic rainfall and rising temperature. Based on its two pillars, germplasm improvement and sustainable intensification, WHEAT is composed of four Flagship Programs (FP) that operate clusters of activities providing: horizontal guidance to WHEAT (FP1), tools for improving genetic gains and breeding efficiency (FP2), improved varieties of spring bread, durum wheat, triticale and winter and facultative bread wheat (FP3), and wheat-systems agronomy to close yield gaps and improve the efficiency in the use of resources (FP4).

Annex 3 summarizes funding during the review period. Funding comprises three Windows (W). W1 includes pooled funding which may be used across the CGIAR System. W2 comprises pooled contributions targeting specific CRPs and/or platforms. W3/bilateral funding is allocated to particular CGIAR Research Centers through the CGIAR Trust Fund, and/or directly to specific projects in CGIAR Research Centers outside the Fund. In 2017-2019, W1W2 accounted for 21-35% of the total funding. The report does not distinguish funding sources in the assessment of quality of science and effectiveness.

1.2.3 Management of R&D Resources: Solving Tensions and Capturing Synergies

"Chemicals act on chemicals to produce new chemicals, goods act on goods to produce new goods, and ideas act on ideas to produce new ideas" (Hanel et al. 2005). This autocatalytic process drives the evolution of both natural and social systems (Figure 1 A). Technological innovation is about finding new combinations of existing stuff. However, owing to Darwinian pre-adaptations[2] the trajectory of these systems is unpredictable (Kauffman 2008; Kauffman 2016). The sample space of tossing a coin is defined; we do not know what will happen, but we know what can happen (tail or head), and work with probabilities. Instead, the sample space of technological innovation is undefinable.

Kauffman's observation is radical: it is not that we cannot predict what will happen, we cannot predict what can happen. This has implications for management of scientific research as we "cannot manage the discovery of the unknown" (Osmond 1995) but cannot afford expensive distractions like nitrogen-fixing cereals (Sadras et al. 2020). Alston et al. (1995) summarize this tension (Box 1).

Instead looking at curiosity- and utility-driven research as antagonistic, exploiting the synergies between them is central to effective management of R&D resources (Osmond 1995; Sadras et al. 2020). Owing to limited resources, competition is unavoidable; the tension between collaboration and competition is also important, and rarely.

Box 1. Alston's equation

There is a demonstrated need for formal economic evaluation of alternative investments and priority-setting procedures but …formal evaluation and priority-setting procedures should not be used as a basis for replacing ingenuity, serendipity and scientific entrepreneurship with bureaucratic procedures...

1.3 Scope of the Review

The review focuses on WHEAT 2017-2019 and assesses its quality of science and effectiveness. The review does not assess individuals, teams, or centers. Emphasis is placed on WHEAT's sphere of control including (1) the quality of inputs, activities and outputs, and (2) influence, defined as short and intermediate outcomes that are expected to lead to a development impact. The 3-year time-window for the review does not match the time frame for development and deployment of technology, in the order of decades (Hall and Richards 2013). For this reason, we have occasionally used metrics of performance that do not necessarily match the review's short time frame.

1.4 Review Questions

The ToR provided three questions[3], two of which are then further defined by sub-questions. These questions focus on quality of science, effectiveness, and future orientation: (1) To what extent does the

[2] Darwinian pre-adaptations preclude prediction of technological innovation; the screwdriver has an initial function (to screw screws), but can also be used to crack nuts or open paint tins. The number of functions of the screwdriver is undefinable.

[3] An initial ToR evaluation question focused on WHEAT's comparative advantage, which is not addressed in this review because it has already been answered comprehensively: "The comparative advantage of WHEAT resides on its global expertise and high-quality research innovation, as noted in the first CGIAR Stakeholder Perceptions Survey. WHEAT involves more than 250 partners worldwide with an excellent track record…"(2015 ISPC Commentary on the WHEAT Phase-II –Preproposal 2017–2022, p. 2).

CRP deliver quality of science, based on its work from 2017 through 2019? (2) What outputs and outcomes have been achieved and what is their importance? (3) What is the evidence for future effectiveness within the life of the program (through 2021) considering the comparative advantages of the CRP and its FPs and drawing on the CRP and FPs progression according to their ToC?

1.5 Method for the Review

We used CGIAR's QoR4D Framework (ISDC 2020; ISPC 2017) that defines and provides metrics for relevance, scientific credibility, legitimacy with a focus on partnerships, and effectiveness in terms of significance of research in the context of ToC. This section describes the methods to assess quality of science and effectiveness, and to outline future orientation. We sourced data and information from documents provided by CGIAR Advisory Services Shared Secretariat (CAS Secretariat) and WHEAT, and from interviews and surveys with managers, scientists and other stakeholders (Annex 5). To solve inconsistencies between data sources (Annex 9) and to ensure robust findings, secondary sources and highly-aggregated data were not taken at face value but used primarily as pointers to primary sources, which were then critically analyzed. Further, CAS Secretariat Evaluation Function and WHEAT checked the review for factual accuracy.

1.5.1 Quality of Science

1.5.1.1 Using Deficient Indices to Quantify Quality of Science

The myriad of indices[4] for the scientific evaluation of individuals, journals, institutions and countries reflects both the multidimensional nature of scientific quality and its elusiveness. The numerous indices also reflect their imperfection; no single index is complete. Furthermore, bibliometric indices can be subject to dishonest manipulation (Wilhite and Fong 2012). The peer-review process and peer-reviewed journals are far from perfect, featuring for example conflicts of interest and gender bias (Budden et al. 2008; Neff 2020).

How do we justify using deficient indices and imperfect journals to rate quality of science? Consider this scenario. As a donor, you seek to invest on a new breeding platform. Platform A is published in *Genome Biology* (1st Quartile; 5-year Impact Factor = 18.36; index h[5] = 207). Platform B is published in *Plant Breeding* (2nd Quartile; 5-year Impact Factor = 1.59, index h = 63). This is all the information you have; make a choice. Despite all its limitations, papers in peer-reviewed journals are central to the credibility of science.

1.5.1.2 Data Sources and Analyses

We analyzed WHEAT journal publications from 2017 to 2019 using the lists provided by WHEAT. The databases were inconsistent between years and between centers. For example, ICARDA 2018 only reported the senior author, and did not include a paper ID; CIMMYT 2018 attributed all papers to a single FP. Data bases included duplications. We solved main inconsistencies and gaps through consultation with FP leaders. As the review focuses on the program, CIMMYT and ICARDA publications were pooled in the analysis.

We used Web of Science (WoS) and SCImago Journal & Country Rank (SJCR) (Annex 6) to assign journals to quartiles using the criterion outlined in Table 1. We retrieved the 5-year impact factor (IF5) from WoS and both index h (Hirsh 2005) and SJR from SJCR. The IF5 is the ratio between the number of citations in the JCR year and the total number of articles published in the five previous years. SJR is the average number of weighted citations received in the selected year by the documents published in the journal in the three previous years.

[4] Journals are scored by Total Cites, Impact Factor, 5-Year Impact Factor, Immediacy Index, Eigenfactor Score, Article Influence Score, Normalised Eigenfactor, Citable Items, % Articles in Citable times, Cited Half-life, and so forth.
[5] Index h, defined as the number of papers with citation number higher or equal to h (Hirsh 2005).

Table 1. Quartile definition

JCR® Category	Rank in Category	Quartile in Category
Agronomy	25 of 89	Q2
Genetics & Heredity	123 of 174	Q3
Horticulture	6 of 36	Q1
Plant Sciences	92 of 228	Q2

Note: The definition of quartiles depends on the category where the journal is classified. The journal Molecular Biology is Q1 in Horticulture, which is unsuitable for WHEAT and is Q3 in Genetics and Heredity, where it competes with medical journals such as Genome Medicine with inherently higher impact. Molecular Biology is Q2 in Agronomy and Plant Sciences. Hence, we used Q2 for this journal as the most suitable for WHEAT.

We complemented bibliometric analysis with FP surveys and interviews. A semi-structured questionnaire was emailed to FP leaders to capture their perspectives and to address information gaps (Annex 7). We then thematically analyzed those data against the evaluative framework. We also interviewed scientists, breeders, and donors to gather additional perspectives.

1.5.2 Effectiveness

Empirical evaluations require a theory; otherwise they are just research with an opinion attached or, at worst, a haphazard process. Theory informs the methods chosen, the decisions made in the field, how data are analyzed, and, importantly, how an intervention is valued. Here we addressed the effectiveness questions using a fidelity approach to evaluation, guided partially by the theoretical framework of Utilization Focused Evaluation (Patton 2008).

Data were gathered using two methods: desk review and interviews. For the desk review we used data sources that (1) described the program and what it is addressing, for whom, where, and why; (2) answered, to the extent possible, the evaluation questions; and (3) revealed data gaps and data that needed triangulation, thus informing a focused research and data-gathering pathway. Semi-structured interviews with CGIAR and WHEAT staff and other stakeholders complemented the desk review (Annex 8). Individual and group interviews were implemented depending on (1) the type of information being requested, (2) timing, and (3) cultural appropriateness. All interviewees were provided with a consent statement stating that (1) their name will appear in the report unless they request otherwise, and (2) they will not be referenced or quoted in the document.

We ensured data saturation (e.g., the same information is found repeatedly, thus suggesting a solid finding) through interviews and document reviews. Data were analyzed thematically against the evaluation questions. There were no specific valuing criteria applied to the evaluation, per CGIAR guidance (see also section 1.8).

1.5.3 Future Orientation

In 2021 2021, WHEAT will transition to a new, yet unknown, modality or modalities. The next two decades are critical for global food security (Fischer and Connor 2018) and the incipient One CGIAR is a likely disruptor in how researchers and stakeholders engage with that challenge. We therefore set a decadal perspective to address the question: to what extent is WHEAT positioned to be effective in the future, from the perspectives of both scientists and end users of R&D outputs. Our analysis was based on data gathered from interviews and surveys (Annex 7, Annex 8).

1.6 Quality Assurance

To ensure the suitability and reliability of data and analytical processes, we used transparent methods with explicit assumptions and limitations (section 1.5) and further criteria presented in Annex 9.

1.7 Organization of the Review Team

Victor Sadras focused primarily on quality of science, and Donna Podems focused primarily on effectiveness. They worked as a team to ensure the integration, coherence, and robustness of the review. CAS and WHEAT provided sustained support to the authors. The authors declare no conflict of interest (Annex 15).

1.8 Limitations

We identified several limitations, including lack of an inception report, lack of a valuing framework, unclear evaluation questions, reliance on secondary data, contradictions between data sources, and inconsistent reporting compromising statistical analyses. For example, the 2018 database for training included "Mexican", "MEXICANA" and "Mexico" as entries for nationality; the original bibliographic dataset had truncated data and duplications; graphs of network analysis used outdated data. Annex 9 explains our approach to ensure quality of data. The terms of reference of the review feature gaps, redundancies, and flaws; for example, the purpose does not mention quality of science, and the identified evaluation approaches are not entirely appropriate (Annex 1). Details of limitations and mitigation strategies, to the extent possible, are provided in Annex 10.

2 Findings

2.1 Quality of Science

2.1.1 Quality of Research Inputs

2.1.1.1 Scientific and Supporting Staff

Key finding: WHEAT comprises ~80 scientists, with a ratio of technical-to-scientific staff from 8.9 to 2.1 among flagship programs, and a ratio of administrative-to-productive staff from 0.04 to 0.10; these ranges are commensurate with the focus and geographical spread of staff of the flagships.

The suitability of resources available to meet outputs and outcomes was rated on a scale from 1 (unsuitable) to 5 (suitable). All FPs, except FP4 led by ICARDA, reported suitable scientific and technical staff to achieve WHEAT's goals, with a score of 4–5. Recovering scientific critical mass is a challenge for ICARDA's agronomy team after leaving Syria. ICARDA is gradually rebuilding a team focusing on field-scale agronomy research; scaling of field-research outputs is incipient. Owing to political instability, programs developed over years have been reduced or discontinued in MENA countries. In some national agricultural research system (NARS) partners, the lack of qualified scientists is only partially related to political instability.

For WHEAT, the number of full-time equivalent (FTE) staff was stable in the period 2017-19 and averaged 82 scientists, 543 technical, and 54 administrative staff (Annex 3), with an additional 5 FTE providing transversal support. FP3 was the largest, and FP1 the smallest. The ratio of technical-to-scientific staff varied from 8.9 in FP3 to 2.1 in FP1, which is consistent with the nature of their research (section 1.2.2). The ratio of administrative-to-productive (scientific + technical) staff averaged 0.09 and varied from 0.04 in FP2 to 0.10 in FP1 and FP3 (Annex 3). The geographical spread of staff explains the difference in administrative-to-productive ratio between flagships: FP2 has staff only in Mexico, whereas the other flagships are regionally spread and thus require local support. This ratio is useful to evaluate administrative overheads (Boon and Wynen 2017; Tainter 1988).

2.1.1.2 Infrastructure, Data Management, and Amount and Predictability of Funding

Key finding: WHEAT's capacity to set its own R&D agenda is constrained by high reliance on bilateral funding and is subject to shifts in donor interests.

All FPs rated the suitability of infrastructure highly, with the exception of ICARDA's component for the reasons outlined above. All FPs share concerns with (1) insufficient and unpredictable funding, (2) shifting donor priorities, (3) limited recognition of the efforts and investment needed to move from proof of concept to mainstream use of resources, and (4) a low (approximately 1:4) ratio between strategic and core funding (Annex 3). Uncertainty in funding constrains WHEAT's capacity to set its own R&D agenda, and this is compounded by the premature ending of CRPs and organizational uncertainty in the transition to One CGIAR.

In 2017–2019, W1W2 accounted for 21–35% of total funding (Annex 3). Bilateral/W3 funding accounted for the balance and varied from U$S52.5M in 2017 to U$S29.4M in 2018. With increased reliance on highly variable W3 funding (Annex 3), priority setting is shifting, with some donors becoming more involved, often to a micromanagement level, to define not only what is done but how projects are implemented. Recently some funders started to allocate W2 to specific flagships within a CRP, further affecting the capacity of the WHEAT MC to allocate funds based on CRP priorities. We further analyze the implications and management of volatile funding in section 2.2.3 in the context of governance.

For FP2 and FP3, capacity to fully explore and leverage large, complex data sets has been a limiting resource. Funding from the Bill and Melinda Gates Foundation (BMGF) and CRP Maize & Wheat for the Enterprise Breeding System (EBS) overseen by EiB is a step forward in resolving it. The first version is currently being rolled out. Data have been stored in Dataverse and are waiting to be transferred into EBS. It is unclear to what extent transversal platforms for data management support FP1 and FP4.

2.1.2 Quality of Process

2.1.2.1 Planning, Documentation, and Monitoring

The 21–35% funding from W1W2 supports essentials to achieve overall WHEAT objectives and is formalized in the program of work and budget (POWB). However, there is a mismatch between the mid- to long-term nature of agR&D from inception to delivery (Hall and Richards 2013) and the annual time-step for planning, budgeting, and reporting W1/W2 outputs and outcomes. Research outputs are documented and monitored at the project level with MARLO and summarized in the CGIAR annual reports, donor-requested reports, technical notes, and scientific publications. Planning, documentation, and monitoring are further explored in section 2.2.3.

Key finding: CGIAR's focus on process risks compromising results.

The tension between processes and innovation has been outlined in section 1.2.4. with emphasis on Alston's equation: ex ante and ex post assessments, monitoring, reporting, and so forth are essential for both accountability and efficient investment of scarce resources, but they must not suffocate scientists (Box 2). Our survey (Annex 8) and interview data showed how these management processes are burdensome in CGIAR. There are CGIAR System reviews on WHEAT managed by the System Management Office, gender review, FP-specific reviews such as breeding program reviews (BPAT = Breeding Program Assessment Tool), funder-demanded project-specific reviews, and reviews like this one. Applying for funding and reporting are time-consuming activities that distract from the core business of scientific and technological innovation. Emphasis on process is apparent in the motivation toward One CGIAR featuring "…a superordinate management structure…" with emphasis on governance, operational structures, and processes across CGIAR[6].

WHEAT uses W1W2 funds to support three annual projects for younger staff whereby candidates submit a 2-page proposal and WHEAT colleagues vote to adjudicate the projects. More broadly, W1W2 allow for

Box 2. Test for decision-makers

The scenario: Cambridge, 1950s. Two young lads knock at the door of our lab. They are interested in our X-ray facility to investigate the structure of the salt of the deoxyribose nucleic acid. We ask them for a business case, an ex ante cost-benefit analysis, and how their project fits into the theory of change of our program. They respond, "…this structure has novel features that are of considerable biological interest…" (Watson and Crick 1953). Would we support them? Would they fit in our box-and-arrow scheme? Is the current climate conducive to support the next Watson and Crick? Can we afford not to?

strategic work canalizing scientific creativity that may be less attractive for bilateral funders. For example, work on wheat blast was initially funded by W1W2 and later attracted ACIAR and USAID funding (Mottaleb et al. 2018). W1W2 was used strategically to engage WHEAT FP1, FP3, FP4, and MAIZE FP2 with a global Biological Nitrification Inhibition Consortium initiated by the Japan International Research Center for Agricultural Sciences. WHEAT-implementing Centers CIMMYT and ICARDA apply, with some flexibility, a 20:80 rule to allow scientists about 20% time for exploratory research.

Historically, WHEAT scientists have enjoyed flexibility to explore the space beyond the limits of contractual obligations. Owing to burdensome management processes and funding volatility, they perceive that this flexibility is currently at risk.

Key Finding: Impact assessments require independent specialists and funding, but must not distract resources from R&D core objectives.

WHEAT is regularly requested to inform donors on high-level impacts of their outputs. These assessments require independent specialists and funding[7]. DNA fingerprinting provides an efficient and objective tool

[6] https://www.cgiar.org/news-events/news/turning-many-into-one-cgiar-network-restructures/; https://www.cgiar.org/news-events/news/whats-next-for-cgiar/

[7] The Standing Panel on Impact Assessment (SPIA) is an external panel of experts in impact assessment subject matter. The panel is supported by a Secretariat and a team of SPIA researchers working on institutionalizing collection

to quantify germplasm adoption (Dreisigacker et al. 2019)[8]. Comparable approaches are currently not available for agronomic technologies. WHEAT scientists and managers noted the merit of an independent unit to regularly assess adoption and impact across centers and programs. However, the opportunity cost of this unit has to be made explicit as it may distract resources from the program's core business—improving the livelihoods of wheat farmers and the productivity of their fields (section 1.2.4).

2.1.2.2 Ethics, Transparency, and Conflict of Interest

Centers draft, and MC and FP leaders implement, policies on ethics, transparency, and conflict of interest. WHEAT addresses these issues pragmatically. For example, the FP3 approach (1) uses Material Transfer Agreements to obtain and utilize non-CIMMYT/ICARDA germplasm with advice from the IP unit, (2) promotes genetic resistance to manage diseases and pests and hence is not involved in promoting chemical use, (3) shares germplasm and results freely to all partners as per CGIAR open access policies, (4) publishes research results in peer-reviewed journals, (5) encourages and supports visits of scientists and other stakeholders to WHEAT facilities, and (6) stores data in open access systems.

2.1.2.3 Partnerships

Key finding: With approximately 80 scientists, WHEAT is subcritical. Over decades, WHEAT has catalyzed a global network of R&D that has delivered and continues to deliver a disproportionate wealth of outputs in relation to investment. Partnerships, and WHEAT reputation as a reliable partner, are vulnerable to funding volatility.

Portugal is a subcritical economy but becomes supercritical in the European Union (Figure 1A). With about 80 scientists (Annex 3), WHEAT is subcritical—hence its reliance on partnerships. WHEAT has a track record of partnerships with purpose. The International Wheat Improvement Network (IWIN) tests wheat genotypes in a network of nurseries in over 90 countries. Breeding, directed toward 12 mega-environments (6 for spring wheat and 6 for facultative and winter wheat), is conducted at strategic hubs to develop around 1,000 high-yielding, disease-resistant lines that are delivered annually as international public goods. Excellence in Agronomy and Excellence in Mechanization are incipient initiatives that illustrate how partnerships bring about trademarks of WHEAT—foresight, advanced science and technology, targeted end users, and focus on impact (section 2.4). Partnerships are the core competitive advantage of WHEAT as reflected in its commitment to allocate part of annual W1W2 to partner grants (Annex 3). Partnerships, and WHEAT's reputation as a reliable partner, are thus vulnerable to funding volatility.

What type of partnership?

Key finding: The network of partners might be scale free – i.e., driven by preferential attachment or the rich-get-richer, with implications for (1) inclusiveness – it may be hard for outsiders to enter the network, and (2) resilience – the system is resilient to random perturbation, but vulnerable to disruption of large nodes such as the centralized breeding system.

Partnerships can be formal or informal, contributions of partners can be financial or in-kind, and their aims are diverse—e.g., strategic thinking, sharing of personnel or facilities for research, political engagement, or funding. An analysis of all these dimensions, and how they influence the efficiency of WHEAT, is beyond the scope of the review. Next we focus on the network's degree distribution and its implications for inclusiveness and resilience.

Networks feature nodes, such as collaborating scientists or institutions, and edges representing their interactions (Figure 1B). For a network's degree distribution, scale free implies a power-law distribution and vice versa (Bak 1996; Broido and Clauset 2019). This means many small nodes and few large ones, returning a linear relationship in a log-log scale. We used this statistical property of networks to assess the partnerships (Figure 1C). Our preliminary analysis with an extremely small sample suggests that

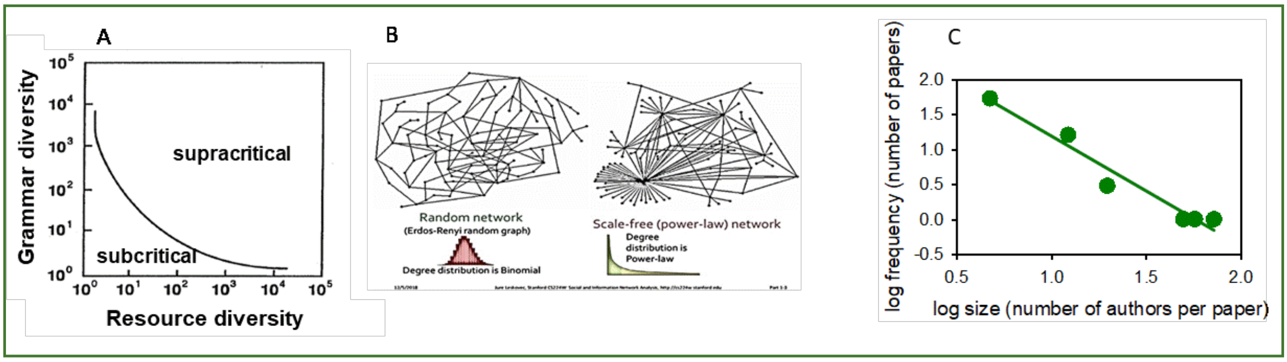
---

of data on CGIAR innovations in national data systems. SPIA works with Impact Assessment Focal Points (IAFPs), nominated by each Center and CRP https://cas.cgiar.org/spia/team. WHEAT competes for SPIA grants for impact assessments, generally resulting in a publication. CAS independent evaluation support for WHEAT comes in the form of the current Phase II review and external reviews of proposals, which result in reports.

[8] Of 560 samples collected from Afghan farmers' fields in 2015-16, 74% were identified as varieties released after 2000, which was more than the number reported by farmers and indicates the general prevalence of use of improved varieties, albeit unknowingly. Farmers correctly identified the variety they were growing in only 59% of cases.

WHEAT partnerships might conform to a scale-free network (Figure 1C), where preferential attachment is a possible driver (Barabási and Albert 1999). Preferential attachment is a process whereby some quantity (e.g., funding, papers) is distributed among partners according to how much they already have—the rich get richer. We can only speculate on the nature of the WHEAT network, but its implications for inclusiveness warrant a deeper assessment—it may be difficult for outsiders to enter the network. The nature of the WHEAT network also has implications for robustness—a scale-free network is robust to random perturbations such as volatile funding but is vulnerable to strategic mistakes such as decentralization of the breeding program.

Figure 1. WHEAT network



(A) Diversity of resources and diversity of grammar (rules for combinations) define a space whereby a hyperbolic curve divides subcritical and supercritical states. WHEAT is subcritical, hence its reliance on partnerships. (B) Random networks and scale-free networks feature characteristic degree distributions. (C) WHEAT network of authors of scientific publications conforms to a power-law degree distribution. Sources: (A) Kauffman (2008), (B) http://web.stanford.edu/class/cs224w/, (C) our analysis.

The tension between competition and collaboration

Key finding: The tension between competition and collaboration is implicit in WHEAT; making the tension explicit may improve integration.

In the period 1994–2014, the public sector accounted for 63% of global wheat varietal releases; this share ranged from 50% in the EU and high-income countries to 99% in South Asia (Lantican et al. 2016). The predominantly public nature of wheat R&D favors collaboration compared with other industries, and WHEAT features a solid network of partnerships. However, limited resources make competition unavoidable, and this can dampen collaboration at all levels—between scientists in the same FP, between FPs, between CRPs, and beyond.

Management of this tension is important and is not explicit in the program. WHEAT leaders have a mostly positive view of their collaborative performance. Improved collaboration between ICARDA and CIMMYT through joint funding applications (e.g., PRIMA on sustainable intensification in North Africa; IFAD on conservation agriculture in North Africa and Latin America) and joint appointments linking CIMMYT, ICARDA, INRA, and Morocco's Mohamed VI Polytechnic University justify this view. In contrast to the mostly positive view regarding competition expressed by WHEAT leaders, some outsiders shared a perception of competition that needs to be made explicit for improved collaboration. For example, some outsiders perceive some duplication of effort between FP2 and FP3.

The tension between collaboration and competition is apparent in the release of varieties bred locally versus selected directly from international trials and nurseries. However, the long history of collaboration between CIMMYT and ICARDA (pre-dating WHEAT) and WHEAT's national partners has fostered a sense of belonging to IWIN that permits release of best varieties irrespective of origin. International nursery testing is a cornerstone of germplasm evaluation that delivers elite lines for NARSs to use as parents or variety release whereas NARSs share their data to inform WHEAT's next crossing cycle. CIMMYT and ICARDA do not ask for name inclusion on variety release papers as the NARSs make the release decision. The One Global Wheat Program (OGWP) has helped to reduce overlaps between CIMMYT and ICARDA, but these efforts are still ongoing and are being hindered by the uncertainty about the future of WHEAT.

CAIGE (CIMMYT Australia ICARDA Germplasm Evaluation) is possibly the most effective of these synergistic collaborations, where Australian prebreeders and breeders benefit from WHEAT germplasm and information, providing in return timely, high-quality data from Australian trials to inform WHEAT breeding decisions. Rapport between leaders of participating organizations and recognition of mutual advantages is at the core of CAIGE's effectiveness. Currently, CAIGE engages five private breeding companies in a pre-competitive setting.

FP1 is the smallest flagship (Annex 3) and is particularly exposed to the perception of WHEAT as a breeding-centered program. To deal with this element of competition with its bigger siblings, FP1 has effectively partnered inside and outside WHEAT. Indeed, our bibliometric analysis shows FP1 features the highest proportion of cross-flagship collaborations (section 2.1.3.2).

2.1.3 Quality of Outputs

2.1.3.1 End Users and Outputs

Key finding: WHEAT and its flagships tailor outputs to well-defined end users.

All flagships clearly align key end users and critical outputs with their ToC as outlined in the WHEAT Handbook (2017) and further analyzed in section 2.2.1. FP1 primary end users are internal colleagues and management. FP2 end users are FP3 breeders, and national and international partners with aligned activities in R4D. FP3 primary end users are national partners (both public and private), and some outputs target policymakers (see below). FP4 targets farmers by scaling up and out with NARSs, development organizations, and NGOs.

Key finding: FP leading scientists consistently rate generation of knowledge as a major output.

In a scale from 1 (lowest) to 6 (highest), all flagships rated generation of knowledge 5 or 6 (Annex 7). Consistently, scientific publications are regarded as important outputs. This is reassuring in the context of quality science driving technological innovation, and the dual motivation of pushing scientific boundaries and utility (section 1.2.4). FP1 also rates policy briefs and outreach highly. The main outputs for FP2 are technological innovations, including bridging germplasm, IT tools, methods, and standards. Elite lines for use by NARS breeders as parent or varieties are the most significant output for FP3. Breeders also emphasize policy and capacity building as these are bottlenecks for effectiveness—seed sector and variety release procedures are a major obstacle for replacement of old varieties in Morocco, and shortage of trained personnel limits the development and promotion of new varieties in the NARS. FP4 also rates policy and training highly, second only to technological outputs. This outline of outputs is further elaborated in the context of effectiveness in section 2.2.1, where outputs are analyzed as the substrate for outcomes for each flagship over the review period, and in section 2.2.2, where selected Outcome Impact Case Reports (OICRs) delve into the impact of R&D in breeding and agronomy.

All flagships target one or more of the following in their training activities: farmers, NARS scientists, and young scientists through MSc or PhD studies and shorter training programs. Training data in annual reports were unsuitable for quantitative analysis because they were highly aggregated or patchy and of uncertain quality (section 1.8). Additional data provided by WHEAT showed that 46 PhD students completed their degrees in 2018. This is about 0.6 graduates per scientist. Section 2.3 expands the analysis of training with a focus on gender and youth.

2.1.3.2 Scientific Publications

Key finding: 79% of the scientific publications were in top 50% of journals; 21% of the publications in the bottom half is a symptom that needs attention.

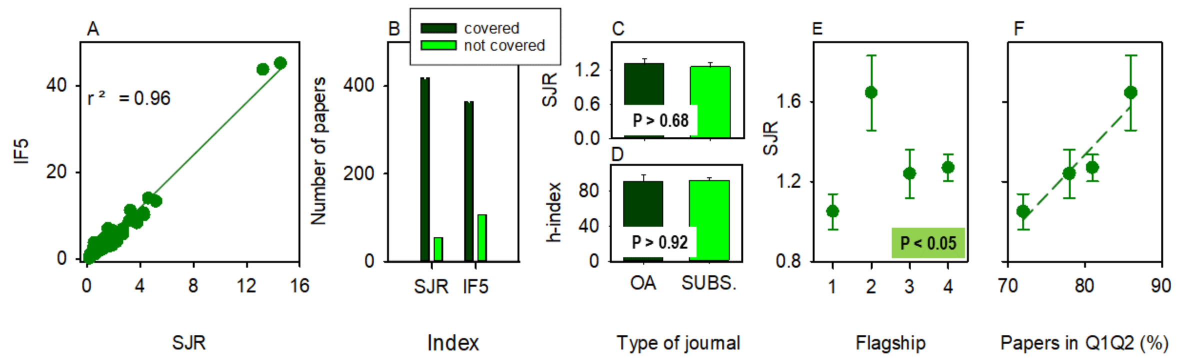
SScientific publications provide a common currency to assess heterogeneous outputs—e.g., a new sociological concept, phenotyping method, or agronomic practice. Here we analyze the performance of WHEAT and its FPs using bibliometric indices for all the 2017–2019 publications and detailed analysis of 2017 papers. In this analysis, we used the database provided by WHEAT (Annex 9).

Sample and indices

Despite some uncertainty in the quality of records (section 1.5.1), a consolidated sample of 469 journal papers is a robust representation of WHEAT scientific output for the period under review. IF5 and SJR provide similar information (Figure 2A), but SJR captured 416 papers with a gap of 53, whereas IF5 captured 364 papers with a gap of 105 (Figure 2B). Altmetric score captures online activity around research items and covered 56% of the papers under analysis compared with an 89% coverage of SJR (Figure 2B). Hereafter, we focus on SJR because it has the broadest and most consistent coverage of WHEAT's journal publications. Importantly, IF5 and SJR are publication-level metrics. They do not apply to individual papers, authors, research groups, or institutions. Furthermore, for agronomy and plant sciences the impact factor of a paper and the impact factor of the journal are unrelated (Slafer 2008).

Of the 469 publications, 53% were open access. The proportion of publications in open-access[9] journals varied with FP (P < 0.05) from 63% in FP2 to 43% in FP4. Both SJR (P > 0.68) and journal h index (P > 0.92) were similar between open-access and subscription journals (Figure 2CD).

Figure 2. Bibliographic analysis of WHEAT, 2017–2019



G

FP	Share of WHEAT (%)	Multiple FP (%)	Papers in journals (%)				SJR	Papers x SJR	Papers x SJR per FTE
			Q1	Q2	Q3Q4 NI	New			
1	21	43	44	27	26	2	1.05	105	9
2	19	24	59	27	14	0	1.65	146	11
3	35	17	50	28	21	1	1.24	205	5
4	25	20	60	21	17	3	1.27	146	8

(A) IF5 and SJR provide the same information, but (B) SJR has broader coverage of WHEAT publications. (C) SJR and h-index of open access (OA) and subscription (SUBS.) journals are statistically undistinguishable. (E) SJR varies with flagship and (F) aligns with proportion of papers in Q1 + Q2 journals. (G) Flagship share of WHEAT publications, proportion of papers signed by multiple FPs, proportion of papers in Q1, Q2, and Q3 + Q4 + not indexed journals, and new journals; average SJR for all papers published per flagship; papers weighted by SJR, and papers weighted by SJR, scaled for FTE. In E and F, error bars are two standard errors.

The SJR varied with flagship program (P < 0.05); it was highest in FP2 and lowest in FP1 (Figure 2E). To check for bias in SJR across FPs, particularly against FP1, with its larger focus on social sciences, we tested the effect of FP and journal Quartile (Q) on SJR with analysis of variance; Q captures discipline-specific variation (Table 1). The lack of FP x Q interaction effect on SJR (P > 0.82) indicates that SJR was unbiased across FP (Annex 11). Furthermore, average SJR aligned with the proportion of Q1Q2 publications (Figure 2F): FP2 averaged higher SJR because it published 86% of papers in Q1Q2 compared with 72% for FP1.

WHEAT featured 53% of papers in Q1 journals, and 79% for aggregated Q1 and Q2 (Figure 2G). The proportion of papers aggregated in Q3, Q4 and non-indexed journals varied from 26% in FP1 to 14% in FP2. Papers weighted by SJR and scaled by FTE varied two-fold, from 11 in FP2 to 5 in FP3, consistent with their R&D scope (section 1.2.2). Weighted by SJR and scaled by FTE, FP1 and FP4 performed similarly. The proportion of publications with multiple FPs averaged 25% for WHEAT and ranged from 42% for FP1 to 16% for FP3 (Fig. 2G); this variation is consistent with the transversal scope of FP1 and the narrower focus of FP3.

[9] The fee for a full paper in *Frontiers in Plant Science*, one of the journals favored by WHEAT scientists, is US$ 2,950. MDPI journals charge between US$ 1,000 and US$ 2,000 for a full paper. Assuming an average of US$ 1,500, the aggregated cost of open access publications for WHEAT was in the order of US$ 370,000 over three years.

To what extent do bibliometric indices reveal aspects of WHEAT scientific output of relevance to its high-level goals?

To explore this question, we reviewed all papers published in 2017.

Key finding: : WHEAT displays a robust matching of questions and methods and a diversity of approaches, including experiments in controlled environments and in the field, surveys, and modeling.

One of the main sources of ineffective investment in agR&D is the mismatch between question and experimental set up (Sadras 2019). For example, glasshouse trials are often reliable for some traits (e.g., herbicide tolerance) but not for other traits (e.g., yield). The literature abounds in question-method mismatch, with a common and trivial conclusion that experiments in controlled environments have to be verified in the field. A trademark of WHEAT, reflected in its publications, is a robust matching of questions and methods and a diversity of approaches with a major focus on field experiments in realistic agronomic conditions.

Key finding: Publications in Quartile 1 journals (Q1, top 25%) were world-class, and some pushed scientific boundaries; publications in Q2 journals were sound but often routine work; publications in Q3Q4 journals commonly featured flaws.

WHEAT papers in Q1 journals were well written and featured clear objectives that flowed from complete, unbiased, and updated reviews of literature, experimental design that matched objectives, replicated experiments to capture environmental variation, rigorous statistical analysis, and conclusions justified by data. Studies published in Q1 journals were largely flawless. Highlights of 2017–2019 research included the quantification of airborne dispersal routes of pathogens over continents to safeguard global wheat supply (Meyer et al. 2017); environmental consequences [10] of and alternative practices to stubble burning (Shyamsundar et al. 2019); an early warning system to predict and mitigate wheat rust diseases in Ethiopia (Allen-Sader et al. 2019); DNA fingerprinting to track adoption of varieties and geographic movement of seed (Dreisigacker et al. 2019); and faster breeding methods (Li et al. 2018). WHEAT scientists have published authoritative reviews and theoretical studies that reflect positively on the dual motivation of curiosity and utility (section 1.2.4). Reviews focused, for example, on complementary practices and enablers to make conservation agriculture more functional for smallholder farmers in Africa (Thierfelder et al. 2018); guides for development of gender-equitable value chains (Stoian et al. 2018); a timely view of meta-analysis with a focus on organic agriculture and conservation agriculture in the context of the political economy of development-oriented agricultural research (Krupnik et al. 2019); and speed breeding and other leading-edge plant breeding technologies (Li et al. 2018).

WHEAT papers in Q2 journals were mostly sound but normally showed routine work and higher frequency of faults such as oversimplification, conclusions not fully supported by data, or unjustified assumptions. For example, a study concluded that a particular enzyme and its functional marker are valuable to improve grain yield in wheat breeding, when the paper did not report yield but focused on grain weight, overlooking the trade-off between grain number and grain weight.

WHEAT papers in lower-quality journals (Q3, Q4, not indexed) represented 21% of the total. Most papers in this category lacked novelty or featured some fundamental flaw in one or more aspects of experimental design, data analysis, or interpretation. For example, studies reported genetic components of phenotypic variance in a single environment, overlooking environmental influences and the interaction genotype-by-environment. Conclusions were drawn about the impact of global warming potential of agronomic practices with no account of environmental variation and the interaction between environment and management. Descriptive studies reported sowing date and variety influences on yield, with no attempt to explain the drivers of the responses. It was observed that treatment A changed the profiles of enzymes in soil and improved yield, concluding that enzymes have a causal influence on yield. Similarly, correlation was mistakenly used to conclude a causal link between use of certified seed and yield. Reviews published in these journals were descriptive, of the form Smith found A and Gregory found B, and some of these used outdated literature. End users would be justified if they ask questions about the novelty and credibility of a new method to identify allelic variations of rust resistance genes published in

[10] In 2015, India had an estimated 1.09 million deaths from air pollution, costing the economy 3% of gross domestic product. In late October and early November, rice crop residue burning from adjacent states contributes 25–70% of the fine-particulate matter pollution in New Delhi (Balwinder et al. 2019).

a Q4 journal. Few papers in lower-quality journals (Q3, Q4, not indexed) were sound, with the potential to suit Q2 journals provided format, presentation, and grammar could be improved. Examples in this class included the analysis of the role of stubble on the potassium budget in soils and an explanation of the persistence of wheat landraces in some regions in terms of preferences for traditional practices, integration with animal husbandry, and flavor of bread.

In conclusion, bibliometric indices captured variation in quality science with real-world implications for credibility and efficiency in the use of agR&D resources. A majority of papers in high-ranked journals consistently featured cutting-edge research, novel technologies and methods, and substantial progress in the frontiers of global science. Clusters of papers in lower-level journals reflected one or more weaknesses that need attention.

Consider again the scenario outlined in section 1.5.1.2. As a donor, you have to select one of two research providers. Institute A publishes 100 papers per year; 80% of them are good to excellent and 20% are mediocre. Institute B publishes 80 papers; all of them first class. This is all the information you have; make a decision.

Surveys and interviews helped to identify two causes for substandard work. One is the system for evaluation of individual scientists. Scientists are evaluated annually, and hence face pressure to produce a larger volume of often lighter papers. This is particularly important for young scientists building a record of publications to ensure the continuity of their careers, and in pockets where critical mass has been compromised. Evaluation and rewarding systems for individual scientists need attention, as they may have unintended consequences (Franzoni et al. 2011). The second is students from scientifically less-mature systems. They may have requirements to publish, and their work could be Q3Q4 level. In these cases, the CRP supervising scientists do not need their name in the byline; this would be the approach of Institute B to signal its standard. Institute B can still claim credit for training.

2.2 Effectiveness

For the period 1994–2014, CGIAR-related varieties as a share of all wheat varietal releases averaged 63% globally, ranging from 48% in former Soviet Union countries to 92% in South Asia. The benefit-cost ratio for CGIAR wheat improvement efforts ranged from 73:1 to 103:1 (Lantican et al. 2016). This demonstrates both the effectiveness of the program and, again, the mismatch between the time frame for the review and the time frame for technological innovation and impact.

This section has four parts. First, we explore the annual achievements of WHEAT and FPs against planned outputs and outcomes. Then we analyze three selected outcome cases to better understand how WHEAT contributed to the global discussion. Next we assess how ToCs were used to guide the program. We conclude with a brief exploration of management and governance.

2.2.1 Achievement of Planned Outputs and Outcomes

Key findings: In 2017–2019, WHEAT mainly achieved its planned outputs and outcomes, and in addition had unplanned outcomes. For the three years reviewed, WHEAT did not drop any research line. Slight changes included the addition of research priorities on mechanization (FP4) and soil-borne diseases (FP3) in 2017.

While the CGIAR requested to use CGIAR Results Dashboard to address this question, our quality assurance approach showed these data were unreliable (Annex 9). The database provided the percentage of milestones completed, not completed, or unknown (Annex 12). For WHEAT in the 2017–2019 period, CGIAR data showed the proportion of milestones completed ranged from 31% to 67%, the proportion of milestones not completed ranged from 28% to 34%, with the balance in the undefined "unknown" class. We cannot interpret these data with any level of accuracy for several reasons.

First, the WHEAT Management Team reported not being familiar with these categories; instead they use "completed," "extended," and "canceled." Second, there is no context for evaluation—e.g., a milestone could be "establish 50 trials" that failed because of a hailstorm, or "deliver a new phenotyping platform" that failed because of technological issues. Both would be given the same weight in the percentage, but the implications are clearly different. Third, we cannot empirically state that not completed milestones were within WHEAT's control. Fourth, interpretation of these data requires mapping each specific milestone against the multitude of possible causes that may have compromised its achievement; this is well outside the scope of the review. Fifth, there are no agreed benchmarks to determine effectiveness based on percentage of milestones achieved/not achieved. Sixth, these data do not capture unplanned achievements typical of scientific inquiry (section 1.2.4). Finally, the CGIAR database figures do not match the data provided by WHEAT. Given that WHEAT provides all the data to the CGIAR, these figures about achievements should match. The WHEAT data show that between 2017-2019, 95% of milestones were either met or extended (compare this to the above figures). The remaining small percentage was canceled mostly owing to lack of funding.

For these reasons, we favored a rigorous assessment of **fidelity,** whereby we mapped the **identified** outcomes and their impact against **planned** outcome **areas** identified in the theory of change. Between 2017 and 2019 WHEAT demonstrated achievements and did not drop any research line, the FPs' impact pathways and ToCs remained unchanged, and contributions in intended outcome areas were identified. In 2017, WHEAT added research priorities on mechanization (FP4) and soil-borne diseases (FP3) (Annual Report, 2017, 2018, 2019). Below we provide a robust assessment of WHEAT's effectiveness based on major annual achievements at flagship level.[11]

FP1: Enhancing WHEAT's Research for Development (R4D) Strategy for Impact

In 2017, FP1 reported that research spanned questions and contexts including changing wheat **consumption dynamics** in rice economies, **wheat blast** threat in South Asia, and modeling to identify hot spots for yield decline with projected warming by 2030–2041 (Asseng et al. 2017). A study on **climate risk** in the Himalayan region of Pakistan found that most farmers were aware of climate and temperature changes, as well as variations in the rainfall patterns and wind, and that predominant climate risk adaptation strategies included adjusting sowing time, adopting tolerant varieties, engaging in off-farm employment, and exploiting crop-livestock interactions (Rahut and Ali 2017). Educated farmers with land rights, large land holdings, and more household assets were most likely to **adopt adaptation strategies**; adoption of those strategies improved incomes and yields and reduced poverty (Annual Report, 2017).

In 2018, FP1 reported achievements in three areas: consumer demand and value chains, management of disease outbreaks, and climate-smart agricultural practices (CSAPs). Research in Bangladesh reported differential projections for rural and urban households by 2030 whereby both would consume more wheat, but urban households would consume less rice compared with 2015 (Mottaleb Khondoker et al. 2018). A new conceptual framework for smallholder value chains, based on complex adaptive systems, was used to study how consumers access and buy processed wheat and maize-based products in Mexico City (Orr and Donovan 2018). An *ex ante* study of wheat blast in Bangladesh, India, and Pakistan estimated an annual potential post-harvest loss of 0.89–1.77 Mt, with 7 Mha at risk (Mottaleb et al. 2018). A study of alternative land uses to control wheat blast in Bangladesh and West Bengal, India, showed a "wheat holiday" would be feasible in the short term only and recommended a focus on disease forecasting and blast-tolerant varieties. Owing to changing input responses under varied weather abnormalities, the economic and environmental benefits of combined CSAPs were found to be superior to individual practices in the wheat-rice systems of the Indo-Gangetic Plains (Kakraliya et al. 2018). Interview data noted that unpublished work done in 2018 had useful outputs related to wheat blast and wheat rust, which built momentum and is being replicated.

In 2019, 2019, FP1 reported progress on **conservation agriculture** adoption in India, and a strong case was made for conservation agriculture in Tunisia. WHEAT reported that Ethiopian farmers sowed more **rust-resistant varieties** and/or increased their variety diversity to reduce the effects of rust re-occurrence after a 2010–2011 yellow rust epidemic. Adopting resistant varieties offers Ethiopian farmers a 29–41% yield advantage even under normal conditions, demonstrating the importance of continued development and deployment of resistant varieties to help smallholders maintain improved yields under rust challenges. A study with a small sample of farmers in Algeria showed that no-till increased gross margin by $84/ha and almost halved working time and fuel consumption in comparison with conventional practices.

FP2: Novel Tools for Improving Genetic Gains and Breeding Efficiency

In 2017, FP2 demonstrated accomplishments in **genetic diversity potential**. The largest-ever genetic characterization of wheat diversity from CIMMYT and ICARDA took place, and a related study identified

[11] While trying to streamline reporting and favor comparisons, CGIAR reporting templates force formats and content that do not necessarily capture the depth of research achievements, with implications for M&E to inform decisions. We thus used annual reports as a primary guide to key publications and analyzed these primary sources to capture the meaning of the achievements.

many landraces with unexplored genetic potential. There were accomplishments in **innovation in breeding research and new methods and tools,** as demonstrated by a study on genomic selection. Further, WHEAT and MAIZE biometricians delivered free software to support breeding decisions. **Innovations in pre-breeding** and their impact on breeding research (which also falls under FP3) had several accomplishments, including identifying several quantitative trait loci (QTL) associated with target traits and through a partnership, 15,000 elite lines genotyped with high-density markers for genomic selection or mapping purposes (Annual Report, 2017).

In 2018, FP2 demonstrated achievements in several areas, mostly through partnerships. For example, WHEAT's achievements are dependent on the CIMMYT **genebanks,** which provided valuable prebreeding material. In one example of how partnerships led to an achievement, research on prebreeding lines based on gene bank accessions delivered novel genetic diversity. Second, FP2 and FP3 scientists delivered four papers of varied quality on **harnessing genetic potential for nutrition and disease resistance.** A robust study used GWAS to reveal 39 marker-trait associations for grain zinc concentration (Velu et al. 2018), whereas a routine, descriptive study reported variation in grain iron and zinc content. A comment paper sketched the implications of new genomic information on *Aegilops tauschii*, the D genome donor of hexaploid wheat (Rasheed et al. 2018). Wheat-rye crosses have delivered one line with high resistance to the Russian wheat aphid (*Diuraphis noxia*) **under controlled conditions**. FP2 and FP3 scientists made further progress on genomic prediction for faster, cheaper, and more precise breeding (Juliana et al. 2019a) and genetic resources for wheat blast resistance (Cruz et al. 2016).

In 2019, a milestone paper in *Nature Genetics* reported the genomic predictabilities of 35 key traits; demonstrated the potential of genomic selection for wheat end-use quality; identified several marker–trait associations for 50 traits evaluated in South Asia, Africa, and the Americas; built a reference wheat genotype–phenotype map; and explored allele frequency dynamics over time and fingerprinted 44,624 wheat lines for trait-associated markers, generating over 7.6 million data points, which together will provide a valuable resource to the wheat community for enhancing productivity and stress resilience (Juliana et al. 2019b). Primary hexaploid synthetics revealed promising adaptations to powdery mildew, leaf/stem rust, and Septoria (Shamanin et al. 2019). In partnership with FAO's International Treaty on Plant Genetic Resources for Food and Agriculture, WHEAT scientists and breeders are researching "lost" wheat landraces. The Hessian fly–resistant variety Faraj released in Morocco in 2008 became commercially available in 2019 and was grown on 20,000 ha.

FP3: Better Varieties Reach Farmers Faster

In 2017, WHEAT reported the release of 63 CGIAR-derived wheat varieties released globally in 19 countries. WHEAT brought together 31 national partners to operate a low-density marker platform to genotype nearly 40,000 wheat lines using from 1 to 10 gene-based markers, generating more than 200,000 data points. SampleTracker software strengthened monitoring of DNA samples linked to germplasm, by providing each study with a unique project ID and each sample an individual ID. The research partnership resulted in a set of heat-adapted durum wheat varieties for the Senegal River Ba. In partnership with Cornell University, the Global Rust Monitoring System was expanded to 30,629 geo-referenced survey records in 20 countries. Between 2006–2007 and 2015–2016, the genetic gain of yield for CIMMYT spring wheats averaged 0.53% yr^{-1} against local checks and ranged from 0.4 1 % yr^{-1} in mega-environment 1 to 1.0 % yr^{-1} in mega-environment 5 (Crespo-Herrera et al. 2017). These rates compare with a target 0.70 % yr^{-1}. A final example of research included an in-house publication of an exhaustive review of recent scientific studies on cereal grains and health.

In 2018, WHEAT reported the release of 48 varieties derived from its breeding research. Rust Tracker, the most comprehensive crop disease monitoring system worldwide, produced over 35,000 survey records from more than 39 countries. Wheat blast research and Delivering Genetic Gains in Wheat (DGGW) had multiple achievements, including more than 4,000 accessions screened for wheat blast in Bolivia and close to 42,000 accessions tested for stem rust in Kenya. Field precision-phenotyping platforms demonstrated effective partnerships with NARSs, returning 10,000+ accessions phenotyped for biotic/abiotic stress traits. A genome-wide association study in synthetic hexaploid wheat revealed 92 marker-trait associations for micronutrients; the apparent lack of trade-offs with yield indicates these markers have potential for biofortication (Bhatta et al. 2018).

In 2019, In 2019, WHEAT reported the release of at least 50 varieties derived from its breeding research, including improvement in wheat blast resistance in Bolivia and India (Annual Review 2019, pp. 7–9), and heat and drought-adapted lines released in Pakistan for which there is an OICR (AR 2019). The established notion that selection for potential yield also improves actual yield in stressful environments (Foulkes et al. 2009) was corroborated in a rigorous study (Voss-Fels et al. 2019). Singh et al. (2019)

highlighted the slow rate of varietal turnover in WHEAT target countries and reviewed the constraints for faster rates. FP3 was part of the Nature Genetics study reported above for FP2 (Juliana et al. 2019b).

FP4: Sustainable Intensification of Wheat-based Farming Systems

In 2017, MAIZE and WHEAT participated in developing the Framework for Sustainable Agricultural Mechanization in Africa. India is the third largest greenhouse gas emitter in the world ,with agriculture accounting for 18% of gross national emissions and the Indo-Gangetic Plains producing food for about 40% of India's 1.2 billion population. In this context, a modeling study focused on high-yield low-emission pathways for cereal production (Sapkota et al. 2018 In a sample of 1,000 households from Bihar, Eastern Indo-Gangetic Plains, only 44% of households knew about zero-till, and both awareness and adoption were biased in favor of larger farms (Keil et al. 2017). Baudron et al. (2017) reported that in the Munesa Forest, Ethiopia, diversity of household diets was inversely related to distance to the forest; this difference was not explained by forest food collection but by biomass flows from the forest to farmlands in the form of feed and fuelwood. WHEAT demonstrated progress on scaling approaches by embedding a new tool in a training package, and WHEAT and MAIZE co-founded a CGIAR Scaling Community of Practice with key partners.

In 2018, WHEAT published several studies on SI, with research conducted in Ethiopia, Mexico, and India. Lopez-Ridaura et al. (2018) found that compared with livestock interventions, conservation agriculture may hold considerable potential in Eastern India, though primarily for wealthier and medium-scale cereal farmers, which are also more vulnerable to drought. Advances were made in mechanization, in partnership with iDE, and to date 191,000 farmers can now access services from a network of nearly 3,000 local providers, representing improved cultivation across 92,000 ha in southern Bangladesh (Annual Report, 2018, p. 8).

In 2019, ICARDA developed a raised-bed machine that can prepare an acre of agriculture land in half an hour, a task it would take 10 people-days to complete manually. As technology transfer of large machinery from high-income countries has been ineffective, mechanization options were scaled to small and scattered plots, and more recently the focus for improved adoption shifted from farmers to service providers. In this context, Van Loon et al. (2020) used the Scaling Scan tool to assess three case studies designed to scale different mechanization service provider models with implications for development interventions in Mexico, Zimbabwe, and Bangladesh. Three complementary pathways may link forests to diets: a direct pathway (e.g., consumption of forest food), an income pathway (income from forest products used to purchase food from markets), and an agroecological pathway (forests and trees sustaining farm production). Baudron et al. (2019b) used piece-wise structural equation modeling to report evidence of a direct pathway in four landscapes (Bangladesh, Cameroon, Ethiopia, and Zambia) and evidence for an agroecological pathway in three landscapes (Bangladesh, Ethiopia, and Indonesia), with no evidence for an income pathway in any of the landscapes sampled. Rasmussen et al. (2019) conducted hypotheses-driven research to explore the relationships between the spatial arrangement of forests and diets and found that (1) the influence of forest on dietary quality extends beyond the proportion of forest in the landscape, (2) the spatial arrangement of forest may influence the consumption of wild vs cultivated fruit, and (3) more diverse diets were associated with greater wealth and the number of crops grown, but higher crop diversity might do little to secure the consumption of a nutritionally important food group, namely fruit.

2.2.2 Demonstrated Importance of Outcomes (Deep Dive on Selected OICRs)

Key finding: The review of three OICRS suggested WHEAT's R&D in agronomy and breeding is relevant to global wheat challenges such as climate change, mechanization, and farmers' income. The cases also demonstrated the importance of both partnerships and long-term (30-year) strategies.

To select the OICRS, we used five sampling criteria: (1) one case on breeding, (2) one case on agronomy, (3) access to key informants, (4) cases with maturity level 2 or 3, and (5) the likely usefulness to the broader agricultural sector. Further explanation for selection can be found in Annex 13.

Critical Outcomes in Wheat Agronomy – OICR 2524

Many of the 2.5 million farmers in northwestern India rely on rice-wheat cropping systems (~4.1 M ha). About 23 Mt of rice residue in their fields is burnt to prepare the land for wheat sowing, contributing to 25–70% of the fine-particulate matter pollution in New Delhi. The Climate Change, Agriculture and Food Security

(CCAFS) platform, WHEAT FP1 and FP4 and their partners from the Nature Conservancy, the University of Minnesota, the Indian Council of Agricultural Research (ICAR), the Borlaug Institute for South Asia (BISA), and other organizations advanced, tested, and promoted technologies to solve this environmental problem. They returned additional economic benefits, making a strong case for conservation agriculture in the context of climate change.[12] As an alternative to burning, the Happy Seeder cuts and lifts rice straw, sows wheat directly into the soil, and deposits the cut straw as mulch over the sown area. A paper in Science explains how systems based on the Happy Seeder reduce air pollution and are on average ~10% more profitable than the most profitable burning option with zero-till seeders and ~20% more profitable than the most common burn system with conventional seeders (Shyamsundar et al. 2019). Signed by 29 authors, the credible milestone research engaged officials from Indian government agencies, farmers, plant managers, and manufacturers. The engagement by the authors and then with key counterparts demonstrates how partnerships are vital for WHEAT to develop agronomic solutions to critical problems associated with poverty, air pollution, and health.

- **What does that mean?**

Scientists in the CSISA program, led by WHEAT among others, made a series of findings that offered evidence-based insights and/or pilots of business models that encouraged adoption of climate-smart farming practices in the region.

- **What is the impact?**

All the findings—whether on the importance of farmer access, awareness, and social inclusiveness of the practices; timely support for small mechanization entrepreneurs; or subsidized machinery sales—had or are having an impact on decision-making within government agencies on ways to encourage climate-smart agricultural practices. However, one big impact is a program to promote similar Happy Seeders in India. That program led to a 2018 government policy recommendation to subsidize the machines and the launch of a special scheme for *in situ* management of crop residues. More recently, it influenced a GOI investment of over 11b INR (€140 million) in subsidies for the machines in three states over two years.

- **Why is this important?**

In northwest India, as much as 22 Mt of rice stubble is burned each year in Haryana and Punjab alone, resulting in pollution that travels for thousands of kilometers. The 14 cities with the highest air pollution globally are in India, and stubble burning is a major contributor. An affordable seeder allows Indian farmers to sow through the stubble and was identified as one of the most cost-effective, innovative, and scalable solutions to manage farm residue on site. In addition, Happy Seeder technology can improve agricultural productivity by 10-15%, reduce labor costs and time, and contributes to recycle nutrients from the crop residue back into the soil.

Critical Outcomes in Wheat Breeding – OICR 3284

112BThe CGIAR synthetic wheat breeding strategy successfully transfers valuable diversity from goat grass to modern wheat, providing farmers with higher-yielding, more reliable varieties . Over three decades, CIMMYT has developed synthetic hexaploid[13] (SH) wheat to improve tolerance to biotic and abiotic stresses. A WHEAT study analyzed the genetic contribution of SH wheat to spring bread wheat germplasm providing a retrospective view into development and utilization of SH in the CIMMYT Global Wheat Program. It found that 20% of the lines sampled in two international yield trials were synthetic derived with an average D' contribution of 15.6% (Rosyara et al. 2019). These results underline the importance of SH for genetic diversity and a more targeted introgression strategy. Because CIMMYT contributions are present in nearly half the wheat sown worldwide, this means that many of the over 2.5 billion people in 89 countries, who consume wheat—over 1.2 billion of whom live on less than US$2 a day and depend on wheat as their primary staple food—are benefiting from diversity and resilience derived from ancient wheat relatives.

[12] OICR 2524 (2019) and OICR 2764 (2018) provide outcomes for the same initiative, at different stages.

[13] Synthetic hexaploid (SH) wheat (AABBD'D') is developed by artificially generating a fertile hybrid between tetraploid durum wheat (*Triticum turgidum*, AABB) and diploid wild goat grass (*Aegilops tauschii*, D'D').

WHEAT's study supported by FP2 and FP3 demonstrates that a long-term approach to research contributes to significant outcomes (better wheat varieties) that have the potential to contribute to impacts (to positively affect farmers and those who consume wheat). These findings provide a strong case for scaling up the synthetic wheat-breeding strategy to meet the urgent global demand for climate-resilient, disease- and pest-resistant, high-yielding wheat. At the same time, the study suggests how sudden shifts in, or levels of, funding could compromise long-term R&D strategies, which are essential to highest level of impacts.

- **What does this mean?**

For more than 30 years, CIMMYT wheat breeders (including WHEAT researchers but predating the CRP) have been using a breeding technique whereby they cross a wild wheat relative with modern durum wheat to produce "synthetic hexaploid wheat," which is then backcrossed with modern bread wheat. The process incorporates the genetic diversity and resilience traits present in goat grass (*Aegilops tauschii*) into modern wheat. *A. tauschii* is a valuable source of disease resistance and nutritional quality, and possibly heat and drought tolerance.

- **What is the impact?**

A 2019 study used state-of-the-art molecular technology to measure the effect of these efforts, and found that 20% of the wheat lines in CIMMYT's global spring bread wheat breeding program contain an average of 15% of the genome segments from the wild wheat relative *A. tauschii*. The only way these segments could have ended up in CIMMYT's modern wheat is through the synthetic breeding approach. And the fact that the "winners" in CIMMYT's rigorous selection process include these segments validates synthetic wheat breeding.

- **Why is this important?**

The relevance of this outcome is that decades of CIMMYT research contributions are present in nearly half the wheat sown worldwide. This means that CIMMYT's improved, diverse, and resilient wheat varieties, including synthetics, benefit farmers throughout the world who rely on wheat production for their incomes. Wheat consumers, who number over 2.5 billion people in 89 countries, including more than 1.2 billion resource-poor consumers who rely on wheat as a staple, also benefit. The scientific validation of this breeding technique has the potential to encourage other breeders to use it to meet the urgent global demand for climate-resilient, disease- and pest-resistant, high-yielding wheat.

2.2.3 CRP Management and Governance

Key findings: WHEAT's roles, responsibilities, and accountabilities are clearly defined and exercised, and governance is sufficiently independent. Some slight challenges related to funding exist. Monitoring and evaluation had some challenges yet appear to adequately support the CRP. Redundant reporting and frequently changing requirements are burdensome.

The partnership

CIMMYT is the lead center, and partners with ICARDA to guide and implement WHEAT. The WHEAT director oversees WHEAT and attends both CIMMYT and ICARDA Board of Trustee meetings to report on progress and raise governance issues. A team of five supports the CRP director, including the WHEAT manager and a senior monitoring, evaluation, and learning specialist (RBM, 2017, p. 3–4). WHEAT is composed of four FPs, each of which has its own managing FP leaders.

Building one global CGIAR wheat program

The Management Committee (W-MC) is WHEAT's executive working committee, while the WHEAT Independent Steering Committee (W-ISC) provides oversight and advice to the Wheat-MC. The W-MC is composed of the CIMMYT and ICARDA program directors and three external members representing strategic R&D partners including ACIAR, BBSRC, and ICAR. The W-MC reviews work plans, budgets, and identifies research gaps, while the W-ISC makes recommendations on these items (WHEAT Handbook, 2015, p. 14). Flagship supervision is assigned to distinct members, yet decisions are taken as one body, which requires endorsement by external members (RBM, 2017, pp. 3–4). Shared FP and CoA leadership reflects progress toward CIMMYT and ICARDA goals (One Global CGIAR Wheat Program) that is driven by a five-year milestone plan (now four years) endorsed by both centers' Boards of Trustees (RBM, 2017, pp. 3–4). The WHEAT Handbook (2017) describes WHEAT's basic management, governance, budgeting, and other related topics, though it is slightly outdated. For example, in 2018 CIMMYT and ICARDA directors general became full voting members of the W-ISC. The Wheat-ISC has developed a fixed point on its agenda during which members are asked to present their research for the benefit and interest of ISC colleagues. Recently the W-ISC took the initiative to draft suggestions for the future of ISCs in the context of One CGIAR construction.

W1/W2 funding challenges

During 2017–2019, WHEAT reported slight issues with W1/W2 funding. In 2018, WHEAT reported volatility and unpredictability, and interview data confirmed that the W-MC buffering budget provided sufficient mitigation. Here are some examples:

- In December 2018, WHEAT learned that it would receive a budget $670k higher than anticipated. The System Management Office (SMO) then had to change W1W2 per CRP FinPlan 2018 figures three times during the year (Annual Report, 2018, p. 10).
- In 2019, WHEAT reported that W1/W2 unpredictability remained an issue, mainly because of in-year donor W2 shifts away from WHEAT and the risk that the total W1 target might not be reached. W-MC maintained a buffering reserve, based on a so-called midi-scenario, until October 2019, when CGIAR SMO confirmed the W1W2 budget. The lower-than-planned W2 would be compensated by the Stabilization Fund, which was created in 2017. W-MC then used the buffer budget to fund more partner grants (Annual Report, 2019, p. 13).

WHEAT uses W1 funding to support a partnership network that mixes cultures, opinions, and research foci. Interview data suggest that the volatility of W1 affects the stability of these partnerships. However, our preliminary finding from network analysis suggests that networks might be scale free driven by preferential attachment—this means the rich get richer; entering the network may be hard for new or minor partners (section 2.1.2.3). This deserves further research. WHEAT has demonstrated a strategic approach to leveraging outputs from regionally focused bilateral funding to effectively achieve global-level outcomes. Further, WHEAT has leveraged advances in upstream academic research to pilot and apply in public wheat R4D. Consequently, lack of flexibility of CRP resources can compromise future opportunities.

One CGIAR Global Wheat Program (OWP)

Both W-ISC and W-MC had a One CGIAR/2030 Plan. A common concern was the lack of a rationale for organizational change (e.g., merge Centers, reporting lines across Centers) and the need to ensure stability, continuity, and the current CRPs' delivery during the transition. In September 2019, CIMMYT and ICARDA revisited the progress made toward OWP. CIMMYT and ICARDA agreed on an OWP that will be reflected in future collaboration from gene banks to prebreeding to breeding, including interdisciplinary teams to tackle specific challenges. The agreement positions the OWP well for a future One CGIAR, irrespective of the direction it will take (Annual Report, 2019, p. 13).

Monitoring, evaluation, learning, and impact assessment (MELIA)

In Phase II, WHEAT M&E had multiple challenges and multiple achievements. One challenge included common CRP indicators being introduced halfway through Phase II. This then meant that FP project leaders and scientists had to be orientated about what these were and what they meant for managing the activities. Other concepts were introduced but did not bring clear definitions or agreed understandings. One such example is "innovation." When terms were not clear, they presented challenges for monitoring, assessment, and reporting. One respondent noted what several others stated:

> *It's a lot of work to mainstream lot of work to mainstream the language and definitions…and when CRPs interpret things differently there's confusion…Even with the MELIA…like how a project is even defined…it is confusing for planning and reporting.*

MARLO

In the Annual Report 2019, WHEAT stated that it is fully utilizing MARLO (Managing Agriculture Research for Learning and Outcomes) to link individual projects and areas of research to FP theories of change and to monitor research progress. MARLO supports WHEAT in operationalizing the POWB by collecting important lessons across projects and incorporating these in program decision-making and institutional learning (p. 16). A brief exploration of the system suggests that it supports well-organized, well-documented, and focused annual reports. For example, the reports became clearer and more focused and contained more data to support statements as the years moved forward. While a few respondents noted that the ToC was not always explicitly used to inform projects, the annual reports identified how the ToC was used to document achievements.

M&E Achievements

WHEAT's reporting context is complex for various reasons, including but not limited to multiple partners, varying uses and kinds of monitoring, differing definitions or lack of definitions, high rate of turnover in management tools, and misunderstanding around WHEAT and M&E terminology. For example, the categories and analysis of milestones promoted by CAS in this report (section 2.2.1) are foreign to WHEAT management. At the same time, WHEAT developed or contributed to various useful M&E processes and products. These include WHEAT contributing to the CGIAR MEL Glossary, the refinement of a quality assurance process for annual reporting, the establishment of a CGIAR MELIA support pack progress on how to implement the projected benefits indicator, agreement on streamlining planning and reporting on MELIA studies, and the creation of new sub-groups to address issues in 2020 (Annual Report, 2019, p. 16).

2.2.4 Progress along ToC (CRP and Flagships)

Key finding: The ToC provides explicit shared thinking about how change comes about in a larger context and is useful for (1) priority setting, (2) assessing the contribution of scientific outputs, (3) seeking and justifying funding, (4) mapping trajectory to impact, and (5) reporting, but is unsuitable for (6) assessing the progress of WHEAT and its flagships.

WHEAT has five theories of change: one for the overall CRP and one for each FP. These were developed collaboratively and, using performance data, reviewed at the end of 2019.[14] That review then informed 2020 planning (Annual Report, 2019, p. 16). While a review of MELIA documents demonstrates links among the WHEAT theory of how change comes about (ToC), how WHEAT planned to bring about change (POWB), and its annual reporting, not all respondents reported using the explicit ToC to construct a theory of action (or POWB). Regardless, the POWBs reflect that explicit ToC, which is then identified in the annual reports.

The evaluation findings provide a slightly different understanding of how the ToC was useful and to whom, which then influenced the evaluation approach that resulted in this section. Neither the flagships nor the CRP can be fairly assessed with regards to their progress along their ToC (question 2.4 in ToR), because that was not the ToC's purpose, though WHEAT scientists' development of FP ToCs included identification of progress indicators that fed into MELIA. Nonetheless, ToCs had critical uses. These are explained in the next paragraph and explored by using FP3 as an example.

The ToC provides explicit shared thinking about how change comes about in a larger context and has several uses. First, it is useful to prioritize and then select key areas to conduct research within an already established and agreed-upon ToC. Second, it clearly demonstrates how scientific outputs contribute to scientific, technological, and social changes. Third, it helps to understand what needs to be done to reach the SLOs and permits the CRP to seek and justify funding needs. Fourth, it provides a map that that can be used to explicitly identify responsibilities to achieve the SLOs (i.e., WHEAT is responsible for A, and another actor for B). Fifth, it provides a clear framework to report against, which clarifies how WHEAT results contribute to sectoral changes.

When asked about the ToC use, one FP leader provided this perspective:

> *The ToC provides a visual representation of the interplay between different components of FPs and between FPs themselves. Different people find the ToC more or less useful depending on their own personal learning style. The ToC has proven to be a useful tool in the articulation to stakeholders of paths to desired impact, the anticipated time taken to achieve impact, and assumptions made within that pathway.*

The FP leader's statement "different people find the ToC more or less useful depending on their own personal learning style" is then elaborated on by two other CRP leaders. Another FP lead brought a different experience:

> *We had used it when designing the activities for CRP Wheat, however since then it is not used…*

[14] The Results Based Management document (2017) states that the "impact pathway serves as the CRP's hypotheses of how change is expected to occur from output to outcome and impact" (p. 4).

A CRP lead had this to say:

> *The ToC illustrates how research will be influenced and how research will influence…we make explicit assumptions and we address them… it's not a good monitoring tool; it is more of a guide.*

The review assessed all FPs against their ToCs to identify how achievements mapped to the ToC. For all FPs, we identified the same key findings. Owing to limited space for the report, we therefore selected one example, FP3, to demonstrate these key points (please refer to Annex 14 when reading this section). While FP3 demonstrated significant achievements and effectiveness, each of which contributed to their ToC (see section 1.2.1), the ToC cannot be used to assess how far FP3 moved along its ToC to make any effectiveness judgment.

A second key point is on cross-cutting issues; note the use of cross-cutting issues in the ToC (Annex 14). Here we select gender to discuss how cross-cutting issues are acknowledged yet not fully engaged, a point that is applicable to all FPs and the overall CRP ToC. The FP3 ToC is clear that gender needs to be addressed to achieve the SLOs; gender literally "splits" the lower-level results from the top-level results. However, it is not clear how change will come about for gender or how gender will bring about what change. Thus, while the sector recognizes the importance of cross-cutting issues, the ToC does not provide a clear path for how these are integrated to achieve the SLOs.

Therefore, FP3 was not assessed against its ToC to determine progress or effectiveness, nor was any FP. Rather, we sought to identify what elements of the ToC had been addressed. In Annex 14, FP3 was not assessed against its ToC to determine progress or effectiveness, nor was any FP. Rather, we sought to identify what elements of the ToC had been addressed. Inwe demonstrate with green highlights where FP3 worked and had successes. Yellow marks highlight what was not done. Boxes with no highlights show no work identified by the CRP (other organizations, governments, research centers, and NARS may have done work in these areas). The ToC was likely useful to the FP (the CRP and other remaining FPs) in providing the five uses identified above.

2.3 Cross-cutting Issues (Capacity Development, Partnerships, Gender, Youth)

Key findings: Overall, WHEAT engages with capacity development and partnerships as mechanisms to achieve milestones, outcomes, IDOs, sub-IDOs, and SLOs. Gender and youth are targeted as a research focus or as research beneficiaries.

This section explores four cross-cutting issues: capacity development, partnerships, gender, and youth. Partnerships are addressed here, and section 2.1.2.3 looks at partnerships with a different analytical lens. While outside the review's scope, we found it useful to assess cross-cutting issues against the relevant ToCs.

2.3.1 Capacity Development

Key Finding: Capacity development contributes to achieve WHEAT's milestones, outcomes, IDOs, sub-IDOs and SLOs.

From 2017 to 2019, capacity development took place through short- and longer-term training, workshops, conferences, and other engagements, with a strong focus on implementing these events through partnerships. For some events, WHEAT flags specific efforts aimed at particular beneficiaries such as women, youth, or farmers in particular areas or regions. MAIZE and WHEAT co-funded a Learning Management System (LMS), which was initiated in 2017 and fully functional in 2019. All CIMMYT staff use the system to plan, support implementation, and document training/learning events of all kinds. Based on LMS data, the CRP reported training 13,070 people in short-term programs and 3,025 people in long-term programs in 2019.

While there are no specific ToC pathways identified for how capacity development change comes about (what leads to capacity development, for whom), capacity development is an integral part of how WHEAT aims to support achievement of the SLOs, particularly through FP1 and FP4, and to a lesser extent FP2.

2.3.2 Partnerships

Key Finding: Partnerships are critical for WHEAT to achieve milestones, outcomes, IDOs, sub-IDOs, and SLOs.

External and internal CGIAR partnerships and collaborations are identified in multiple aspects of the WHEAT initiative, suggesting that WHEAT is effective in engaging in partnerships, as planned. Furthermore, data strongly suggest that WHEAT would not have achieved its intended outputs or outcomes without its various strategic partners (e.g., universities, governments). For example, between 2017 and 2019, WHEAT engaged in partnerships and/or collaborations (data sources do not distinguish between these words) to achieve a large part of their intended results, providing strong evidence that these various partnerships are vital to WHEAT's achievements. A statistical analysis of scientific networks suggests preferential attachment, with implications for inclusiveness (section 2.1.2.3). With this exception, we did not explore the kinds of partnerships (e.g., formal, informal, financial, in-kind) nor their extent (e.g., strategic thinking, sharing of personnel, partners' contribution level, political engagement). Thus, we did not identify what kind of partnership is most effective (or efficient) in achieving WHEAT's intended results.

This evaluation and the 2020 MOPAN self-evaluation identified that CIMMYT and ICARDA work closely and regularly with a wide range of stakeholders from national and regional levels to identify areas of research, to conduct research, and to share findings. Partner relationships have been developed and maintained over time and are significant to WHEAT's accomplishments.

> *[Partnerships] help to ensure that the programmes are relevant to the contexts in which they are implemented, to build on and extend the capacity of implementing partners and to provide regular opportunities for reflection on progress and changes. The coverage of risk management, cross-cutting issues and sustainability in the planning of CRPs is currently adequate (MOPAN, 2020, p. 32).*

Exploring the ToC provides a different picture, where partnership and collaboration roles vary within the FP. For example, in the WHEAT overarching ToC and in the FP1 and FP2 ToC, there is a focus on strengthening partners (i.e., national partners and beneficiaries enabled) that takes place at different levels, and FP4 has a focus on "enhanced capacity for innovation in actors involved in SI" and "enhanced institutional capacity of actors in SI" (ToC FP4). The FP2 ToC focuses mostly on engaging with partners to produce a product or conduct a capacity development exercise.

2.3.3 Gender

Key Finding: "Gender" is used as a research focus, and "gender" is used to identify a research or capacity development beneficiary; gender achievements in research, while often notable, were often siloed.

WHEAT aimed to achieve specific outcomes to contribute to gender equality and empower all women and girls (SDG 5), although we note the male : female ratio in WHEAT scientific leadership is 4.7. In the Annual Report 2017 (p. 1), WHEAT stated that a specific gender outcome (IDO) was to improve understanding of gender dynamics in wheat-based systems and subsequently identified specific milestones to accomplish that outcome. The review and synthesis of data from 2017–2019 strongly suggest that WHEAT has been effective in exploring gender dynamics in wheat-based systems and sharing that understanding with a wide variety of stakeholders, which then contributes to the achievement of SDG5.

In 2018, WHEAT reported that the CIMMYT-led gender 11-CRP GENNOVATE program came to a close, with the release of a special issue in the Journal for Agriculture, Gender and Food Security (note that this journal is not indexed, which raises questions about credibility; section 2.1.3) and 17 tools or guidance notes appropriate for non-gender specialist researchers within and outside the CGIAR (Annual Report, 2018, p. 20). In 2018, WHEAT reported one policy contribution though papers and workshops aimed at influencing policy with regard to gender. In 2019, an additional result in gender was the completion of the FP1 2019 milestone *Cross-regional assessment of the influence of gender norms and agency on men/women's capacity to innovate in WHEAT AFS-based livelihoods* with the production of 5 papers and 10 reports (Annual Report, 2019, p. 38).

Gender appears in the ToC for all FPs, though at different levels. There are no specific pathways identified for how change comes about for gender, how gender brings about change, or how addressing gender is needed to reach the SLOs. However, our interviews revealed that gender is recognized as an important piece of the puzzle. For example, some data suggest efforts to identify gender-specific traits that would influence breeding priorities (e.g. quality traits), but no such traits were identified or trade-offs precluded their realization (Weiner 2019). For example, some Nepalese women favored grain with softer texture for easier milling whereas others were concerned with susceptibility to weevils associated with softer grain. For other traits, women's preferences aligned with male preferences. Other data suggest that WHEAT had contributed most data in GENNOVATE.

WHEAT addressed gender through capacity building and research, with a smaller focus on youth. Data reviewed suggest that these two areas (gender and youth) appear to be a special focus (silo), with little integration into the whole of WHEAT. For example, for some reviewed projects WHEAT reported on the number of farmers but not women or youth, and there is little evidence of how WHEAT encourages the use of its gender findings to inform other WHEAT interventions or research. The 2020 MOPAN evaluation provides similar findings for CGIAR, noting that while gender-disaggregated data are collected, it is not clear that these are being used. This evidence suggests a more systemic issue in the CGIAR with regard to gender; achievements in gender often represent isolated pockets of good practice (p. 41).

2.3.4 Youth

Key Finding: the CGIAR approved WHEAT's approach to Youth (along with gender) and results on youth reflect that agreement, with a notable achievement in 2017 that informs youth-oriented R4D.

While the review focuses on 2017–2019, it is the 2016 WHEAT proposal that clarifies how WHEAT intended to engage with youth, as "an integral constituent of our understanding of gender. Gender as a relational concept intersects with other social identities, including youth. As such, gender and youth are not mutually exclusive, but often overlap, depending on the specific context, situation and parties involved" (p. 24). The CGIAR accepted that explanation, and WHEAT implemented it accordingly, specifically noting that SI knowledge/technology portfolios differentiated for gender, youth, and resource-poor communities will be developed (2016 WHEAT Proposal, p. 20).

While there are few IDOs, milestones, or outcomes to assess youth achievements (i.e., FP1 has one outcome), results on youth are identified in all three years reviewed, including a total of four contributions to policy in 2018 and 2019. A joint MAIZE and WHEAT investment resulted in a 2017 published paper on youth in Africa and agriculture research that identified where and in which cases specifically youth-oriented R4D makes sense, as opposed to social inclusion/gender. That paper informs an approach to youth-oriented agricultural research that has wide applicability to R4D and youth. The 2017 paper looked at the potential role for the CRP WHEAT (and MAIZE), and an interview with WHEAT clarified that youth were identified as "more of a structural and rural transformation issue than youth specific research agenda per se." Further, since youth initiatives are sometimes subsumed in other work, youth-related CRP achievements may have been overlooked. For example, the paper *Understanding gender in wheat-based livelihoods for enhanced WHEAT R4D Impact in Afghanistan, Pakistan and Ethiopia* highlights gender, yet further digging identifies that youth is also addressed (which coincided with WHEAT's approach, noted above).

Youth is not prominent in most of the ToCs, and the ToC appears to provide a youth result with no clear pathway for that achievement or explanation of how youth contributes to achievement of higher-level results.

2.4 Future Orientation

Key Finding: Wheat as a crop is bound to be central to global food security in the foreseeable future. WHEAT as a R&D agent has a track record of delivering local solutions with a global perspective and is well positioned to continue this trajectory in the next decade. There are opportunities and challenges for the way ahead, including the risk of fragmenting the global breeding program, restrictions to exchange germplasm and ideas, the opportunity to integrate R&D in agronomy, misguided emphasis on minor crops, and CGIAR's focus on process at the expense of results.

Millions of farming families in Asia, the MENA region, Sub-Saharan Africa, and Latin America depend on wheat as the core source of energy, protein, and income, and on WHEAT and its partners to improve their productivity and livelihoods. Here we briefly discuss WHEAT's transition during its final year, touch on some of the challenges for WHEAT's next phase in the context of global agriculture, and summarize the roles of wheat as a crop and as a component of the CGIAR R4D portfolio, to address those challenges. We conclude with a rationale for management, monitoring and evaluation with purpose.

2.4.1 Transition Phase

WHEAT combines breeding and agronomy R4D toward a target rate of yield gain of 1.4% yr^{-1}. The One CGIAR Wheat Breeding Program can improve yield and tolerance to biotic and abiotic stresses globally through its wide adaptation strategy, international germplasm exchange, and partnerships with largely public sector regional breeders. Agronomic practices to capitalize on improved varieties and close yield gaps will likely remain a "locally adapted" affair (CRP ToC), with comparative R4D to enable learning across farming systems and regions. Without international germplasm exchange and dedicated breeding research partnerships, a globally impacting WHEAT breeding program is not feasible. Transversal breeding and agronomy programs can reinforce WHEAT impact (section 2.4.3).

2.4.2 The Challenges Ahead

The goals of global agriculture in the next two decades are (1) for all at all times, abundant, affordable, healthy and nutritious food; (2) for farmers, stable incomes, in line with the rest of society, from sustainable farming with less drudgery; (3) for the non-farm environment, absence of encroachment and of contamination by farming; (4) for rural communities, viable support and attractive landscapes; and (5) for the world, maintenance of non-agricultural biodiversity (Fischer and Connor 2018).

In the context of these global goals, our interviews highlighted operational challenges and challenges related to fashionable but not necessarily scientifically robust propositions (Sadras et al. 2020). These include:

1. **The risk of fragmenting the global breeding program into regional units.** A fragmented program would weaken current synergies and compromise economies of scale. Breeding is expensive; many national programs lack critical mass and economies of scale to justify full independence and will continue to benefit from WHEAT as a reliable provider of germplasm.
2. **Restrictions to germplasm exchange.** Exchange of material and information is paramount for a competitive wheat industry, which is largely in the public sector in WHEAT's target countries.
3. **The opportunity to integrate R&D in agronomy.** In the words of one of WHEAT's leaders "…it is insane to separate components of farming systems…".
4. **Misguided emphasis on minor crops.** Wheat, together with maize and rice, are the backbone of global food security, and reallocation of effort to minor crops could prove an expensive distraction, while a net increase in CGIAR breeding research scope on pulses/legumes and select "minor" cereals, such as barley, would broaden the CGIAR's offerings for marginal environments and crop rotation alternatives.
5. **CGIAR's focus on process risks compromising results**, as outlined in sections 2.1.2.3 and 3.2.3.

2.4.3 Wheat as a Crop and WHEAT CRP as an R&D Agent

Wheat is, together with maize and rice, the backbone of global food security. The diversity of wheat-based and "including wheat" systems, with their different rotations, farming practices, and agroecological and sociopolitical environments, continues to offer comparative R4D potential. Wheat is the most widely adapted crop, growing in diverse environments ranging from sea level to regions as high as 4,570 m.a.s.l. in Tibet and from the Arctic Circle to the equator, but most suitably at the latitude range of 30° to 60°N and 27° to 40°S (Tadesse et al. 2019). In most targeted countries, wheat R&D remains in the public sector. A commonality of interests favors an open and synergistic interchange of materials, information, and technology and dampens the tension between competition and collaboration in comparison with other industries (section 2.1.2.3). Wheat has wild relatives that have already contributed to agriculture in the form of chromosome translocations largely targeted for resistance to diseases and pests (section 2.2.2). This source of diversity will continue to support crop improvement. Wheat is highly responsive to technological solutions to climate change. Modeled potential benefits from genotypic adaptation to future climates are in the order of 60% for wheat compared with 12% for both maize and rice (Ramirez-Villegas et al. 2020).

WHEAT focuses on delivering varieties and practices to improve adaptation to heat and drought, diseases and insects, with prebreeders and breeders improving yield potential and adaptation to biotic and abiotic stresses; agronomists helping to close the yield gap by delivering technologies tailored to local conditions; economists and social scientists framing agR&D in contexts of gender and social inclusion and guiding more effective allocation of limited resources (sections 1.2.2 and 2.1.3.1); and all of them contributing to capacity building in target countries and regions (sections 2.1.3.1 and 2.3). WHEAT scientists are world class and resilient; they have a track record of delivering solutions to improve the livelihoods of wheat farmers and the productivity of their fields (sections 2.1.3.2, 2.2.1, and 2.2.2) despite the unsettling cycles of change in CGIAR. The program's competitive advantages include purpose-driven global and regional networks (sections 2.1.2.3 and 2.3) and a pipeline of innovation in the two pillars of breeding and agronomy illustrated in the next sections.

Modernization of the breeding process. A process has been initiated to accelerate the rate of genetic gain for grain yield and other traits under the guidelines set by CtEH and EiB with the support of two large W3 projects ensuring continuity for the next five years, giving a sense of stability to scientists and partners. A sample of deliverables includes gene editing technology to address the needs of jurisdictions not served by multinational seed businesses by leveraging cutting-edge proprietary technology from the private sector; DNA fingerprinting to objectively quantify rate of adoption, with spinoffs such as tracking the movement of seed at different scales; new opportunities from centralization, curation, standardization and application of data; methods to apply sparse testing and other efficiency-enhancing experimental designs; genomic selection, index selection and molecular markers; enhanced conservation and use of genetic resources with a broadened network (CIMMYT GRP, CIMMYT genebank, ICARDA, INIFAP-Mexico, INTA-Argentina, Universidad Politécnica de Madrid, ARO-Volcani Center, National Gene Bank of Tunisia, Genetic Resources Institute of Azerbaijan).

Integrated approaches to sustainable intensification. Excellence in Agronomy, an incipient platform, responds to public and private demand for scalable agronomic solutions as an engine for agricultural development. It combines big data, sensing technologies, geospatial decision tools, and innovative partnerships to improve spatially explicit agronomic recommendations. Science is informed by sustainability, climate change, behavioral economics, and scaling pathways. A 2-year incubation phase (2020–2022) has been co-developed between nine CGIAR centers. Excellence in Mechanization, a less-mature initiative, presents an entry point to profitability; post-harvest processing; access to markets; and efficient use of resources, including energy, nutrients, water, and labor, with implications for gender and drudgery. Labor is a major limiting factor for smallholder productivity (Baudron et al. 2019a; Dahlin and Rusinamhodzi 2019) and will be more pressing in the post-pandemic world—hence the critical role of targeted mechanization.

2.4.4 Management, Monitoring, and Evaluation with Purpose

In international development programs, the prevailing focus of evaluation is often the program model. Programs such as WHEAT are instructed to make meaning of what is done by reducing complex dynamic systems into linear logic. This approach limits WHEAT (and the CGIAR as a whole) to effectively engage with, and assess what WHEAT aims to address, such as poverty, climate change, and sustainability. Traditional program evaluation (e.g. indicators, assessing planned versus achieved) are necessary (i.e. donors demand it) yet not necessarily useful for evaluating WHEAT's efforts towards achieving dynamic, adaptive, and resilient sustainability.

By forcing WHEAT to assess its merit with traditional project boxes aimed at standardization, predictability, and simple, linear attribution, donors, managers, and evaluators inhibit innovation, adaptation, and responsiveness (Box 1). In the future, we suggest combining more appropriate approaches, namely Principles Focused Evaluation (Patton 2018) and Developmental Evaluation (Patton 2011), in the context of Alston's equation.

- Principles-focused evaluation examines the extent to which statements of principles (1) provide meaningful guidance, (2) are useful in decision-making, (3) are inspirational, (4) support adaptation and development, and (5) are evaluable.
- Developmental Evaluation is a strategic learning tool that supports innovation and social change in complex or uncertain environments. The evaluator's role facilitates real-time, or close to real-time, feedback in a continuous, empirical, learning, improvement, and development loop.

These two approaches are complementary and method neutral and can work in tandem with traditional indicator-focused approaches.

3 Conclusions

3.1 Quality of Science

With approximately 80 scientists, WHEAT relies on collaborations to tap global science and technology. Over decades, WHEAT has catalyzed a worldwide network of R&D that has delivered and continues to deliver a disproportionate wealth of outputs in relation to investment. More rigorous study is required to test the hypothesis of a scale-free network and the analysis of its implications. Owing to the dominant role of public R&D, the tension between competition and collaboration is intrinsically lower in wheat than in other industries. Nonetheless, it needs to be managed. CAIGE is an excellent model, illustrating the synergies from international collaborations, the opportunities to engage private actors in a pre-competitive setting, and the importance of rapport between leaders. It may be an unsuitable model for many countries of interest, but it might apply to some of them, e.g., India.

WHEAT scientists consistently rate generation of knowledge as a major output and demonstrate they are indeed pushing boundaries in some fields. This is reassuring because routine work would starve technological innovation and compromise the achievement of high-level outcomes. We identified opportunities to improve quality of science. Scientific credibility is increasingly important to connect public policy and technological change in our society, where misinformation and disinformation prevail in the public sphere (Iyengar and Massey 2019).

3.2 Effectiveness

3.2.1 Achievement of Planned Outputs and Outcomes

Between 2017 and 2019 WHEAT demonstrated achievements in each area of its ToC, and within each FP. As with any program of this nature, not all planned outputs were achieved in the year planned, while unplanned achievements occurred. Over the period under study, the CRP did not drop any research line and the FPs' impact pathways and ToCs remain unchanged. These findings suggest that WHEAT is largely effective in contributing towards moving the global wheat program towards higher level achievements, and thus doing what they said would do (the main effectiveness criterion).

3.2.2 Demonstrated Importance of Outcomes

Three cases provided a small window into the importance of WHEAT's outcomes. These cases support WHEAT's relevance to global wheat challenges and opportunities in agronomy and breeding. Exploration of the three cases identified the critical importance of supporting long-term strategies and the critical role of multiple partnerships both in research, publishing and in funding. While the section aimed to explore the importance of WHEAT's outcomes, the cases highlighted the significant role played by WHEAT (CIMMYT and ICARDA) in the sector, the critical nature of long-term strategies (here 30 years), and the essence of partnerships at multiple levels, which are critical to achieve relevant and substantial outcomes.

3.2.3 CRP Management and Governance

Despite some minor challenges mostly related to funding, limited evaluation data suggest that WHEAT's roles, responsibilities, and accountabilities are clearly defined and exercised, and governance is sufficiently independent. While monitoring and evaluation had some challenges, such as introducing indicators and new terminology midway through, it appears to have adapted and adequately supports the CRP, its flagships, and related reporting requirements. While the amount of reporting and frequently changing requirements are viewed as burdensome to WHEAT and its scientists, most interview data suggest that these requirements are the nature of CGIAR and its current context.

In addition, WHEAT has demonstrated a strategic approach to leveraging outputs from regionally focused bilateral funding to effectively achieve global-level outcomes. Further, WHEAT has leveraged advances in upstream academic research to pilot and apply in public wheat R4D. Some evidence suggests that without the flexibility of CRP resources, many opportunities would have been missed or delayed.

3.2.4 Progress along ToC (CRP and Flagships)

The ToC provides explicit shared thinking about how change comes about in a larger context and has several uses. **First**, it is useful to prioritize and then select key areas to conduct research within an already established and agreed-upon ToC. **Second**, it clearly demonstrates how outputs contribute to the scientific, technological, and social changes. **Third**, it supports an explicit understanding with regard to what needs to be addressed in order to reach the SLOs and supports the CRP to seek and justify funding for these needs. **Fourth**, it provides a map that that can be used to explicitly identify responsibilities to achieve the SLOs (i.e., WHEAT is responsible for A, and another actor for B). **Fifth**, it provides a clear framework to report against, which clarifies how WHEAT results contribute to sectoral changes. Therefore, while a ToC is important for several reasons, it is unsuitable for assessing (judging) the progress of WHEAT toward its SLOs (as in how far it made it toward the top) and its flagships. Stated another way, it is not useful to assess how far "up the chain" WHEAT has identified accomplishments (the ToC does not appear to be defined for this purpose); rather, it is useful for exploring what WHEAT accomplished within its selected areas of the ToC, how that contributed to the global effort, and why it's important.

3.3 Cross-cutting Issues (Capacity Development, Partnership, Gender, Youth)

The WHEAT CRP engages with all four cross-cutting issues. Capacity development and partnerships are used as mechanisms for change, which are lumped together with gender and youth (which are then also at times combined) and added on. Partnerships are critical to WHEAT's identified achievements, and various kinds of partners (many of which are long-term partners) are almost always reflected in these achievements. Youth and gender provide focus areas for research, beneficiaries of the research, or are targeted in capacity development efforts. Significant yet often siloed findings were identified for gender, with fewer youth-focused results identified in the three-year period. The critical finding, however, is not how WHEAT engages with cross-cutting issues, but how cross-cutting issues are valued systemically; with strategies and reports reviewed for this evaluation, and even in this evaluation, suggesting a strong focus on science, and then the "rest" (gender, youth and social inclusion) without a comprehensive (or seemingly sustained effort) for engagement with how each supports and benefits the science achievements.

3.4 Future Orientation

Wheat as a crop is bound to be central to global food security in the foreseeable future. WHEAT as an R&D agent has a track record of delivering local solutions with a global perspective and is well positioned to continue this trajectory in the next decade. There are opportunities and challenges for the way ahead, including the risk of fragmenting the global breeding program, restrictions to exchange germplasm and ideas, the opportunity to integrate R&D in agronomy, misguided emphasis on minor crops, and CGIAR's focus on process at the expense of results..

4 Recommendations

Here we recap key findings that require recommendations for specific user groups; the targets of these recommendations are managers and researchers, unless specified. Quality of Science is discussed first, followed by Effectiveness. In each section, findings and their recommendations are provided in order of importance.

4.1 Quality of Science

Key finding: With approximately 80 scientists, WHEAT is subcritical. Over decades, WHEAT has catalyzed a global network of R&D that has delivered and continues to deliver a disproportionate wealth of outputs in relation to investment. Partnerships, and WHEAT reputation as a reliable partner, are vulnerable to funding volatility.

Recommendation: Support strategic investment in network development and maintenance (donors and decision-makers).

Key finding: The network of partners might be scale free – i.e., driven by preferential attachment or the rich-get-richer, with implications for (1) inclusiveness – it may be hard for outsiders to enter the network, and (2) resilience – the system is resilient to random perturbation, but vulnerable to disruption of large nodes such as the centralized breeding system.

Recommendation: Investigate the nature of the network using larger samples and complementary metrics beyond authorship (scientists, managers). Use this information to protect the network from both random and targeted attacks (managers, science leaders). Consider opportunities for expanding networks beyond current nodes (scientists).

Key finding: Publications in Quartile 1 journals (Q1, top 25%) were world-class and some pushed scientific boundaries; publications in Q2 journals were sound but often routine work; publications in Q3Q4 journals commonly featured flaws.

Recommendation: Set targets (time frame, rates) to shift a proportion of Q2 papers to Q1. Set targets (time frame, rates) to phase out Q3 and Q4 papers; set up mentoring systems to avoid work that leads to lower-quality papers in the first place. Revise evaluation and reward system to improve the quality-to-volume ratio of scientific output—i.e., fewer but better papers in the context of different stages in the scientist's career. All these recommendations aim at science leaders.

Key finding: CGIAR's focus on process risks compromising results.

Recommendations: Use Alston's equation (Box 1) to avoid disproportionate process (donors, managers). Constructively argue against unreasonable bureaucratic requirements (scientists) and openly listen to those arguments (donors, managers). Defend (scientists) and support (managers, donors) the 20:80 rule to allow scientists about 20% time for exploratory research.

Key finding: Impact assessments require independent specialists and funding but must not distract resources from R&D core objectives.

Recommendation: An independent unit for transversal impact assessment should be funded with new money (One CGIAR decision-makers). Assessments must be proportional to the R&D effort, not becoming an end in themselves (donors, managers). *Ex ante* assessments need to acknowledge that the sample space of technological innovation is undefinable (section 1.2.4 and Kauffman 2008, 2016).

Key finding: WHEAT comprises ~80 scientists, with the ratio of technical-to-scientific staff from 8.9 to 2.1 among flagships, and the ratio of administrative-to-productive staff from 0.04 to 0.10; these ranges are commensurate with the focus and geographical spread of staff of the flagships.

Recommendation: Regularly check the evolution of the technical-to-scientific and administrative-to-productive staff ratio to ensure support to R&D is proportional to the focus and needs of each FP.

4.2 Effectiveness

Key findings: In 2017–2019, WHEAT mainly achieved its planned outputs and outcomes, and in addition had unplanned outcomes. For the three years reviewed, WHEAT did not drop any research line. Slight changes included WHEAT adding research priorities on mechanization (FP4) and soil-borne diseases (FP3) in 2017.

Recommendation: Establish how WHEAT, or any CRP, will be assessed for effectiveness when the proposal is submitted and approved, and set clear criteria for judging effectiveness. For alternative suggestions for assessing WHEAT's effectiveness, which can also be used for improving the intervention, management decisions, and judgment, see section 2.4.

Key finding: The review of three OICRS suggested that WHEAT's R&D in agronomy and breeding is relevant to global wheat challenges such as climate change, mechanization, and farmers' income. The cases also demonstrated the importance of both partnerships and long-term (30 year) strategies.

Recommendation: None related to the cases. Recommendation is to focus on more useful approaches to assessing, understanding, and identifying outcomes. See section 2.4.

Key findings: WHEAT's roles, responsibilities, and accountabilities are clearly defined and exercised, and governance is sufficiently independent. Some slight challenges related to funding exist. Monitoring and evaluation had some challenges yet appear to adequately support the CRP. Redundant reporting and frequently changing requirements are burdensome.

Recommendation: Indicators and other approaches that use quantitative data need to continue, as they are expected and considered necessary by most donors. Reducing the focus on these, and increasing the use of impact case studies (which include numbers and words, or statistics and narratives) that focus on countries or themes (e.g., Ethiopia or mechanization) will likely bring broader perspectives, a more informed understanding of WHEAT's effectiveness, and therefore improved decision making. See section 4.3 and the recommendation above for an independent assessment unit.

Key finding: "Gender" is used as a research focus, and "gender" is used to identify a research or capacity development beneficiary; gender achievements in research, while often notable, were often siloed.

Recommendation: Select a few key gender findings that are useful to specific FP interventions or research, integrate these findings, and ensure that these gender aspects are included in effectiveness assessments. Require that gender statistics are collected and reported for all training, workshops, and conferences; use statistics to better understand and improve participation levels where appropriate, not as a box-ticking exercise.

Key finding: The CGIAR approved WHEAT's approach to youth (along with gender), and results on youth reflect that agreement, with a notable achievement in 2017 that informs youth-oriented R4D.

Recommendation: Use the 2017 research paper to consider how to bundle youth-relevant R4D across CRPs, and provide a core budget.

Key finding: The ToC provides explicit shared thinking about how change comes about in a larger context and is useful for (1) priority setting, (2) assessing the contribution of scientific outputs, (3) seeking and justifying funding, (4) mapping trajectory to impact, and (5) reporting, but is unsuitable for (6) assessing WHEAT's or its flagships' effectiveness by judging their progress toward the SLOs.

Recommendation: Continue to refine the ToC as needed and recognize its five uses in its current form. Do not assess the CRP's progress toward the identified SLOs to judge WHEAT's effectiveness. Stated another way, it is not useful to assess how far "up the chain" WHEAT has identified accomplishments; rather, explore what WHEAT accomplished within its selected areas of the ToC, how these areas were selected, how results contributed to the global effort, and why it is important.

4.3 Future Orientation

Key finding: Ensure support to both modernization of the breeding process and integrated approaches to sustainable intensification including mechanization. For management, monitoring, and evaluation with purpose, consider integrating three elements based on three well-known social and scientific theories: (1) Principle-focused Evaluation to examine the extent to which statements of principles provide meaningful guidance, are useful in decision-making, are inspirational, support adaptation and development, and are evaluable; (2) Developmental Evaluation as a strategic learning tool that supports innovation and social change in complex or uncertain environments; and (3) Alston's equation – evaluations before, during and after investments are critical but must not suffocate scientists.

4.4 CGIAR System-level Recommendation

Our core recommendations above targeted not only the CRP but also donors and higher-level management in CGIAR. Further, given that the CRP phase will end in 2021, little time remains for any shifts in CRP management, governance, or resources/budgets. As one respondent noted,

> *Everyone is waiting for One CGIAR, and we will then make shifts accordingly.*

5 References

Allen-Sader C et al. (2019) An early warning system to predict and mitigate wheat rust diseases in Ethiopia. Environmental Research Letters 14:115004

Alston JM, Norton GW, Pardey PG (1995) Science under scarcity: principles and practice for agricultural research evaluation and priority setting. Cornell University Press, Ithaca

Asseng S et al. (2017) Hot spots of wheat yield decline with rising temperatures. Global Change Biology 23:2464-2472

Bak P (1996) How nature works: the science of self-organised criticality. Copernicus, Springer-Verlag, New York

Balwinder S, McDonald AJ, Srivastava AK, Gerard B (2019) Tradeoffs between groundwater conservation and air pollution from agricultural fires in northwest India. Nature Sustainability 2:580-583

Barabási A-L, Albert R (1999) Emergence of Scaling in Random Networks. Science 286:509-512

Baudron F, Duriaux Chavarría J-Y, Remans R, Yang K, Sunderland T (2017) Indirect contributions of forests to dietary diversity in Southern Ethiopia. Ecology and Society 22

Baudron F, Misiko M, Getnet B, Nazare R, Sariah J, Kaumbutho P (2019a) A farm-level assessment of labor and mechanization in Eastern and Southern Africa. Agronomy for Sustainable Development 39

Baudron F, Tomscha SA, Powell B, Groot JCJ, Gergel SE, Sunderland T (2019b) Testing the Various Pathways Linking Forest Cover to Dietary Diversity in Tropical Landscapes. Frontiers in Sustainable Food Systems 3

Bhatta M et al. (2018) Genome-Wide Association Study Reveals Novel Genomic Regions Associated with 10 Grain Minerals in Synthetic Hexaploid Wheat. International Journal of Molecular Sciences 19:3237

Boon J, Wynen J (2017) On the bureaucracy of bureaucracies: Analysing the size and organization of overhead in public organizations. Public Administration 95:214-231

Broido AD, Clauset A (2019) Scale-free networks are rare. Nature Communications 10

Budden AE, Tregenza T, Aarssen LW, Koricheva J, Leimu R, Lortie CJ (2008) Double-blind review favours increased representation of female authors. Trends in Ecology & Evolution 23:4-6

Crespo-Herrera LA et al. (2017) Genetic Yield Gains In CIMMYT's International Elite Spring Wheat Yield Trials By Modeling The Genotype × Environment Interaction. Crop Science 57:789-801

Cruz CD et al. (2016) The 2NS Translocation from Confers Resistance to the Pathotype of. Crop Science 56:990

Dahlin AS, Rusinamhodzi L (2019) Yield and labor relations of sustainable intensification options for smallholder farmers in sub-Saharan Africa. A meta-analysis. Agronomy for Sustainable Development 39

Dreisigacker S et al. (2019) Tracking the adoption of bread wheat varieties in Afghanistan using DNA fingerprinting. BMC Genomics 20

FAO (2016) FAOSTAT. Statistics Division of the Food and Agriculture Organization of the United Nations http://www.fao.org/faostat/en/#home.

Fischer RA, Byerlee D, Edmeades GO (2014) Crop yields and global food security. Will yield increase continue to feed the world? ACIAR, Canberra

Fischer RA, Connor DJ (2018) Issues for cropping and agricultural science in the next 20 years. Field Crops Research 222:121–142

Foulkes MJ, Reynolds MP, Sylvester-Bradley R (2009) Genetic improvement of grain crops: yield potential. In: Sadras VO, Calderini DF (eds) Crop physiology: applications for genetic improvement and agronomy. Academic Press, San Diego, pp 355-385

Franzoni C, Scellato G, Stephan P (2011) Changing Incentives to Publish. Science 333:702-703

Hall AJ, Richards RA (2013) Prognosis for genetic improvement of yield potential and water-limited yield of major grain crops. Field Crops Research 143:18-33

Hanel R, Kauffman SA, Thurner S (2005) Phase transition in random catalytic networks. Physical Review E 72

Hirsh JE (2005) An index to quantify an individual's scientific research output. Proceedings of the National Academy of Science of USA 102:16569

ISDC (2020) Quality of Research for Development in the CGIAR Context, Technical Note January 2020.

ISPC (2017) Quality of Research for Development in the CGIAR Context, Brief N. 62. Rome: Independent Science and Partnership Council.

Iyengar S, Massey DS (2019) Scientific communication in a post-truth society. Proceedings of the National Academy of Sciences of the United States of America 116:7656-7661

Juliana P et al. (2019a) Integrating genomic-enabled prediction and high-throughput phenotyping in breeding for climate-resilient bread wheat. Theoretical and Applied Genetics 132:177-194

Juliana P et al. (2019b) Improving grain yield, stress resilience and quality of bread wheat using large-scale genomics. Nature Genetics 51:1530-1539

Kakraliya SK et al. (2018) Performance of portfolios of climate smart agriculture practices in a rice-wheat system of western Indo-Gangetic plains. Agricultural Water Management 202:122-133

Kauffman SA (2008) Reinventing the Sacred: A New View of Science, Reason, and Religion Basic Books, Philadelphia

Kauffman SA (2016) Humanity in a creative universe. Oxford University Press

Keil A, D'Souza A, McDonald A (2017) Zero-tillage is a proven technology for sustainable wheat intensification in the Eastern Indo-Gangetic Plains: what determines farmer awareness and adoption? Food Security 9:723-743

Krupnik TJ, Andersson JA, Rusinamhodzi L, Corbeels M, Shennan C, Gérard B (2019) Does size matter? A critical review of meta-analysis in agronomy. Experimental Agriculture 55:200-229

Lantican MA et al. (2016) Impacts of International Wheat Improvement Research, 1994-2014. CIMMYT, Mexico, D.F.

Li H, Rasheed A, Hickey LT, He Z (2018) Fast-Forwarding Genetic Gain. Trends in Plant Science 23:184-186

Lopez-Ridaura S, Frelat R, Van Wijk MT, Valbuena D, Krupnik TJ, Jat ML (2018) Climate smart agriculture, farm household typologies and food security. Agricultural Systems 159:57-68

Meyer M et al. (2017) Quantifying airborne dispersal routes of pathogens over continents to safeguard global wheat supply. Nature Plants 3:780-786

Mottaleb KA, Singh PK, He X, Hossain A, Kruseman G, Erenstein O (2019) Alternative use of wheat land to implement a potential wheat holiday as wheat blast control: In search of feasible crops in Bangladesh. Land Use Policy 82:1-12

Mottaleb KA et al. (2018) Threat of wheat blast to South Asia's food security: An ex-ante analysis. PLOS ONE 13:e0197555

Mottaleb Khondoker A, Rahut Dil B, Kruseman G, Erenstein O (2018) Evolving food consumption patterns of rural and urban households in developing countries: A Bangladesh case. British Food Journal 120:392-408

Neff MW (2020) How academic science gave its soul to the publishing industry. Issues in Science and Technology 36:35-43

Orr A, Donovan J (2018) Introduction to special issue: smallholder value chains as complex adaptive systems. Journal of Agribusiness in Developing and Emerging Economies 8:2-13

Osmond CB (1995) Quintessential inefficiencies of plant bioenergetics: Tales of two cultures. Australian Journal of Plant Physiology 22: 123-129

Patton MQ (2008) Utilization-Focused Evaluation. 4th edition. Thousand Oaks, Ca: Sage Publications

Patton MQ (2011) Developmental evaluation: Applying complexity concepts to enhance innovation and use. Guilford Press., New York, NY

Patton MQ (2018) Principles-Focused Evaluation: The GUIDE. Guilford Press, New York

Rahut DB, Ali A (2017) Coping with climate change and its impact on productivity, income, and poverty: Evidence from the Himalayan region of Pakistan. International Journal of Disaster Risk Reduction 24:515-525

Ramirez-Villegas J et al. (2020) CGIAR modeling approaches for resource-constrained scenarios: I. Accelerating crop breeding for a changing climate. Crop Science 60:547-567

Rasheed A, Ogbonnaya FC, Lagudah E, Appels R, He Z (2018) The goat grass genome's role in wheat improvement. Nature Plants 4:56-58

Rasmussen LV et al. (2019) Forest pattern, not just amount, influences dietary quality in five African countries. Global Food Security:100331

Rosyara U et al. (2019) Genetic Contribution of Synthetic Hexaploid Wheat to CIMMYT's Spring Bread Wheat Breeding Germplasm. Scientific Reports 9

Sadras VO (2019) Effective Phenotyping Applications Require Matching Trait and Platform and More Attention to Theory. Frontiers in Plant Science 10

Sadras VO et al. (2020) Making science more effective for agriculture. Advances in Agronomy 163

Sapkota TB, Aryal JP, Khatri-Chhetri A, Shirsath PB, Arumugam P, Stirling CM (2018) Identifying high-yield low-emission pathways for the cereal production in South Asia. Mitigation and Adaptation Strategies for Global Change 23:621-641

Shamanin V et al. (2019) Primary hexaploid synthetics: Novel sources of wheat disease resistance. Crop Protection 121:7-10

Shyamsundar P et al. (2019) Fields on fire: Alternatives to crop residue burning in India. Science 365:536-538

Singh RP, Chintagunta AD, Agarwal DK, Kureel RS, Kumar SPJ (2019) Varietal replacement rate: Prospects and challenges for global food security. Global Food Security:100324

Slafer GA (2008) Should crop scientists consider a journal's impact factor in deciding where to publish? European Journal of Agronomy 29:208-212

Stoian D, Donovan J, Elias M, Blare T (2018) Fit for purpose? A review of guides for gender-equitable value chain development. Development in Practice 28:494-509

Tadesse W et al. (2019) Wheat breeding handbook at ICARDA. https://hdl.handle.net/20.500.11766/10723

Tainter JA (1988) The collapse of complex societies. Cambridge University Press, Cambridge

Thierfelder C et al. (2018) Complementary practices supporting conservation agriculture in southern Africa. A review. Agronomy for Sustainable Development 38

Van Loon J, Woltering L, Krupnik TJ, Baudron F, Boa M, Govaerts B (2020) Scaling agricultural mechanization services in smallholder farming systems: Case studies from sub-Saharan Africa, South Asia, and Latin America. Agricultural Systems 180:102792

Velu G et al. (2018) Genetic dissection of grain zinc concentration in spring wheat for mainstreaming biofortification in CIMMYT wheat breeding. Scientific Reports 8

Voss-Fels KP et al. (2019) Breeding improves wheat productivity under contrasting agrochemical input levels. Nature Plants 5:706-714

Watson JD, Crick FHC (1953) Molecular structure of nucleic acids - a structure for deoxyribose nucleic acid. Nature 171:737-738

Weiner J (2019) Looking in the wrong direction for higher-yielding crop genotypes. Trends in Plant Science in press

Wilhite AW, Fong EA (2012) Coercive citation in academic publishing. Science 335:542-543.

.

CGIAR Advisory Services

CGIAR Advisory Services (CAS) Secretariat

Via dei Tre Denari, 472/a, Maccarese (Fiumicino), Italy

tel: (39) 06 61181 - email: cas@cgiar.org

https://cas.cgiar.org/