

1 **Evaluation capacity building in response to the agricultural research impact agenda: Emerging**
2 **insights from Ireland, Catalonia (Spain), New Zealand, and Uruguay**

3 **Abstract**

4 Performance-based funding and calls for public-funded science to demonstrate societal impact are
5 encouraging public research organisations to evaluate impact, the so-called impact agenda. This paper
6 explores evaluation methods of four fully or partially public-funded agricultural research organisations
7 and how they are building evaluative capacity to respond to the impact agenda. Drawing on cross-
8 organisational comparison of the readiness of each organisation to implement evaluation, the
9 implications for improving evaluative capacity building (ECB) are discussed. This study extends the
10 current literature on ECB, as very little has focussed on research organisations in general, and
11 particularly agricultural research.

12 Driven by the impact agenda, the organisations are beginning to emphasise summative evaluation.
13 Organisational leaders valuing the demonstration of impact and commitment to building evaluation
14 capacity are important precursors to other aspects of organisational readiness to implement
15 evaluation. However, organisational emphasis remains on using evaluation for accountability and to
16 improve efficiency and allocation of funding. The organisations have yet to systematically embed
17 evaluation processes and capabilities for learning at programme and organisation-levels. There is,
18 therefore, an opportunity to develop organisation and programme-level evaluation processes that
19 inform each other and the pathways to impact from science.

20 To realise this opportunity, organisations could strengthen internal and external networks of
21 evaluation practitioners and academics to bridge the gap between the theory and practice of
22 monitoring and evaluation for learning (MEL) and to begin to reshape organisational culture by using
23 evaluation methods that are grounded in co-production and integrated scientific and societal values.

24 **Keywords:** agricultural research impact; monitoring and evaluation, evaluation capacity building

25 **Introduction**

26 Over the past two decades there has been an increasing call for public-funded science to deliver and
27 demonstrate impact (Penfield et al., 2014), going beyond the excellence of academic outputs to the
28 benefits these produce for society (Donovan, 2011). This is challenging the social contract with science
29 through which scientific freedom was exchanged for the expectation of socially beneficial impacts
30 (Joly et al., 2016; Owen et al., 2012). Simultaneously, the diffusion of new public management rules
31 (Gauand et al., 2015) has encouraged performance-based funding for public-funded science (New
32 Zealand Ministry of Business, Innovation and Employment, 2015; Organisation for Economic Co-
33 operation and Development, 2010).

34 Delivering and demonstrating impact is especially important for agricultural science (Midmore, 2017),
35 with agriculture at the nexus of significant societal challenges, such as food security, rural resilience,
36 and environmental sustainability (Tilman et al., 2002). Given the nature of these challenges
37 agricultural science needs to deliver social, cultural, and environmental benefits (Donovan, 2011;
38 Kelley et al., 2008), as well as economic (Joly et al., 2015). For the purposes of this study, research
39 impact is defined as: “the demonstrable contribution that research makes to the economy, society,
40 culture, national security, public policy or services, health, the environment, or quality of life, beyond
41 contributions to academia” (Australian Research Council cited in Midmore (2017)).

42 The reasons research organisations may evaluate science impact can be organised into the four “As”:
43 advocacy, accountability, allocation, and analysis (Morgan Jones, Manville, and Chataway 2017;
44 Penfield et al., 2014):

- 45 1. Advocacy to demonstrate the value of research to government, stakeholders, and the public
46 (Joly et al., 2015; Penfield et al., 2014).
- 47 2. Accountability to funders, and more broadly society, by monitoring and managing research
48 organisation performance to contribute to society (Penfield et al., 2014).
- 49 3. To inform funding allocation, based on the potential value of research (Penfield et al., 2014).

50 4. Analysis to understand how agricultural science generated contributes to changes in practices
51 and policies in the agri-food system (Joly et al., 2016; Midmore, 2017; Pollock, 2012).

52 The “impact agenda” therefore calls for both summative (advocacy and accountability) and formative
53 (allocation and analysis) functions of evaluation (Joly et al., 2015). Formative evaluation takes place
54 during the development of a programme, with the intention of improving the value or effectiveness
55 of the programme, while summative draws lessons from a completed programme (Superu, 2017a).

56 While agricultural research organisations are in the emergent stages of meeting the impact agenda,
57 other research organisations have made progress (e.g. CSIRO, 2015; Joly et al., 2016; Kelley et al.,
58 2008; Maredia et al., 2014; Menon et al., 2009). The Consultative Group on International Agricultural
59 Research (CGIAR) has been undertaking ex-post Impact Assessments since the early 1970s (Watts et
60 al., 2008), including non-economic impacts (Kelley et al., 2008). Utilising a suite of evaluation methods
61 (Kelley and Gregersen, 2003), CGIAR has moved from an accountability focus to evaluation for learning
62 (like analysis) (Watts et al., 2008) and includes participation of stakeholders and organisations in
63 evaluation (Mackay and Horton, 2003; Watts et al., 2008; Mayne and Johnson, 2015).

64 Research organisations working to meet the impact agenda face several challenges including:
65 evaluation of multiple possible impact pathways (Joly et al., 2016); multiple impacts (Bozeman and
66 Sarewitz, 2011; Kelley et al., 2008) at different levels from programme to organisation; attributing
67 impacts in multi-actor innovation systems (Douthwaite and Hoffecker, 2017; Horton and Mackay,
68 2003; Joly et al., 2016); and time-lags between research outputs and impacts (Kelley et al., 2008). To
69 address these challenges agricultural research organisations have begun to invest in evaluation
70 capacity building (ECB) (Joly et al., 2016; Maredia et al., 2014; Stone-Jovicich et al., 2019), which is
71 defined as “the intentional work to continuously create and sustain overall organisational processes
72 that make quality evaluation and its use routine” (Hueftle Stockdill et al., 2002, p.14). The goals of ECB
73 include designing, undertaking, and managing evaluation projects; developing and using evaluation
74 knowledge and skills; encouraging organisational learning as an ongoing process; organisational

75 accountability; creating awareness and support for evaluation within the organisation (Cousins et al.,
76 2014; Preskill and Boyle, 2008); and sustaining evaluation practice over time (Wade & Kallemeyn,
77 2020).

78 However, to what extent, and how, are public-funded agricultural research organisations building
79 evaluation capacity to respond to the impact agenda? To address this question our aim is to describe
80 the current evaluative capacity of four partially to fully public-funded agricultural research
81 organisations in terms of:

- 82 1. The evaluation methods these organisations are using; and
- 83 2. Their comparative evaluation capacity.

84 This article proposes to extend the current body of literature on ECB, as very little has focussed on
85 research organisations in general, and agricultural research in particular (Joly et al., 2016). As far as
86 we are aware, this is the first time that four publicly funded research organisations have documented
87 their approaches to ECB. Drawing on our cross-organisational comparison of evaluation capacity,
88 implications are discussed to improve ECB within agricultural research organisations.

89

90

91

92

93

94 **Background**

95 This section provides a brief background to each of the four research organisations that participated
96 in this study. The four organisations established a learning network on evaluation capacity building
97 that was operationalised with workshops in 2017 and 2019 that provided data for the analysis in this

98 paper (see Methodology). The network was based on the organisation’s commonalities, particularly
 99 public research organisations in the agri-food sector focusing on pastoral systems.

100 Each of the organisations is described in terms of their purpose, agri-food sectors served, key research
 101 and extension capabilities, major funding sources (Table 1), overarching evaluation frameworks and
 102 core principles, key characteristics of their organisational impact agenda (Table 2), and timeline of
 103 activities to build evaluative capacity (Figures 1 to 4). The description of the organisations is for the
 104 period 2015-2018. Given the dynamic nature of each organisation the information presented here will
 105 differ from the current reality.

106 **Table 1.** Major funding sources for each case study organisation – average for 2015-2018

Income Source	Contribution to the total (%)			
	AgResearch	INIA	IRTA	Teagasc
Structural ¹ Funds ¹	27.0	84.0	32.0	69.0
Competitive ²	17.6	5.0	12.0	17.0
Contracts for services and product sales ³	44.1	10.0	39.0	14.0
Subsidies ⁴	0.0	0.0	7.0	0.0
Other funds	11.2	1.0	10.0	0.0

107 ¹Funding allocated directly to the organisation to then allocate, ² Funds that are bid to a funder for, ³
 108 One off contracted projects, e.g. for an agricultural company, and ⁴ Funds that are received from the
 109 provincial councils and for funding researchers from national and regional programs. The availability
 110 of structural research funding from the state depends on the organisational and funding
 111 characteristics of each country’s science system.

112

113

114 ***Teagasc, Ireland***

115 Teagasc – the Agriculture and Food Development Authority – is the national body providing integrated
 116 research, advisory and education services to the agriculture and food industry and rural communities
 117 in Ireland. It is a key node in Ireland’s Agricultural Knowledge & Innovation System (AKIS) and has long-
 118 established and strong collaborations with many other public and private sector AKIS actors. Teagasc

119 has 1,260 staff based in seven research centres, 52 advisory offices and four colleges of further
120 education around the country.

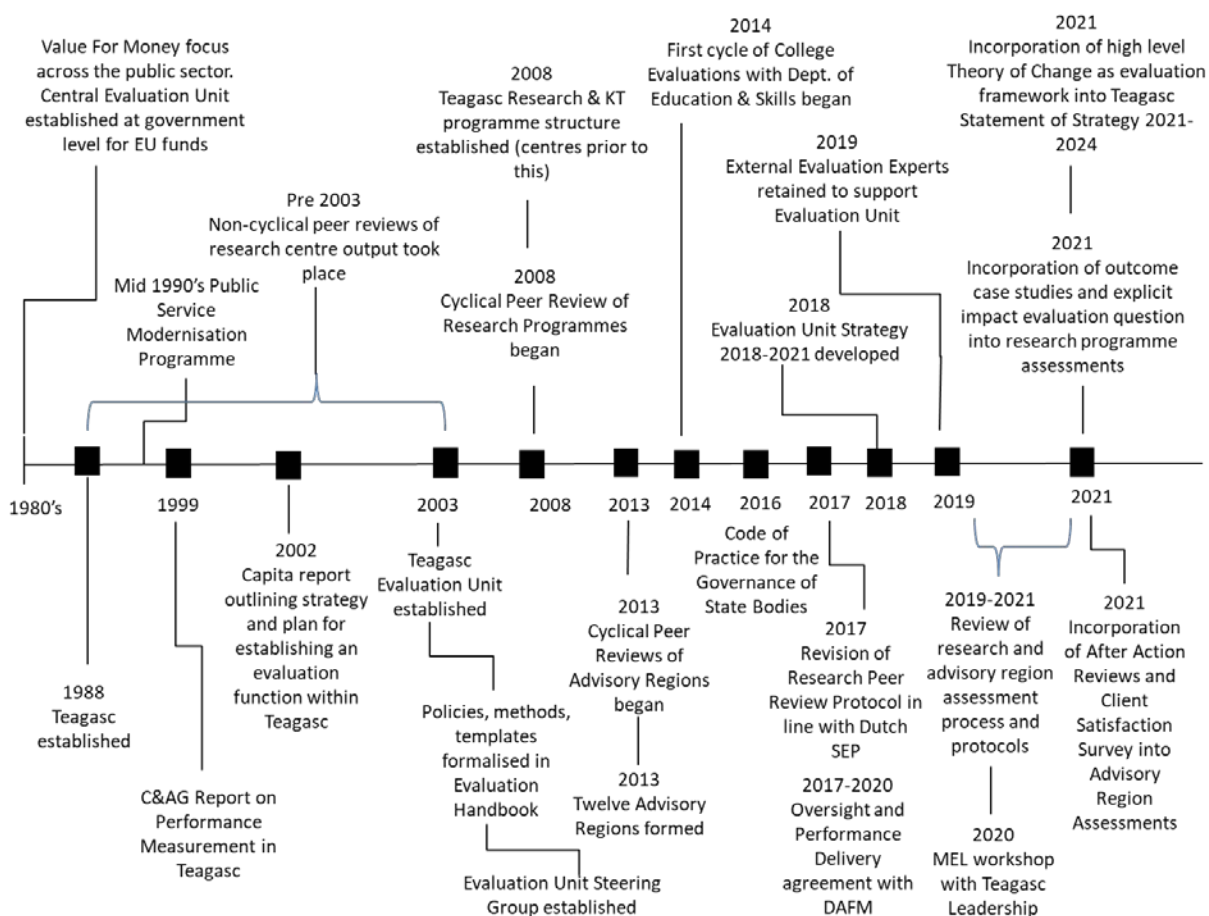
121 An Evaluation Unit was established in Teagasc in 2003 in response to recommendations made by the
122 Comptroller and Auditor General (C&AG) in a 1999 report on Performance Measurement in Teagasc
123 (see Figure 4). In 2002, consultants (Capita) outlined a strategy and plan for establishing an evaluation
124 function across Teagasc, not just research evaluation. Although no formal Evaluation Unit existed in
125 the organisation before 2003, research and advisory programmes were reviewed, and other
126 organisational functions evaluated, over the years since Teagasc was established in 1988.

127 After 2003, a regular cycle of peer assessments of research centres, then research programmes
128 (established in 2008) and advisory regions (established in 2013) began. Organisational functions were
129 evaluated on a non-cyclical basis. In 2014, Teagasc approached the Department of Education & Skills
130 to develop an evaluation framework for its colleges. Externally, in 2016, the Code of Practice for the
131 Governance of State Bodies issued by the Department of Public Expenditure and Reform led to the
132 Oversight and Performance Delivery Agreement (2017 – 2020) with the Department of Agriculture,
133 Food and the Marine (DAFM). This Agreement formalises the arrangements between DAFM and
134 Teagasc for oversight and reporting on performance. Internally, a series of revisions of evaluation
135 protocols, framework, methodologies and networks has shaped the evaluation function from 2017 to
136 2021.

137 Teagasc uses agricultural innovation systems as an overarching framework for understanding the
138 organisation's contribution to the transformation of Ireland's AKIS. To evaluate its research
139 programmes and advisory service Teagasc uses ex-post outcome assessment and peer assessments
140 (Table 2, Figure 1). Key principles that guide evaluation activities in Teagasc are:

- 141 1. Using M&E for accountability and learning.
- 142 2. Incorporating stakeholder participation.
- 143 3. Undertaking M&E of economic, social, and environmental impacts.

- 144 4. A focus on improvement of service design and delivery.
- 145 5. Strengthening organisational governance.



146
 147 **Figure 1:** Timeline of key external and organisational events in building evaluation capacity in Teagasc.

148
 149
 150

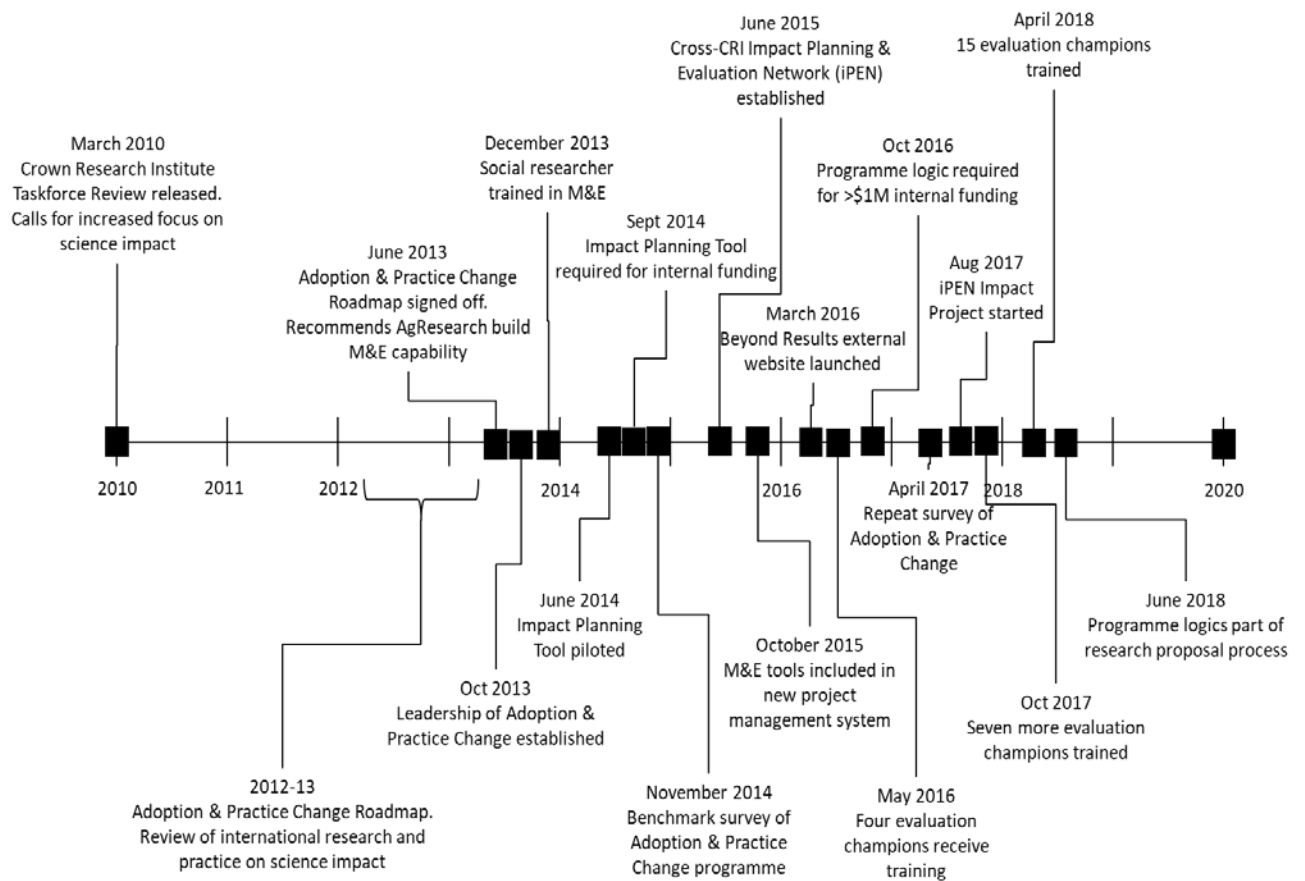
151 **AgResearch Ltd., New Zealand**

152 AgResearch is one of seven Crown Research Institutes (CRI), Government-owned organisations tasked
 153 with providing research and technologies to deliver outcomes for New Zealand. AgResearch focuses
 154 on the pastoral and agri-food sectors, *“to enhance the value, productivity and profitability of New*

155 *Zealand's pastoral, agri-food and agri-technology sector value chains to contribute to economic*
156 *growth and beneficial environmental and social outcomes for New Zealand."* (AgResearch, 2018).

157 AgResearch achieves this by providing research and the transfer of knowledge and technologies in
158 partnership with key Māori (the indigenous peoples of New Zealand), industry and Government
159 partners (AgResearch, 2020). AgResearch has 286 scientists and 178 technicians (2020), but does not
160 have staff dedicated to technology transfer. Instead, AgResearch partners with stakeholders with
161 these capabilities. .

162 In 2010 the CRI Taskforce Review (Crown Research Institute Taskforce, 2010) (Figure 2), recommended
163 changes to encourage realising impact from science. Recommendations included, improve
164 partnerships between research organisations and industry, increase internal funding managed by each
165 CRI to align with sector needs, and increase each CRI's accountability for impact (Turner et al., 2013).
166 To encourage the last each CRI reports on the percentage of relevant end-users who have adopted
167 their knowledge and technology and provides annual impact case-studies (AgResearch, 2020) (Table
168 2).



169

170 **Figure 2:** Timeline of key external and organisational events in building evaluation capacity in
 171 AgResearch.

172 In response to the CRI Taskforce Review recommendations the AgResearch Adoption and Practice
 173 Change Roadmap (2012-13) developed organisational recommendations to increase science impact,
 174 including:

- 175 1. Partnering with stakeholders to deliver research outcomes.
- 176 3. Planning science programmes with stakeholders to identify issues, outcomes and get a clear
 177 understanding of who the stakeholders will be and their roles in achieving impact (Impact
 178 Planning Tool).
- 179 4. Monitoring and evaluating progress towards outcomes and impact within research
 180 programmes (Percy et al., 2015) for accountability to funders and to support learning and

181 adaptation of research activities (Botha et al., 2017), by requiring Impact Planning Tool and
182 programme logics for internal and Government funded programmes.

183 5. Embedding evaluation within the organisation, rather than through external evaluators or a
184 separate evaluation unit, by training internal Evaluation Champions, and making monitoring
185 and evaluation tools available on Beyond Results website¹ and included in a new
186 organisational project management system.

187 ***Instituto Nacional de Investigación Agropecuaria (INIA), Uruguay***

188 The purpose of INIA is to “produce and adapt knowledge and technologies to contribute to the
189 sustainable development of Uruguay and the agricultural sector, considering State policies, social
190 inclusion, and market and consumer demands”. To achieve this, INIA is committed to research, science
191 and technology, and collaborates in the design of public policies for the development of the
192 agricultural sector. Furthermore, through publications and technological developments, it aims at
193 diffusing and transferring scientific knowledge to the agricultural sector and society. INIA’s research
194 activities are particularly focused on livestock, agriculture, dairy, forestry, and plant production from
195 extensive family farming through to intensive production, as well as integrated crop-livestock
196 production systems. INIA has 565 permanent staff, 174 university staff, of which 121 are effective
197 researchers, and 391 support staff. It has five experimental stations distributed throughout the
198 country and a mixed unit with the Pasteur Institute in Montevideo.

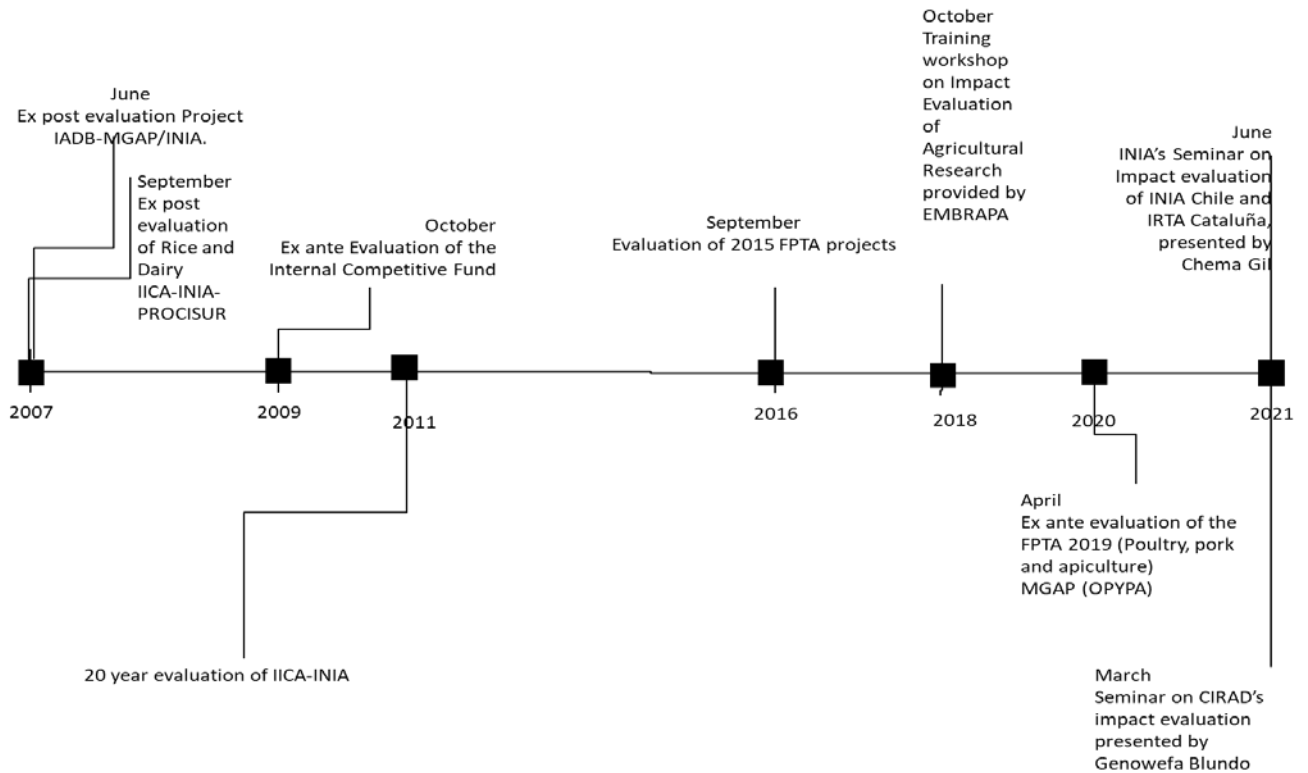
199 INIA has created a system of evaluation certification of technologies with participation of external
200 actors in order to improve the adoption and the impact of the technologies (Vasen et al., 2021).

¹ <https://www.beyondresults.co.nz/>

201 The private sector provides the main source of structural funding through a tax on the sale of
202 agricultural and livestock assets and Government agricultural departments. This amount represents
203 approximately 82% of the funding managed by INIA (Table 1).

204 Research evaluation, and impact evaluation in particular, have been an area of concern, and, as seen
205 in the timeline (Figure 3)², a series of activities have been implemented to advance this. An
206 institutional assessment was carried out in 2011, and it has recently strengthened its Planning,
207 Monitoring and Evaluation capacity by having a staff member dedicated to M&E. Although different
208 innovation perspectives coexist internally, co-innovation is becoming increasingly important. At INIA
209 co-innovation involves the participation of relevant stakeholders in all phases (design,
210 implementation, and evaluation) with the goal of achieving greater adoption and adaptation of
211 technologies, as well as a greater impact.

² The timeline is based on Costa (2022).



212

213 **Figure 3:** Timeline of key external and organisational events in building evaluation capacity in INIA-
 214 Uruguay³.

215 ***Institute for Food and Agricultural Research and Technology (IRTA), Catalonia (Spain)***

216 IRTA is a research institute of the Government of Catalonia (Spain) whose purpose is to contribute to
 217 the modernisation, improvement and promotion of competitiveness and sustainable development in
 218 the agriculture, food, and aquatic sectors, providing safe and quality foods to the final consumer, and

³ IADB-MGAP/INIA - Banco Inter-Americano de Desarrollo. Ministerio de Ganadería, Agricultura y Pesca. Instituto Nacional de Investigación Agropecuaria. IDB-MGAP/INIA - Inter-American Development Bank. Ministry of Livestock, Agriculture and Fisheries. National Institute of Agricultural Research. IICA-INIA-PROCISUR - Instituto Interamericano de Cooperación para la Agricultura. Instituto Nacional de Investigación Agropecuaria. Programa Cooperativo para el Desarrollo Tecnológico Agroalimentario y Agroindustrial del Cono Sur. Inter-American Institute for Cooperation on Agriculture. National Institute of Agricultural Research. Cooperative Program for the Technological Development of Agrofood and Agroindustrial of the Southern Cone. FPTA projects - Proyectos financiados con el Fondo de Promoción de Tecnología Agropecuaria. Projects financed with the Fund for the Promotion of Agricultural Technology. EMBRAPA - Brazilian Agricultural Research Corporation. Empresa brasileña de investigación agropecuaria. MGAP (OPYPA) - Ministerio de Ganadería, Agricultura y Pesca (Oficina de Programación y Política Agropecuaria). Ministry of Livestock, Agriculture and Fisheries (Agricultural Programming and Policy Office).

219 contributing to the global improvement of human welfare. IRTA is the leading organisation for agri-
220 food applied research in Catalonia. Aligned with the Catalan/State/EU strategy, IRTA must address the
221 main challenges of the agri-food sector thus generating new scientific knowledge and transferring it
222 to producers and industry to generate innovation.

223 IRTA's activities are to promote research and technological development in agri-food, facilitate the
224 transfer of scientific findings, and increase its own technological advances while maximising public
225 and private sector coordination. In the spirit of collaboration as the most efficient way to strengthen
226 capacities and improve societal impact, IRTA established and consolidated a network of partnerships
227 with public and private institutions. IRTA is economically supported by its own funds from royalties
228 and other sources, as well as by regional, state, and international funding (Table 1). Research and
229 development activities are conducted in ten research centres and eight experimental stations
230 throughout Catalonia, as well as two other centres in association with universities and other
231 organisations. IRTA employs approximately 700 staff.

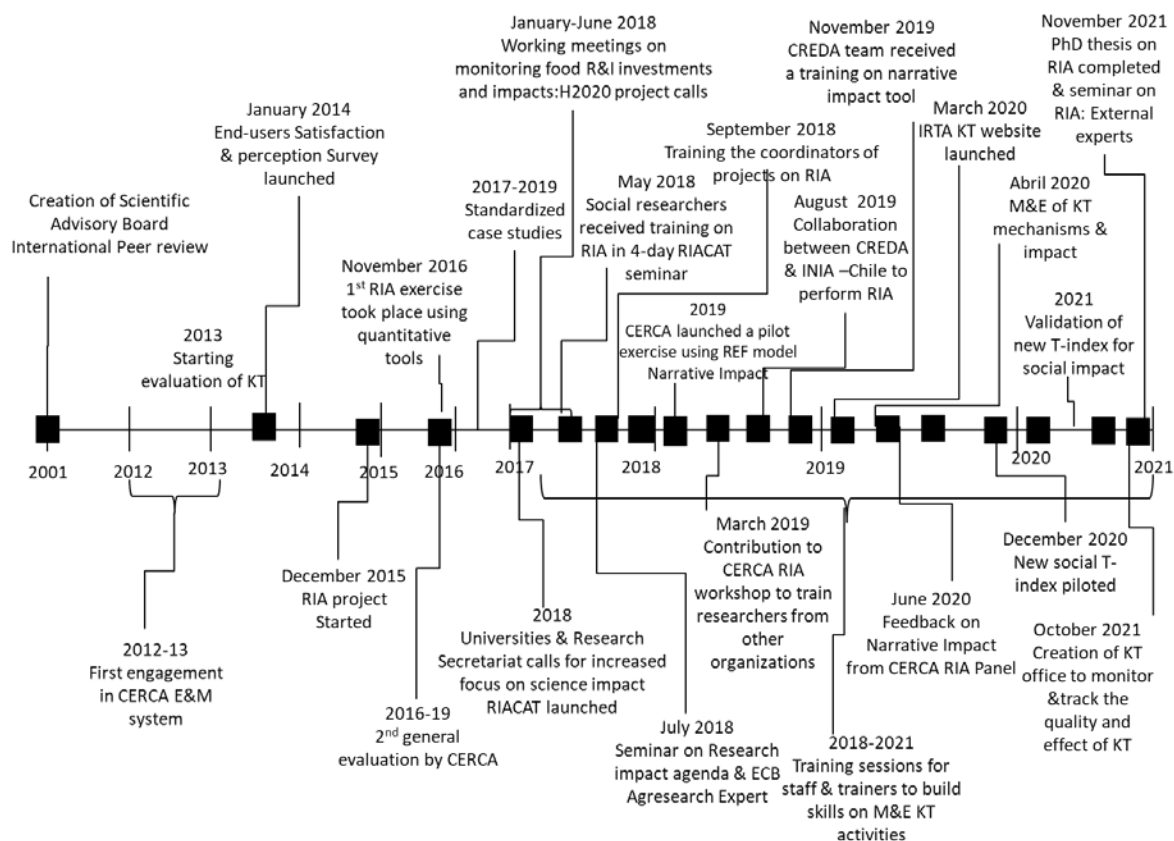
232 IRTA's evaluation of research is based on ex-post economic impact assessment and social returns
233 approaches to consider multiple impacts of the organisation's research (Table 2, Figure 4). The impact
234 on Catalan agricultural productivity of the research conducted by IRTA has recently been analysed for
235 the period 1985-2015. The main conclusions of the study (Guesmi and Gil, 2021) were:

- 236 1. IRTA plays a relevant role in strengthening the system of agricultural technology in Catalonia.
- 237 2. Elasticity of Total Factor Productivity⁴ with respect to the public knowledge stock was
238 approximately 0.15, indicating that a one per cent increase in the public R&D knowledge is
239 likely to lead to a 0.15 per cent increase in agricultural productivity.

⁴ Total factor productivity represents the part of output growth that cannot be explained by the growth of inputs used in the production. It is measured as the ratio of total output to total inputs.

240 3. The social rate of return on investment⁵ in IRTA ranges from 15% to 28%, depending on
 241 assumed lag times and discount rate.

242



243

244 **Figure 4:** Timeline of key external and organisational events in building evaluation capacity in IRTA.

245

246

247

248

249

⁵ The social rate of return could be defined as a percent return on each currency unit spent on R&D investments. It provides an estimate of the benefits from a one-off increase in R&D expenditures, which could be useful as an ex-post measure of the returns achieved and an ex-ante tool to assist in resource allocation (Sheng et al., 2011).

250

251

252

253

254 **Table 2:** Characteristics of the agricultural research impact agenda in each organisation – definition, weighting and incentives for an impact focus

Characteristics	AgResearch	INIA	IRTA	Teagasc
Definition of impact	Improved productivity, profitability and value from pasture-based production while reducing environmental footprint.	Contribute to the sustainable development of Uruguay and the agricultural sector considering State policies, social inclusion, and market and consumer demands, strengthening the linkages between research and the productive sector.	Modernisation, improvement, and promotion of competitiveness and sustainable development in the agriculture, food and aquatic sectors.	Underpin profitability, competitiveness and sustainability of agri-food sector and bioeconomy.
Weighting of impact in research assessment	50% in ex-ante assessment of public research funding to show potential benefits to New Zealand (economic, environmental and social).	Almost all efforts are oriented to research monitoring. Ex-ante evaluation and impact are not yet established practices at INIA.	Depends on the type of project. High weights attributed to research related to business activities.	50% in ex-post outcome assessment.
Funder mechanisms for encouraging an impact focus	Ministry of Business, Innovation and Employment.(MBIE) competitive funds require ex-ante estimate of impact and credible impact pathway in funding proposals. Funding contracts with MBIE competitive funds include an outcome statement against which projects are annually assessed. Impact case studies (ex-post) are required as part of annual reporting on all AgResearch science to MBIE. AgResearch provides a selection of annual case-studies as examples of economic, social, and environmental impact.	INIA's last institutional impact evaluation carried out in 2010, through IICA and a multidisciplinary team. In 2015 INIA's Fund for the Promotion of Agricultural Technology (FPTA) had an impact evaluation process with institutional funding.	Evaluation against outcomes (indicators): adopted 10 years ago, though it is no longer used.	Annual reporting to Department of Agriculture, Food and the Marine on Teagasc's performance in delivering on its goals/objectives as set out in its multi-annual Statement of Strategy and the associated annual High Level Business Plan. Focus on outcome evaluation. Ex-ante impact pathway.

Sources	(AgResearch, 2020; Ministry of Business, 2015, 2017)	www.inia.org.uy/online/site/96788211.php www.iica.int/es/prensa/noticias/iica-evalua-20-anos-de-investigacion-de-inia-uruguay	(IRTA, 2017)	Teagasc (2017)
---------	--	--	--------------	----------------

255

1 The next section presents the study methodology, defining evaluative capacity building and the
2 analytical framework for comparing the types and uses of evaluation, and evaluative capacities. The
3 results of the comparative analysis are then presented and discussed. The paper finishes with key
4 conclusions regarding how public-funded agricultural research organisations can build evaluation
5 capacity.

6 **Methodology**

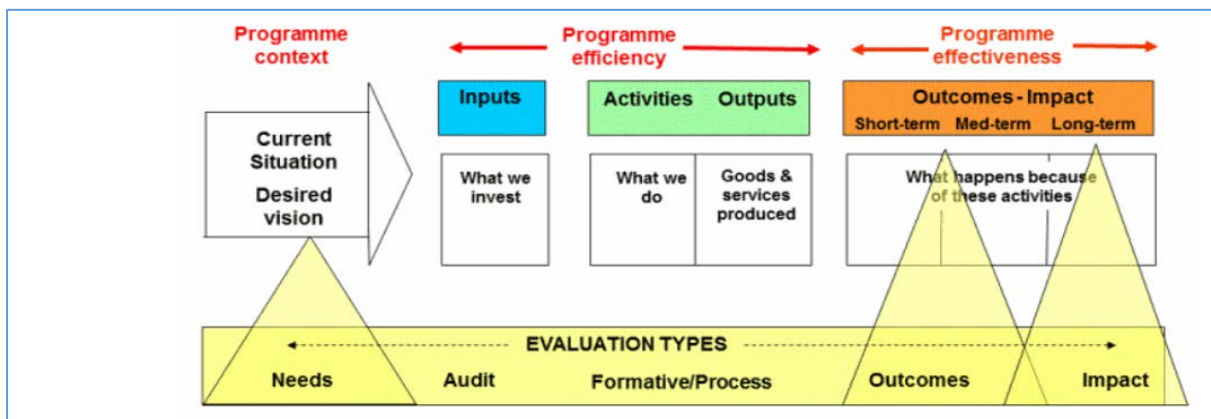
7 The Evaluation Capacity Building (ECB) literature points to the need to increase the capacity within the
8 individual, organisation, and wider science system to do and use evaluation (Bourgeois & Cousins,
9 2013; Chaplowe & Cousins, 2016). However, as every organisation works and operates within different
10 contexts there is no single approach to building evaluative capacity (Bourgeois et al., 2015). White,
11 Percy & Small (2018) suggest using a hybrid model of ECB based on the multi-disciplinary model of
12 Preskill & Boyle (2008) and Cousins, Goh, Elliott & Bourgeois (2014) as a useful foundation for science
13 organisations building an ECB culture. This aligns an evaluative culture of “*doing*” that includes such
14 things as organisational support systems, structure, and leadership; capacity to do evaluation within
15 the organisation; and capacity to do evaluation within programmes, with a culture of “*using and*
16 *valuing*” that involves organisational learning capacity; programme and individual learning capacity;
17 and capacity to use the evaluation by the organisation and within programmes. This hybrid model
18 formed the basis of the analytical framework described below.

19 **Analytical framework**

20 For the purposes of this study, evaluation is defined as the structured interpretation of predicted or
21 actual impacts of proposals, organisations, programmes, projects, or individuals (Taylor-Powell &
22 Henert 2008). It looks at original objectives, and at what is either predicted, or what was accomplished
23 and how it was accomplished. To understand the types of evaluation in each organisation, we
24 estimated the allocation of funding to internal and external evaluations, and the proportion of
25 evaluation undertaken as (Taylor-Powell & Henert 2008; Superu, 2017b):

- 1 1. Needs – Characterisation of the needs and priorities of the population targeted by a
- 2 programme, as well as the potential barriers and enablers to the programme meeting these.
- 3 2. Audit – Evaluation of a programme for the purpose of verifying adherence to a set of pre-
- 4 defined processes.
- 5 3. Formative (Process) – Method of assessing how a programme is being implemented. Process
- 6 evaluation focuses on the programme’s operations, implementation, and service delivery.
- 7 4. Outcome – Focuses on the effectiveness of the programme and its outcomes. Outcomes can
- 8 be produced in the short or long term.
- 9 5. Impact – Assesses programme effectiveness in achieving its ultimate goals. This is done several
- 10 years after the programme is implemented.

11 **Figure 5:** Types of evaluation.



12 Source: Taylor-Powell & Henert (2008)

13
14 To understand why each organisation undertakes evaluation, the frequency of evaluation for each of
15 the four “As” was estimated using a four-point Likert scale (never/occasionally/frequently/always):

- 16 • Allocation – Evaluation is used to instigate changes within a project or programme.
- 17 • Advocacy – Evaluation is used to support a position or to justify action or inaction.
- 18 • Analysis – Evaluation is used to learn about the programme, its staff, its operation, or
- 19 outcome.

- Accountability – The process of evaluation is used to support programme management and engagement (e.g. use of programme logics, Theory of Change, Evaluation Plans etc.).

To consider the readiness of each organisation to implement evaluation, questions adapted from the Checklist for Organisational Readiness for Evaluation (Superu, 2017b) were used (Table 3). For each of the questions in Table 3 readiness was assessed as:

- Emerging: There is very little or no evidence of this in the organisation.
- Developing: there is some evidence of this in the organisation – it happens sometimes but is not embedded in routines.
- Consolidating: This is seen occurring regularly in the organisation, however, not everyone is on board.
- Highly developed: This is evident at all levels of the organisation and appears strongly embedded in the organisation’s governance, leadership, system, structures, and practices.

Table 3: Questions used to consider organisational readiness for evaluation.

Our organisation uses evaluation to be accountable to our funders and/or internal and external stakeholders
Our organisation uses evaluation for learning and reviewing to improve the ways our organisation delivers impact from science and/or investments in science
Our organisation understands the value of demonstrating the difference we have made from science for our stakeholders
Leaders and/or managers are committed to developing evaluation capacity
Systematic M&E occurs at all levels of our organisation (if not at all levels, specify which, e.g. project-level, organisation-level)
Our organisation has the systems and processes in place to collect data for evaluation
Our organisation has the skills and knowledge to collect data for evaluation
In our organisation, we have the skills and knowledge we need to understand and use data required for evaluation
In our organisation, staff have time, funds and/or resources to collect and use data required for evaluation
In our organisation we (decision-makers and staff) have opportunities to involve a wide range of people (internal and external to our organisation) in evaluation
In our organisation, we participate in external networks of learning to increase our knowledge of evaluation

Source: Superu (2017b)

1 ***Data collection and analysis***

2 Two workshops were held on November 2017 and February 2019 at Teagasc (Carlow, Ireland) and
3 IRTA-CREDA (Barcelona, Spain), respectively. At the first workshop there were three participants from
4 Teagasc, two from CREDA-UPC (who provided information on IRTA), one from INIA, and two from
5 AgResearch. At the second workshop there was one participant from Teagasc, three from CREDA-UPC
6 (who again provided information on IRTA), one from AgResearch, one from Scotland, and one from
7 Udelar (University of the Republic) (who participated virtually and provided information on INIA). In
8 the second workshop participants reviewed and updated data collected in the first workshop. All the
9 workshop participants are tasked with undertaking evaluation within their organisations, as well as
10 one participant each from Teagasc, INIA and AgResearch being tasked with building evaluation
11 capacity in their organisations.

12 At the 2017 workshop the participants mapped the reasons evaluation is undertaken, the types of
13 evaluation, approaches and activities, and actions to build evaluation capacity (see Analytical
14 Framework). The selection of questions (Table 3) were circulated prior to the workshop so that
15 workshop participants could consult within their organisation to provide an assessment of
16 organisational readiness.

17 At the workshop each organisation used a colour-coded dot to show where they were positioned
18 (Table 6) and then a facilitated discussion was held around the gaps and trends emerging. After the
19 workshop the participants undertook verification with key staff in their respective organisations.

20 Having identified the types of evaluation and evaluative capacities, workshop participants then
21 identified the initiatives, strategies, actions, and key resources each organisation engaged in to
22 support evaluation at the stages of planning, monitoring and communication.

23 **Findings and Discussion**

24 ***Purposes for evaluating research***

1 The four organisations undertook evaluation for all four purposes, though with different frequency
2 (Table 4). Evaluation for allocation purposes ranged from always in Teagasc, to frequently in INIA, to
3 occasionally in the other organisations. Teagasc has more initiatives in place to assist with being able
4 to make changes within a project or programme, including a formal cyclical peer review system of
5 research programme objectives and evaluation of knowledge transfer services. AgResearch uses its
6 annual programme reporting and funding processes to reflect on programme contributions to the
7 Science Plan and identify changes to allocation of funding to different research areas. IRTA also use
8 evaluation to allocate resources to different research programmes based on ex-ante estimation of the
9 potential impact of research. INIA are looking to increase the use of evaluation for informing funding
10 allocation at the programme-level.

11 Accountability is traditionally a strong driver for undertaking evaluation (Midmore 2017; Penfield et
12 al., 2014), and agricultural research organisations are no different. INIA, Teagasc and AgResearch had
13 strong accountability drivers. In contrast, the remaining organisation identified that while evaluation
14 for accountability was occurring, it was not well embedded throughout the organisation. For example,
15 through its Annual Reports to the New Zealand Government, AgResearch reports on organisational
16 performance against Key Performance Indicators (KPIs), such as commercial revenue, science
17 publications per scientist, and understanding of stakeholder strategies. Another example is INIA which
18 intends to hold a workshop with producer associations and the government, at the end of its five-year
19 strategic plan, to evaluate performance in seven production systems against the objectives and goals
20 proposed at the beginning of the plan. In the case of Teagasc, through its oversight agreement with
21 DAFM, it provides an evaluation of its performance in delivering on its goals/objectives as set out in
22 its multi-annual Statement of Strategy and the associated annual High Level Business Plan for the
23 period under review.

24 Another significant driver for evaluation is learning (or analysis) (Midmore 2017; Pollock 2012; Wade
25 & Kallemeyn, 2020). While this driver was evident within Teagasc, AgResearch, and INIA, IRTA is at the

1 very early stages of realising organisationally the potential for evaluation to improve their science
 2 programmes through evaluation for analysis (Table 4). To support this use of evaluation AgResearch
 3 is building evaluation into all major research programmes; though this is at an early stage (Table 8).

4 **Table 4:** Frequency with which evaluation is used in each research organisation for the four purposes.

Purposes for evaluating research	Never	Occasionally	Frequently	Always
Allocation		AgResearch IRTA	INIA	Teagasc
Advocacy		Teagasc IRTA	AgResearch INIA	
Analysis		IRTA INIA	AgResearch Teagasc	
Accountability		IRTA	AgResearch Teagasc INIA	

5
 6 Since this research was undertaken the four “As” have been extended by Parks et al. (2019) to include
 7 acclaim (comparing and recognising the value of higher education institutions and the research
 8 conducted within them) and adaptation (to steer change in organisational structures, behaviours and
 9 cultures, and research activities and priorities). Future research should consider the evaluative
 10 capacities needed to support evaluation for acclaim and adaptation.

11 ***Internal and external evaluation***

12 The four organisations have different balances of evaluation undertaken by internal and external
 13 parties (Table 5). For IRTA and AgResearch, evaluation is predominantly undertaken internally. In
 14 contrast, evaluation is equally split between external and internal in Teagasc, and in the case of INIA
 15 is predominantly undertaken externally.

1 **Table 5:** Estimated proportion of internal and external evaluation undertaken within each organisation

Organisation	% undertaken internally	% undertaken externally
AgResearch	~ 85	~ 15
INIA	40	60
IRTA	90	10
Teagasc	50	50

2

3 Teagasc and INIA evaluations always involve an external provider in collaboration with internal staff.
4 These organisations also use external peer review panels chosen to represent the different areas of
5 the programme being evaluated. Using external providers and review panels enables evaluations to
6 be completed more rapidly and the evaluation is considered more impartial. This emphasis on external
7 evaluation may be related to Teagasc and INIA both emphasising the use of evaluation for
8 accountability purposes (Table 4).

9 ***Types of evaluation***

10 For most of the organisations, evaluation is concentrated on internal research performance: needs,
11 auditing, and formative evaluation (Table 6 and Figure 5). INIA, IRTA, and AgResearch are all
12 undertaking similar proportions of evaluation in these areas. For example, AgResearch reports on
13 milestones, deliverables, and case studies of research programmes in its Annual Reporting. The
14 emphasis on formative (or process) evaluation, relative to summative (or outcome) evaluation, may
15 be due to impacts occurring post-programme and the difficulty of allowing for the time lag to when
16 impacts occur, and lack of investment in evaluation of impacts post-project.

17 There appears to be less analysis of changes in the social, economic and production context of
18 stakeholders to analyse research effectiveness and changes beyond the research organisations.
19 Evaluation of impact is, therefore, limited in all organisations. However, the organisations are
20 beginning to increase their emphasis on summative evaluation. Teagasc is strategically moving into

- 1 formative and summative evaluation and INIA recently incorporated the evaluation of outcomes and
- 2 the goal is to strengthen the evaluation of impact in the future.

Table 6: Proportion and examples of types of evaluation being done in each organisation.

Types of evaluation	AgResearch	INIA	IRTA	Teagasc
Needs	<p>30%</p> <p>Stakeholder participation in programme logic development; participatory project development with stakeholders.</p>	<p>20%</p> <p>5-year strategic plan defining research areas and technology transfer activities.</p>	<p>25%</p> <p>2020-2023 Strategic Plan.</p>	<p>1%</p> <p>Stakeholder consultation on research programmes and advisory services.</p>
Audit	<p>25%</p> <p>Reporting on programme milestones and deliverables usually to meet requirements of funders.</p> <p>Organisational performance is annually reported on against KPIs for annual reporting</p>	<p>45%</p> <p>Research projects and outputs are annually monitored against organisational priorities and KPIs¹.</p>	<p>30%</p> <p>Individual evaluations of the programmes by external panels (2015).</p>	<p>1%</p> <p>Annual progress reporting on milestones, deliverables and KPIs.</p>
Formative	<p>35%</p> <p>Planning and monitoring of outputs and activities and outcomes within individual projects. Requirement to record in project management system.</p>	<p>20%</p> <p>Mainly implemented in relation to the Fund for the promotion of Agricultural Technology (FPTA) projects. Currently an evaluation strategy is being developed with the Policy Unit of the Ministry of Agriculture.</p>	<p>35%</p> <p>Annual objectives of programmes (scientific, economic & transfer).</p>	<p>50%</p> <p>Part of cyclical review process of research programmes and advisory regions.</p>

¹ KPIs – Key Performance Indicators

Table 6: Proportion and examples of types of evaluation being done in each organisation (continued)

Types of evaluation	AgResearch	INIA	IRTA	Teagasc
Outcomes	<p>5%</p> <p>Programme case studies in AgResearch Annual Report. Some outcome reporting for funders for individual projects.</p>	<p>10%</p> <p>Research products are evaluated against organisational KPIs, and from there to individuals using a cascade-approach ⁶.</p>	<p>5%</p> <p>Being part of the CERCA system in 2018 (The Agency for the Research Centres of Catalonia).</p>	<p>43%</p> <p>Research evaluated against organisational KPIs. Case studies in annual reports and part of cyclical review process.</p>
Impact	<p>5%</p> <p>Programme case studies in AgResearch Annual Report.</p>	<p>5%</p> <p>INIA is currently defining its evaluation strategy, and the impact evaluation approach. In 2015 the FPTA fund and selected individual programmes were evaluated.</p>	<p>5%</p> <p>CREDA's ex-post evaluation of the societal impact of research and innovation (2016-2019).</p>	<p>5%</p> <p>Annual Research Impact Highlights Cyclical review process</p>

⁶ The cascade approach goes from the macro (organisations) to the meso and then the micro-level of individuals, through stakeholders working at each level, e.g. managers and regional directors, and then through the programme directors.

1 **Evaluative Capacity Building Activities**

2 While not an entirely comprehensive list, Table 7 provides an indication of specific initiatives,
3 strategies, actions, and key resources that the four organisations used to implement evaluation
4 capacities in planning, monitoring, and communication (including reporting). In the planning stages,
5 all organisations had strategy documents, and most utilised an advisory board to support them.
6 AgResearch has targeted moving evaluation to the programme-level and indicate several initiatives to
7 assist with this. Monitoring activities were evident around organisational KPIs and annual reporting.
8 At the programme-level, reviews or evaluations were also a key feature in all four organisations. To
9 communicate impact, all four organisations utilised both widespread communication, as well as more
10 focused means. M&E activities that have been easier to implement provide an obvious and immediate
11 value to staff or management. For example, in INIA-Uruguay and IRTA this has been evidencing the
12 economic return on research investment for accountability purposes. In Teagasc this has been M&E
13 as part of cyclical reviews of programmes. In AgResearch this has been programme logics because
14 they have helped research teams to prepare proposals for funding, which is evidenced by the tools
15 page of AgResearch's Beyond Results website being the site's second most visited page.

16 A key evaluative capacity building activity in all four organisations has been establishment of groups
17 or teams (internal in INIA-Uruguay, Teagasc and AgResearch and external for IRTA) with dedicated
18 budgets, time, and staff to undertake M&E (Table 7 and Figures 1 to 4). This has demonstrated the
19 organisational commitment to M&E.

20

21

22

23

1 **Table 7:** Organisational activities to support planning, monitoring, and communicating for impact as of 2017 baseline workshop.

Evaluation stage	AgResearch	INIA	IRTA	Teagasc
Plan	Adoption & Practice Change programme funding Benefits/impact assessment for large science programmes Training (internal and external) to develop capability Advisory Board Participatory programme logics for planning research (outcome focus) Evaluation champions initiative to support planning Processes to support incorporation of evaluation in application process Science Plan/Statement of Core Purpose Co-innovation guidelines and success principles Impact planning tool for new programmes to understand stakeholders Internal community of practice around evaluation champions	Strategic Plan 2017-2020 International Advisory Board	Strategic Plan 2017-2020 International Advisory Board	Teagasc Statement of Strategy Evaluation Unit Strategy Annual Business Plans

2

3

4

1 **Table 7:** Organisational activities to support planning, monitoring, and communicating for impact as of 2017 baseline workshop (continued)

Evaluation stage	AgResearch	INIA	IRTA	Teagasc
Monitor	Evaluation plans for large science programmes Organisational KPIs Ad hoc reviews, surveys, etc. for different parts of the organisation	Project level experience for +20 years Organisation level KPIs (in 2017 used an international board)	Annual reports Impact on society document Three indices: Research, Economic, Transfer	Business Planning mid- and end-year reviews Internal audit Cyclical peer reviews Ad hoc reviews
Communicate	AgResearch website and publications Impact case studies for annual report AgResearch Annual Report/KPI reporting In development – upskilling/training/resources of staff to tell impact story Publications regarding process of evaluation Beyond Results resources available (www.beyondresults.co.nz)	Press conference Report on 20 years of impact evaluation Surveys – about image of INIA in society Surveys – about knowledge application and satisfaction with technology	Invited speakers Publications Mass media Activities Newsletter	Board of Directors (Teagasc Authority) Website publications Action Plans Research impact highlights Project teams

2

3 Good practice examples in the organisations (Table 7) have involved M&E being built into project or
4 programmes as part of cyclical reviews (e.g. Teagasc and IRTA) so that programmes learn from and
5 are guided by M&E. In the case of Teagasc cyclical reviews are an established organisational process.
6 In AgResearch there are examples of projects including M&E at the proposal stage and then revisiting
7 M&E through the life of the project.

8 Another good practice example of activities for building evaluative capacity is using opportunities for
9 internal and external practitioner and academic feedback on M&E practices as part of continuous
10 improvement (Table 7 and Figures 1 to 4). For example, the CREDA team working with IRTA have
11 improved M&E practice through collaboration with other international research centres (e.g. INIA-
12 Chile), feedback from IRTA staff and management, and academic feedback through conferences and
13 peer-review of publications on M&E methodologies. This builds a network, including social scientists,
14 to systematise activities to bridge the gap between academic research and the practice of M&E, which
15 has been highlighted by Joly et al. (2016) as an opportunity to strengthen evaluation capacities in
16 agricultural research organisations. This was also highlighted in Hall's (2018) review of AgResearch in
17 recommending "The need to create a critical mass of thought and practice leadership dedicated to
18 the programme to anchor and frame it and to act as a bridge between theory and practice. This is an
19 issue of adequate resourcing and appropriate skills".

20 **Lessons Learned**

21 ***Leadership is important for organisational readiness for evaluation***

22 The degree to which the organisations have systems and capabilities in place to collect and use
23 evaluation data appears to be positively related to the degree to which the organisations value
24 demonstrating impact and in which leaders are committed to building evaluation capacity (Table 8).
25 For example in a review of AgResearch practices for science impact Hall (2018, p. 16 highlighted "The
26 importance of senior management level support and championing of capacity building for impact
27 programs in research organisation". Teagasc has already begun to consolidate the systematic use of

28 evaluation at all levels of the organisation, including processes for collecting data. Teagasc's
29 recognition of the value of demonstrating impact is also reflected in the organisation being at the
30 consolidating stage of skills, knowledge, and resourcing to collect and use evaluation data. The other
31 organisations are emerging or developing the skills, knowledge, and resources to collect and use
32 evaluation data.

33 Perhaps reflecting that leadership commitment to building evaluation capacity is still developing in
34 AgResearch, and INIA, these organisations are at the stage of developing processes for the systematic
35 use of evaluation but are still emerging in having processes to collect evaluation data. For example, all
36 AgResearch competitively funded programmes are required to develop a programme logic. However,
37 in most cases this has yet to be translated into evaluation plans and collection of evaluation data for
38 funded programmes. Leadership and valuing the demonstration of impact, therefore appear to be
39 important precursors to other aspects of organisational readiness to implement evaluation as well as
40 sustaining evaluative capacity, as has previously been observed by Stone-Jovicich et al. (2019), Wade
41 & Kallemeyn (2020), and White et al. (2018). This is not surprising given the importance of these
42 factors in the success of organisational change initiatives, as highlighted by research on successful
43 organisational change (e.g. Kotter (2012) and Brinkerhoff and Morgan (2010)).

44

45 **Table 8:** Degree of organisational readiness to implement evaluation capacities as of 2018.

Evaluation capacities	Emerging	Developing	Consolidating	Highly developed
Organisation values demonstrating impact		IRTA INIA	AgResearch	Teagasc
Leadership committed to building evaluation capacity	IRTA	AgResearch INIA	Teagasc	
Systematic M&E occurs at all levels of organisation	IRTA	AgResearch INIA	Teagasc	
Systems and processes to <u>collect</u> M&E data	AgResearch IRTA		Teagasc INIA (in projects)	
Skills and knowledge to <u>collect</u> M&E data		AgResearch IRTA	Teagasc INIA	
Skills and knowledge to <u>use</u> M&E data	AgResearch IRTA	INIA	Teagasc	
Sufficient resourcing to <u>collect</u> M&E data	AgResearch IRTA	INIA	Teagasc	
Opportunities to involve stakeholders in M&E		AgResearch INIA IRTA	Teagasc	

46

47 ***Internal and external support networks to build evaluative capacity***

48 All four organisations have benefited from having both internal and external support for M&E
 49 activities (Table 7 and Figures 1 to 4). Internal support included senior management resourcing and
 50 championing the inclusion of M&E in organisational processes. External support included bringing in
 51 M&E experts to facilitate evaluation or train social scientists in evaluation best practice, particularly
 52 through “train the trainers” and “learning by doing”. For example, IRTA has worked closely with a team
 53 of social researchers at CREDA who are experts in research impact assessment.

54 Also beneficial was growing internal and external networks of M&E champions to build legitimacy for
 55 M&E and share good practices. The latter has previously been recommended by Joly et al. (2016) and
 56 our findings highlight the benefits of this for creating an enabling environment for M&E. Internally,
 57 science staff leading large research programmes or with experience in evaluation can be influential
 58 champions. For example, Teagasc has supported evaluators to move into other roles in the

59 organisation to embed evaluative thinking. Three of the organisations (INIA, IRTA and AgResearch)
60 have benefited from joining a network of other research organisations and agencies implementing
61 M&E. This includes bringing in external experts to participate in seminars and workshops on
62 evaluation methodologies to build awareness and evaluation capacity, an approach also used by
63 Teagasc.

64 ***Build evaluation practices into existing planning and review processes***

65 Each organisation had unique reasons for the effectiveness of activities to build evaluative capacity.
66 These reflect the specific characteristics of existing evaluation and science impact activities. Teagasc
67 has embedded evaluation activities into the cyclical reviews of research programmes and knowledge
68 transfer activities that it conducts. These reviews provide an opportunity for staff and stakeholders to
69 reflect on and learn from programmes outside of the annual business planning process. IRTA has been
70 effective in embedding evaluation of impact by having a dedicated knowledge transfer unit that is
71 assessed on its knowledge transfer activities through end-user assessment, technology transfer
72 measures and use of this to inform further improvement. This has included increased resourcing and
73 staff to undertake M&E of knowledge transfer activities. In contrast AgResearch did not have staff
74 dedicated to knowledge transfer nor conducts cyclical reviews. In this context embedding M&E into
75 organisational planning and funding processes, including making M&E tools easily available, has been
76 most effective for increasing M&E activities. These differences in programme and project-level
77 evaluation are like the differences observed in the five agricultural research organisations studied by
78 Joly et al. (2016).

79 A common theme across the organisations though is that M&E activities were built into existing
80 organisational processes (cyclical reviews, knowledge transfer, project planning) and incentivised
81 (KPIs, funding). The value of cyclical reviews by IRTA and Teagasc for incentivising M&E of impact and
82 reflection and learning suggests this could be an opportunity for other research organisations to
83 encourage on-going M&E. Teagasc is seeking to deepen the integration of M&E with regular business

84 planning activities. There is a risk, however, that as was experienced by AgResearch, operationalising
85 M&E means that it can become a ‘tick box’ activity for compliance. Overall however, evaluative
86 capacity building will benefit from having organisational processes for planning and review, as well as
87 incentives for delivering impact, already in place into which M&E activities can then be embedded.

88 ***Different drivers of organisation- and programme-level evaluation***

89 Given different organisational emphasis on evaluation for accountability, advocacy, allocation, or
90 analysis (Table 4), and how these may drive a focus on organisation or programme-level evaluation,
91 none of the organisations would consider themselves highly developed in systematically embedding
92 evaluation at all levels from programme to whole of organisation (Table 8). Out of the four
93 organisations in this study, INIA, and IRTA appear to emphasise accountability and advocacy, which
94 focuses evaluation on organisation-level processes and reporting. For example, strategic plans inform
95 KPIs used in reporting on organisational performance to funders (Table 6). AgResearch appears to
96 emphasise allocation purposes, and a focus on programme-level evaluation, through cyclical peer
97 reviews of individual programmes or areas of research by advisory boards. Teagasc appears to be the
98 most advanced in undertaking evaluation at organisation and programme-levels for accountability,
99 advocacy, allocation, and analysis.

100 This suggests there is an opportunity to develop organisation and programme-level evaluation
101 processes that inform each other. Organisation-level evaluation sets longer-term outcomes
102 (accountability and advocacy) to which multiple programmes deliver, and then supports reporting on
103 the contribution of a suite of programmes to these outcomes. Programme-level evaluation therefore
104 assists with the allocation of funds to the mix of programmes to realise outcomes.

105 There is less use of evaluation for analysis (or learning) to better understand the processes that could
106 lead to greater impact from research. However, since this research was undertaken, Teagasc has
107 begun to address this gap by incorporating a theory-led approach into its research evaluations. This
108 role of evaluation in supporting increased research impact has recently been emphasised in other

109 agricultural research organisations such as CIRAD (Blundo-Canto et al., 2019), the International
110 Livestock Research Institute (Kristjanson et al., 2009), Wageningen University, and the Dutch
111 Agricultural Research Institutes (Spiertz & Kopff, 2011). In the case of the four research organisations
112 studied here, the emphasis remained on using evaluation for improving the efficiency and allocation
113 of funding within the organisations and with sporadic participation of stakeholders (Table 8).
114 Evaluation provides the potential for organisations to work differently with stakeholders to realise
115 impact by keeping them more engaged in programmes, their progress, and impact (Blundo-Canto et
116 al., 2019; Percy et al., 2015; Stone-Jovicich et al. 2019). This practice ranges from emerging (INIA)
117 through to highly developed (Teagasc). For example, Teagasc uses stakeholder engagement both in
118 programme development and evaluation. These programme logics and evaluation plans can then be
119 used to assess how a project is progressing toward desired outcomes, provide accountability for public
120 investment in research, and produce evidence of the benefits of this investment.

121 ***Evaluation capacities for learning remain limited compared with capacities for accountability***

122 M&E for learning remains under-resourced relative to evaluation for accountability and the
123 organisations lack the necessary structures to embed M&E for learning (see also Joly et al. (2016)).
124 The four organisations have therefore found it challenging to implement M&E to understand and learn
125 from the pathways by which science contributes to wider societal impact. This challenge has
126 previously been highlighted by Joly et al. (2016) in a review of M&E practices in five agricultural
127 research organisations. In the current study, two reasons for this challenge were identified.

128 Firstly, M&E for learning is more complex as it needs to reflect the multitude of contexts, actors,
129 interactions, impact pathways and outcomes of individual projects. This places greater demands on
130 organisational resources and capacities, for example to undertake case studies, support the
131 information systems to collect and track a wider variety of data at programme, project, and individual
132 levels, and to standardise findings across case studies to report on organisational-level impact (see

133 also Joly et al. (2016)). It also makes analysis of attribution of individual organisations challenging, if
134 not impossible.

135 Secondly, it was highlighted that strong cultural obstacles to embedding learning from M&E remain.
136 For example, a senior manager in one of the organisations highlighted that “one of the major
137 difficulties is ... still deeply rooted belief that science already fulfils its objectives by generating
138 knowledge and that impact is not a priority or even it is assumed that it will happen if the science is
139 good, a fact that rarely happens.” This may reflect a deeply rooted belief in the separation of scientific
140 and societal values and linear model of knowledge production (Smit & Hessels, 2021). That is, that
141 science is focused on knowledge generation and therefore is not immediately about impact or that
142 impact is something that will automatically happen if research meets science excellence criteria. This
143 culture is emphasised by the tendency for the organisations to focus on traditional measures of
144 scientific excellence (e.g. h-index, citations), stakeholder satisfaction surveys, economic impacts
145 evaluated at the end of a project, or include evaluation only at the planning stage. Thus, there has
146 been limited on-going monitoring and use of that data for reflection and learning, with these activities
147 seen as a “nice to do” that compete with science activities and funding. As Smit and Hessels (2021)
148 highlight in a review of ten evaluation methodologies, the culture influences the types of evaluation
149 methods prioritised by research organisations.

150 There is an opportunity to build capacity amongst in-house M&E teams to conduct M&E for learning
151 by deeper engagement with the academic research and researchers on this type of evaluation. To
152 then begin to reshape organisational culture, evaluation methods that are grounded in co-production
153 and integrated scientific and societal values should be emphasised, e.g. contribution mapping and
154 evaluative inquiry (Smit & Hessels, 2021). Teagasc has begun to address on-going evaluation for
155 learning by introducing three high-level impact pathways to describe the ways Teagasc contributes to
156 change in the agri-food sector. AgResearch also identified the need for clear narratives of the

157 pathways to impact from its research (Hall, 2018), and has found that the process of project teams
158 developing programme logics raised awareness of the existence of different impact pathways.

159 ***Evaluation capacity building increases the frequency of evaluation in agricultural research***
160 ***organisations***

161 For evaluation to thrive within an organisation, it must be valued (Stone-Jovicich et al., 2019) and
162 enabled (Preskill and Boyle, 2008; Wade & Kallemeyn, 2020; White et al., 2018). Figure 6 shows some
163 consistent trends regarding the four organisations' positioning in this regard. At the time of the study,
164 IRTA was in an emerging space whereby the skills, knowledge, resources, and opportunities were yet
165 to be developed; however, it is developing this area. Teagasc had some of the key enablers to collect
166 evaluation data (systems, processes, skills, knowledge) within the organisation, but not yet
167 systemically distributed from the programme to organisation-levels. AgResearch and INIA were
168 positioned in between, considering themselves to be in an emerging to developing position.
169 Organisational evaluation capacity readiness is also reflected in the commitment of leadership. Within
170 AgResearch, INIA and Teagasc, organisational leaders and managers have shown commitment to
171 developing evaluation capacity. Leadership within IRTA has yet to mandate further internal
172 development of systems and processes in place to collect evaluation data and capacity for
173 evaluation. **Figure 6:** Organisations' positioning towards Evaluation Capacity Building.



174

175 While the four organisations included M&E for learning in guiding principles, all were predominantly
 176 focused on M&E for accountability and advocacy. This was reflected in a disconnect between
 177 evaluation at the programme- and organisational-level, with an emphasis on evaluation of
 178 organisational KPIs. For example, IRTA has focused on embedding research impact assessment. This
 179 was prioritised for accountability and reporting at specified periods. M&E for learning at the
 180 programme-level appeared to be left to individual programmes and received less resourcing and
 181 capacity within the organisations, and where it did occur was a point in time event, rather than a
 182 continuous activity. The experience of Teagasc in embedding M&E as part of cyclical programme
 183 reviews could be an opportunity to address the programme- and organisation-level disconnect by
 184 evaluating suites of programmes and projects delivering to a common objective (e.g. Sustainability
 185 Development Goals) and involving project and programme teams in reflecting on these evaluations.

186 **Conclusion**

187 To the best of our knowledge, the present study is one of the first attempts to provide empirical
 188 evidence of how publicly funded agricultural research organisations put research evaluation principles

189 into practice. Drawing on our analysis, accountability and allocation are strong drivers of evaluation
190 in the four agricultural research organisations studied. For most of the organisations, evaluation is
191 concentrated on internal research performance, with less focus on evaluation of societal,
192 environmental, and economic impacts. However, perhaps reflecting the influence of the impact
193 agenda, the organisations are beginning to increase their emphasis on summative evaluation. This is
194 reflected in organisational leaders valuing the demonstration of impact and commitment to building
195 evaluation capacity. These appear to be important precursors to other aspects of organisational
196 readiness to implement evaluation.

197 Given different organisational emphasis on evaluation for accountability, advocacy, allocation, or
198 analysis, and how these may drive a focus on organisation or programme-level evaluation, none of
199 the organisations would consider themselves highly developed in systematically embedding
200 evaluation processes and capabilities at all levels of the organisation from programme to whole of
201 organisation. This suggests there is an opportunity to simultaneously develop organisation- and
202 programme-level evaluation processes that inform each other. There is also less use of evaluation by
203 the organisations to better understand the processes that could lead to greater impact from research.

204 The four organisations have demonstrated their ability to develop evaluation capacities and to
205 establish cycles of periodical impact assessment within their organisations. Given the difference in
206 development across organisations, we believe that more on-going collaboration between the four
207 organisations would provide the opportunity to learn constantly through exchanging information and
208 insights of ‘what works’ in other research organisations and create a network to improve the
209 evaluation capacity, which in turn helps sustain future science impact. Finally, our empirical findings
210 reveal that the research organisations’ challenge is mainly to adapt evaluation practices in terms of
211 skills, knowledge, time, data, resourcing, and supportive structure to build an ECB culture within their
212 own context.

213 **Acknowledgements**

214 Thank you to Teagasc and IRTA for hosting the workshops, and to Penny Payne and Roxanne Henwood
215 for reviewing an early draft of the manuscript.

216 **Funding:** This work was supported by the Strategic Partnership; Strategic Science Investment Fund,
217 New Zealand; Scottish Government; Teagasc; Instituto Nacional de Investigación Agropecuaria (INIA)
218 in Uruguay; and the Societal Impact of R&D Investments [IRTA project-61095].

219

220 **References**

- 221 AgResearch (2020). *Annual Report 2020*. AgResearch Ltd., Lincoln, New Zealand.
- 222 AgResearch (2018). *AgResearch: Statement of Core Purpose*. AgResearch Ltd, p. 2.
- 223 Blundo-Canto, G., Triomphe, B., Faure, G., Barret, D., De Romemont, A., & Hainzelin, E. (2019). Building
224 a culture of impact in an international agricultural research organization: Process and reflective
225 learning. *Research Evaluation*, 28(2), 136-144.
- 226 Botha, N., Coutts, J., Turner, J.A., White, T., & Williams, T. (2017). Evaluating for learning and
227 accountability in system innovation: Incorporating reflexivity in a logical framework. *Outlook on
228 Agriculture*, 46, 154-160.
- 229 Bourgeois, I., & Cousins, J. B. (2013). Understanding dimensions of organizational evaluation capacity.
230 *American Journal of Evaluation*, 35(4), 579-593.
- 231 Bourgeois, I., Whynot, J., & Thériault, É. (2015). Application of an organizational evaluation capacity
232 self-assessment instrument to different organizations: Similarities and lessons learned. *Evaluation and
233 Program Planning*, 50, 47-55.
- 234 Bozeman, B., & Sarewitz, D. (2011). Public value mapping and science policy evaluation. *Minerva* 49,
235 1-23.
- 236 Brinkerhoff, D.W., & Morgan, P.J. (2010). Capacity and capacity development: Coping with complexity.
237 *Public Administration and Development: The International Journal of Management Research and
238 Practice*, 30(1), pp.2-10.
- 239 Chaplowe, S. & Cousins, B. (2016). *Monitoring and evaluation training: A systemic approach*.
240 California: Sage Publications Inc.
- 241 Costa, M. (2022). *Experiencia sobre evaluación en INIA Uruguay* INIA, Montevideo.
- 242 Cousins, J.B., Goh, S.C., Elliott, C., Aubry, T., & Gilbert, N. (2014). Government and voluntary sector
243 differences in organizational capacity to do and use evaluation. *Evaluation and Program Planning*, 44,
244 1-13.
- 245 Cousins, J., Goh, S. C., Elliott, C. J., & Bourgeois, I. (2014). Framing the capacity to do and use
246 evaluation. *New Directions for Evaluation*, 141, 7-23.
- 247 Crown Research Institute Taskforce (2010). How to Enhance the Value of New Zealand's Investment
248 in Crown Research Institutes. *Report of the Crown Research Institute Taskforce*. Wellington: Ministry
249 of Business, Innovation and Employment.
- 250 CSIRO (2015). *Impact Evaluation Guide*. Canberra: CSIRO.
- 251 Donovan, C. (2011). State of the art in assessing research impact: introduction to a special issue.
252 *Research Evaluation*, 20, 175-179.
- 253 Douthwaite, B., & Hoffecker, E. (2017). Towards a complexity-aware theory of change for participatory
254 research programs working within agricultural innovation systems. *Agricultural Systems*, 155, 88-102.
- 255 Gaunand, A., Hocdé, A., Lemarié, S., Matt, M., & Turckheim, E.D., 2015. How does public agricultural
256 research impact society? A characterization of various patterns. *Research Policy*, 44, 849-861.
- 257 Guesmi, B., & Gil, J.M. (2021). The impact of public R&D investments on agricultural productivity.
258 *Review of Economics and Finance*, 19(1), 284-291.

- 259 Hall, A. (2018). Forward Looking Review of AgResearch's Adoption and Practice Change Program.
260 Report to AgResearch Ltd. AgResearch Ltd, Lincoln, New Zealand
- 261 Horton, D., & Mackay, R. (2003). Using evaluation to enhance institutional learning and change: Recent
262 experiences with agricultural research and development. *Agricultural Systems*, 78, 127-142.
- 263 Hueftle Stockdill, S., Baizerman, M., & Compton, D.W. (2002). Toward a definition of the ECB process:
264 A conversation with the ECB literature. *New Directions for Evaluation*, 2002, 7-26.
- 265 Joly, P.-B., Gaunand, A., Colinet, L., Larédo, P., Lemarié, S., & Matt, M. (2015). ASIRPA: A
266 comprehensive theory-based approach to assessing the societal impacts of a research organization.
267 *Research Evaluation*, 24(4), 440-453.
- 268 Joly, P.-B., Colinet, L., Gaunand, A., Lemarié, S., & Matt, M. (2016). "Agricultural research impact
269 assessment: Issues, methods and challenges", OECD Food, Agriculture and Fisheries Papers, No. 98,
270 OECD Publishing, Paris. <http://dx.doi.org/10.1787/5339e165-en>
- 271 Kelley, T., & Gregersen, H. (2003). NRM impact assessment in the CGIAR: meeting the challenge and
272 implications for ICRISAT, Methods for assessing the impacts of natural resource management
273 research. *A summary of the proceedings of an International Workshop, 6-7 Dec 2002*, International
274 Crops Research Institute for the Semi-Arid Tropics (ICRISAT), Patancheru 502 324, India.
- 275 Kelley, T., Ryan, J., & Gregersen, H. (2008). Enhancing ex post impact assessment of agricultural
276 research: the CGIAR experience. *Research Evaluation*, 17, 201-212.
- 277 Kotter, J.P. (2012). *Leading Change*. Brighton: Harvard Business Press.
- 278 Kristjanson, P., Reid, R.S., Dickson, N., Clark, W.C., Romney, D., Puskur, R., MacMillan, S., & Grace, D.,
279 2009. Linking international agricultural research knowledge with action for sustainable development.
280 *Proceedings of the National Academy of Sciences*, 9, 5047-5052.
- 281 Mackay, R. & Horton, D. (2003). Expanding the use of impact assessment and evaluation in agricultural
282 research and development. *Agricultural Systems*, 78, 143-165.
- 283 Maredia, M.K., Shankar, B., Kelley, T.G., & Stevenson, J.R. (2014). Impact assessment of agricultural
284 research, institutional innovation, and technology adoption: Introduction to the special section. *Food*
285 *Policy*, 44, 214-217.
- 286 Mayne, J., & Johnson, N. (2015). Using theories of change in the Agriculture for Nutrition and Health
287 CGIAR research program. *Evaluation*, 21(4), 407-428.
- 288 Menon, S., Karl, J., & Wignaraja, K. (2009). *Handbook on planning, monitoring and evaluating for*
289 *development results*. UNDP Evaluation Office, New York, NY.
- 290 Midmore, P. (2017). The science of impact and the impact of agricultural science. *Journal of*
291 *Agricultural Economics*, 68, 611-631.
- 292 Ministry of Business, Innovation and Employment (2015). *National Statement of Science Investment:*
293 *2015-2025*. Wellington: Ministry of Business, Innovation and Employment.
- 294 Ministry of Business, Innovation and Employment (2017). *The Impact of Science, Discussion Paper*.
295 Wellington: Ministry of Business, Innovation and Employment.
- 296 Morgan Jones, M., Manville, C. & Chataway, J. (2017). Learning from the UK's research impact
297 assessment exercise: a case study of a retrospective impact assessment exercise and questions for the
298 future. *The Journal of Technology Transfer* 10, 1-25

- 299 Organisation for Economic Co-operation and Development (2010). *Performance-based funding for*
300 *public research in tertiary education institutions: Workshop proceedings*. Paris: OECD Publishing.
- 301 Owen, R., Macnaghten, P., & Stilgoe, J. (2012). Responsible research and innovation: From science in
302 society to science for society, with society. *Science and Public Policy*, 39, 751-760.
- 303 Parks, S., Rodriguez-Rincon, D., Parkinson, S., & Manville, C. (2019). The changing research landscape
304 and reflections on national research assessment in the future. *Research England, 2019*.
305 https://www.rand.org/pubs/research_reports/RR3200.html.
- 306 Penfield, T., Baker, M.J., Scoble, R., & Wykes, M.C. (2014). Assessment, evaluations, and definitions of
307 research impact: A review. *Research Evaluation*, 23, 21-32.
- 308 Percy, H., Turner, J.A., Botha, N., White, T., Small, B., Fear, A., & Cooper, A. (2015). Beyond results:
309 Planning science for meaningful change. *Rural Extension and Innovation Systems Journal*, 11, 173.
- 310 Pollock, C. (2012). Repairing a fractured pipeline: Improving the effectiveness of agricultural R&D in
311 the UK. *International Journal of Agricultural Management*, 2, 1-4.
- 312 Preskill, H., & Boyle, S. (2008). A multidisciplinary model of evaluation capacity building. *American*
313 *Journal of Evaluation*, 29, 443-459.
- 314 Sheng, Y., Gray, E.M., & Mullen, J.D. (2011). Public investment in R&D; and extension and productivity
315 in Australian broadacre agriculture (No. 422-2016-26880). Australian Agricultural and Resource
316 Economics Society (AARES), 9–11 February 2011, Melbourne, Victoria
- 317 Smit, J. P., & Hessels, L. K. (2021). The production of scientific and societal value in research evaluation:
318 a review of societal impact assessment methods. *Research Evaluation*, 30(3), 323-335.
319 doi:10.1093/reseval/rvab002
- 320 Spiertz, J.H.J., & Kropff, M.J. (2011). Adaptation of knowledge systems to changes in agriculture and
321 society: The case of the Netherlands. *NJAS - Wageningen Journal of Life Sciences*, 58, 1-10.
- 322 Stone-Jovicich, S., Percy, H., McMillan, L., Turner, J., Chen, L., & White, T. (2019). Evaluating
323 monitoring, evaluation and learning initiatives in the New Zealand and Australisan agricultural
324 research and innovation systems: The MEL² framework. *Evaluation Journal of Australasia*, 19(1), 8 -
325 21.
- 326 Social Policy Evaluation and Research Unit (Superu) (2017a). *Making sense of evaluation: A handbook*
327 *for everyone*. Wellington, New Zealand. Available at: <https://dpmc.govt.nz/sites/default/files/2018-03/Evaluation%20Handbook%20Dec%202017.pdf> (accessed 6 August 2021).
- 329 Social Policy Evaluation and Research Unit (Superu) (2017b). *Getting Your Organisation Ready to Do*
330 *Evaluations: Assessment Guide and Templates*. Wellington, New Zealand. Available at:
331 <https://thehub.swa.govt.nz/assets/documents/Toolkit-print-FINAL-03-17.pdf> (accessed 6 August
332 2021).
- 333 Taylor-Powell, E., & Henert, E. (2008). Developing a logic model: Teaching and training guide. *Benefits*,
334 3, 22.
- 335 Teagasc (2017). Statement of Strategy 2017 -2020. *Supporting Science-based Innovation in Agriculture*
336 *and Food by driving sustainable profit from productivity*.
337 <https://www.teagasc.ie/media/website/publications/2017/Statement-of-Strategy-2017-2020.pdf>
- 338 Tilman, D., Cassman, K.G., Matson, P.A., Naylor, R., & Polasky, S. (2002). Agricultural sustainability and
339 intensive production practices. *Nature*, 418, 671.

340 Turner, J.A., Klerkx, L., Rijswijk, K., Williams, T., & Barnard, T. (2016). Systemic problems affecting co-
341 innovation in the New Zealand Agricultural Innovation System: Identification of blocking mechanisms
342 and underlying institutional logics. *NJAS - Wageningen Journal of Life Sciences*, 76, 99-112.

343 Turner, J.A., Rijswijk, K., Williams, T., Barnard, T., & Klerkx, L. (2013). Challenges to effective interaction
344 in the New Zealand agricultural research and extension system: An innovation systems analysis.
345 *Extension Farm Systems Journal*, 9, 89-98.

346 UK Research and Innovation (2022). Case for support – impact funding.
347 [https://www.ukri.org/councils/stfc/guidance-for-applicants/what-to-include-in-your-proposal/case-](https://www.ukri.org/councils/stfc/guidance-for-applicants/what-to-include-in-your-proposal/case-for-support-impact-funding/)
348 [for-support-impact-funding/](https://www.ukri.org/councils/stfc/guidance-for-applicants/what-to-include-in-your-proposal/case-for-support-impact-funding/) (Accessed 24th February, 2022)

349 Vasen F, Sierra M, Paruelo JM, Negro C, Nolla F, Lapetina J, Salvagno M. (2021). Evaluation of
350 technical production in agricultural sciences: a new certification scheme in Uruguay. *Agrociencia*
351 Uruguay [Internet]. 2021 25(2):e491. Available from: [http://agrocienciauruguay.](http://agrocienciauruguay.uy/ojs/index.php/agrociencia/article/view/491)
352 [uy/ojs/index.php/agrociencia/article/view/491](http://agrocienciauruguay.uy/ojs/index.php/agrociencia/article/view/491).

353 Wade, J., & Kallemeyn, L. (2020). Evaluation capacity building (ECB) interventions and the
354 development of sustainable evaluation practice: An exploratory study. *Evaluation and Program*
355 *Planning*, 79, 1-9.

356 Watts, J., Horton, D., Douthwaite, B., La Rovere, R., Thiele, G., Prasad, S., & Staver, C. (2008).
357 Transforming impact assessment: beginning the quiet revolution of institutional learning and change.
358 *Experimental Agriculture* 44, 21-35.

359 White, T., Percy, H., & Small, B. (2018). Creating an evaluation culture through capacity building: A
360 new frontier in a science organisation. *Evaluation Matters – He Take Tō Te Aromatawai*, 4, 111-135.

361