

Creating an evaluation culture through capacity building: A new frontier in a science organisation

Toni White, Helen Percy, and Bruce Small

Evaluative capacity building is a critical element of weaving the action of evaluation, evaluative activity, evaluative thinking, and appreciation of evaluation into the fabric of organisations. AgResearch, the Crown Research Institute for the New Zealand pastoral farming sector, is embarking on the evaluative capacity building journey following an internal review identifying the need for evaluation to enable better measuring and monitoring of impact from science research programmes.

Evaluation has not traditionally been a core component of New Zealand science programmes. However, with funders and stakeholders requiring science organisations to demonstrate impact and outcomes from research, this situation is changing. The newness of an evaluation culture, whereby monitoring and evaluation are built into science research programmes for learning purposes as much as accountability, to many science organisations within New Zealand means that this is an era for development, learning, and investment in evaluative activity.

This article focuses on three key realms of exploration: a) Who and what creates demand for evaluation in science? b) How can science organisations, such as AgResearch, create the capability to engage in and use evaluation? c) What are the challenges and opportunities for science organisations and science teams as they engage in the new practices of programme monitoring and evaluation?

Key findings include the importance of organisational readiness to facilitate evaluation access, the need to incentivise and encourage scientists to engage in evaluation, and the need to build evaluative capacity throughout the organisation and its programmes. The article also highlights the need to identify and work towards addressing the multiple challenges that project teams face when first introduced to the field of evaluation.

Introduction

Evaluation is a young field in New Zealand in comparison to many other Western countries (Duignan, 2002). Despite an increase in evaluation in many sectors over the past two decades, it has been suggested that, in general, the field of New Zealand research and science still has a considerable way to go before achieving a robust evaluation culture (New Zealand Government, 2015). The biophysical science sector has not been devoid of evaluation, yet it would be difficult to find evidence that evaluation is common-place. Consequently, Preskill's reference to the occurrence of "a social epidemic of evaluation" in her 2008 presidential speech (as cited in Ensminger, Kallemeyn, Rempert, Wade, & Polanin, 2015, p. 129) is certainly not yet true for biophysical science in New Zealand. The relative scarcity of evaluative processes and thinking within the New Zealand biophysical science sector provided an opportunity for AgResearch to initiate a process of evaluative capacity building (ECB).

The documented experiences of international agricultural science, or science-related, organisations and programmes embedding evaluation provide some optimism for New Zealand science organisations. These include the experiences of the Consultative Group on International Agricultural Research (CGIAR) who have evolved a long history of impact assessment with diverse stakeholders (Kelley, Ryan, & Gregersen, 2008), the European Initiative for Agricultural Research for Development (EIARD) which focuses on maximising the impact of agricultural research in developing countries (EIARD, 2003), and environmental organisations such as Environment Canada (Brown, 2011) among others (Clark, Mitchell, & Cash, 2006). The Department for International Development in the UK (DFID), for example, has devised an evaluation strategy (DFID, 2014a) which sets out their direction and mandate for embedding evaluation into the organisation by increasing staff capability, creating standards for measuring against, and undertaking rigorous impact evaluations (DFID, 2014b). Such accounts also highlight the particular challenges faced by science in general, and agricultural science in particular, to achieve impact in a demanding environment with increasingly tight funding conditions (Midmore, 2017; Norton, 2015). Challenges include the complexity of issues that agricultural science seeks to address. For example: environmental sustainability; the fact that the impact of science is closely linked to market success (or failure) within the agricultural sector; considerable time lags that occur between the development of new innovations and adoption by end-users such as farmers; and the diverse actor network system within which science organisations are situated which makes it difficult to attribute impact to individual actors or interventions (Ekboir, 2003; Midmore, 2017; Norton, 2015).

Further, by delving into the fundamental theoretical backgrounds of underpinning so-called “hard” science, positivism often

values impact through peer-reviewed publications and output counts as some of the main validations of knowledge. However, participatory research leans towards constructivism whereby validation of knowledge comes through its use and application (Douthwaite, Kuby, van de Fliert, & Schulz, 2003). These factors all point to the need to develop evaluation capacity and capability of best fit within science-focused organisations and programmes (Patton, 2012).

AgResearch is a government-owned Crown Research Institute (CRI) with a focus on research for New Zealand's pastoral farming sector. AgResearch employs about 685 staff across four New Zealand campuses and has over 700 active science projects each year (AgResearch, 2016a). This article draws on ECB literature to address three focus questions:

1. Who and what creates demand for evaluation in science?
2. How can science organisations such as AgResearch create the capability to engage in and use evaluation?
3. What are the challenges and opportunities for science organisations and science teams as they engage in the new practices of programme monitoring and evaluation?

This article outlines how AgResearch addressed these questions and initiated a process of embedding ECB into the organisation. The article is organised into the following sections to mirror these questions: Drivers for creating evaluation demand; Creating and embedding evaluative capacity; and Challenges to ECB. We conclude with the insights and lessons learnt that could be extrapolated into the wider field of ECB practice.

Drivers for creating evaluation demand

The literature points to a number of external and internal drivers that motivate organisations to consider adopting evaluation and building evaluative cultures. These drivers include:

- securing funding or resources (Bakken, Núñez, & Couture, 2014; Labin, Duffy, Meyers, Wandersman, & Lesesne, 2012)
- showing effective management (Cousins, Goh, Elliott, Aubry, & Gilbert, 2014)
- being accountable for commitments and resources (Bakken et al., 2014; Bourgeois & Cousins, 2013; Cousins, Bourgeois, & Associates, 2014)
- transparency of process (Labin et al., 2012)
- proving adaption to a constantly changing environment (Bakken et al., 2014)
- the organisation's values or an overall desire to improve and learn (Cousins, Bourgeois et al., 2014).

The diversity of drivers contributing to ECB indicates that these will influence organisations differently depending on the context in which they are operating (Bourgeois, Whynot, & Thériault, 2015). For any organisation there needs to be a clear value proposition that motivates staff to view evaluation as an opportunity rather than an obligation (Leviton, 2014; Rennekamp & Engle, 2008) and shows evaluation as a benefit rather than a cost. The New Zealand science community is no exception to this.

Cousins, Goh, Elliott, Aubry et al. (2014) identify that governments have long been a driver of evaluation. In New Zealand the government is a key investor in science and there is a clear message emerging that government expects recipients of science funding to prove the value and societal impact from that investment. For

example, AgResearch and the other CRIs must report on performance against a Statement of Core Intent; a statement outlining the organisational strategy to deliver to its core purpose (AgResearch, 2016b). External drivers of accountability and transparency for funders remain a key justification for evaluation. In addition, and in response to the external drivers for an increased focus on impact from research, AgResearch developed the Adoption and Practice Change Roadmap (AgResearch, 2013). The Roadmap provided the strategic direction for an increased focus on impact within the organisation, of which evaluation of that impact is a key component.

One of the six recommendations from the Roadmap was that AgResearch would build capability to undertake monitoring and evaluation of projects for all of its key impact areas (AgResearch, 2013). AgResearch had previously conducted a few ad hoc evaluations at the project level and in order to contribute to annual reporting requirements (for example, AgResearch, 2015; White & Sheath, 2009). However, this strategic approach was new and required some thought on how AgResearch might build evaluative capacity into a science organisation with limited experience in evaluation. As part of implementing the Adoption and Practice Change Roadmap (from 2013 onwards), AgResearch identified what evaluative capability was required within the organisation, and initiated steps towards the development of an evaluative culture.

Creating ECB in a science organisation

There is no recipe for building evaluative capacity because every organisation works and operates within different contexts (Bourgeois et al., 2015). Another complication is the diverse definitions within the literature, making it difficult to compare one ECB construct to another (Naccarella et al., 2007). A definition fit

for the purposes of a science organisation is provided by Stockdill, Baizerman, and Compton (2002):

ECB is a context-dependent, intentional action system of guided processes and practices for bringing about and sustaining a state of affairs in which quality program evaluation and its appropriate uses are ordinary and ongoing practices within and/or between one or more organizations/programs/sites. (p. 8)

Further to this definition, it should be highlighted that ECB is also a process where the opportunity is created to reflect on and improve practice(s) to optimise the outcomes and impacts that can be achieved.

The initial question posed when considering ECB is: Who should be doing this? Should this be conducted and overseen by an external evaluator, or should this be co-ordinated internally? External evaluators bring a diversity of experience and knowledge to the table. However, as claimed by Ensminger et al. (2015), often organisational evaluation activity is driven internally rather than by externals. Bakken et al. (2014) reported that, when employees participate in evaluation, they take a greater sense of ownership suggesting that an internal approach may encourage greater buy-in from participants. It might be argued that an external evaluator may be less biased than an internal one, yet de Laat (2014) points out that unbiased evaluators do not exist; rather, it is the mode of operation and conduct instigated in the evaluation by the evaluator which manages for bias. Further, every organisation has its own politics, and often negotiating space for evaluation is better done from within an organisation rather than from the outside (Stockdill et al., 2002). Another reason for some science organisations using internal evaluators to build ECB is that they already have social science expertise which they may adapt to engage in the field of evaluation (Rennekamp & Engle, 2008). Indeed, Mackay and Horton (2003) identify that research-oriented evaluation can be defined as the “systematic application of social

science research procedures” (p. 145). Although well situated within their own fields of expertise, the social sciences and evaluation are, in the opinion of the authors, not unlike kissing cousins—they are genetically related, that is, through common data-collection methodologies and some theories, but boundaries or social norms are set within the wider family to ensure the purity of the field is maintained. The social scientist in the family seeks to understand and explore, but the evaluators in the family are often more judgemental, concerning themselves with the merit and worth of an evaluand.

Organisations new to evaluation and ECB need to define the important factors to successfully operationalise an ECB initiative. This reflects a utilisation-focused evaluation approach (Patton, 2012). There are two key functions that are equally required to build a successful culture of evaluation within an organisation: firstly the ability to do evaluation; and secondly the fortitude to utilise evaluation findings to prove or improve what has been evaluated. It is only by embracing these two key functions, that the optimal value for the evaluand, organisation, and associated stakeholders will be realised. It can be argued that organisational decision makers need to actually experience the benefits from an evaluation before they will commit to it and “embrace it as a leverage of change” (Cousins, Bourgeois et al., 2014; Cousins, Goh, Elliott, & Bourgeois, 2014). Organisations need to put support systems, structures, and leadership in place before they can effectively undertake evaluations either at the organisational or programme level. The addition of skills, resources, and processes will aid this. The ability to undertake evaluations (design, collect data, analyse, and make recommendations) leads to having access to information that effectively generates greater ability to critique the progress and achievement of the evaluand. The development of such evaluative inquiry enhances the likelihood of those findings being used to either make adjustments in the programme being evaluated

or to make decisions about its progression and future. As a result of this, learning becomes a very important function within a successful evaluation culture.

Preskill and Boyle (2008) contribute a multidisciplinary model that provides greater detail about how to support ECB. They identify specific strategies for ECB including coaching, internship, written materials, technology, meetings, appreciative inquiry, communities of practice, training, technical assistance, and involvement in evaluation (Preskill & Boyle, 2008). They also link a number of sustainable evaluation practices that help underpin the conceptual model by Cousins, Goh, Elliott, and Bourgeois (2014), such as the need for dedicated evaluation resources, strategic planning, and sharing and learning about evaluation.

Preskill and Boyle (2008) promote the concept of evaluative culture. Building an overall culture of evaluation within an organisation embedding evaluative practice (Bourgeois & Cousins, 2013). It is also useful to remember that ECB, as well as the resulting evaluative learning and thinking, is required not only at the organisational level, but also at multiple levels throughout the organisation, including the programme or individual level (Preskill & Boyle, 2008).

The two models (from Cousins et al. and Preskill & Boyle) together provide a useful foundation for science organisations for building an ECB culture of *doing* as well as *utilising* evaluation. Figure 1 brings the ECB picture together acknowledging both organisational- and programme-level impacts and linkages for “doing” and “using” evaluation. The next section explores the components of these models to our experience to date of embedding an evaluation culture at AgResearch.

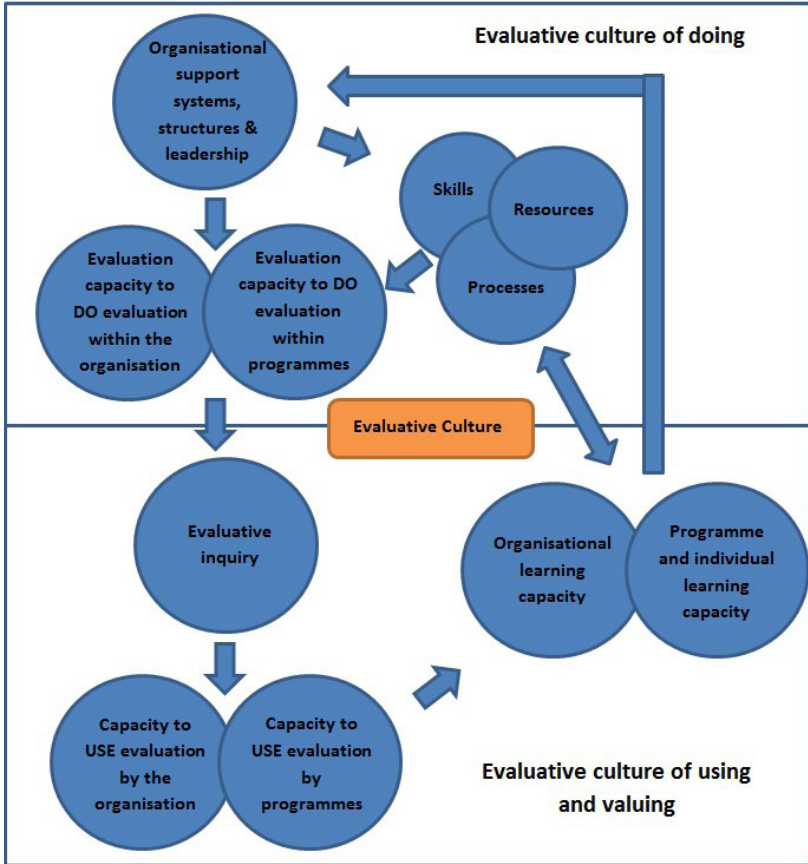


Figure 1. A model of ECB culture building within organisations and programmes (adapted from Cousins, Goh, Elliott, & Bourgeois, 2014 and Preskill & Boyle, 2008)

Organisational support systems, structures, and leadership in AgResearch

In accordance with Figure 1, organisational support systems have been a critical part of evaluation capacity building in AgResearch. The leader of the wider Adoption and Practice Change programme has

had a critical role to play in facilitating the building of organisational capacity for evaluation. This included ensuring that organisational support was in place, such as communicating with key people within the organisational structure, and ensuring the evaluation-focused members of the team had sufficient time and resources to participate in the programme. Further, the programme leader played a critical role in identifying where organisational learning could be built as a result of the ECB efforts, and facilitated access to science teams to encourage this.

The capacity to use evaluation was also facilitated by the Adoption and Practice Change programme leader by engaging key organisation decision makers, increasing their awareness of the opportunity and value of evaluation. This resulted in senior management buying into and supporting the need for evaluation to prove and improve the impact of our science. This became more tangible as senior management ensured direction for evaluation would be aligned with the new company-wide project management system. Management also agreed that significant funding proposals should be aligned with evaluation processes in order to assist project planning, and planning for evaluation as they are developed. Aligning outcome evaluation at the programme and project level had not previously been present in the organisation, and hence most biophysical scientists had not had the opportunity to use evaluation as a reflection and learning tool.

These three factors—organisational support structures, learning capacity, and capacity to use evaluation—are all key organisational features for successful ECB. They provide the necessary underpinnings for widespread evaluative activity and evaluation use (Bourgeois & Cousins, 2013; Cousins, Goh, Elliott, & Bourgeois, 2014) (see Table 1 below).

Embedding evaluation in AgResearch

Two key areas were important for ECB development at project level. The first was capacity building within the Adoption and Practice Change (A&PC) team. Project team members undertook training to build on their existing knowledge in evaluation, which enabled the members of the A&PC project to establish processes as well as evaluation tools, and communication tools to engage wider staff members in evaluation uptake. An internal evaluation plan was developed by the A&PC project team to enable tracking of progress towards identified outcomes. This was based on a theory of change developed within the programme and targeted objectives and milestones to work towards. An external evaluator was engaged to provide a critical review of the AgResearch programme, an important practice identified by Labin et al. (2012). This provided a wider outside perspective and helped counter any unidentified biases of the internal evaluators supporting this programme. Wider support was provided within the organisation including from technical IT and communication staff. External support assisted through the formation of the Impact, Planning and Evaluation Network (iPEN), a network of fellow CRI evaluators with the purpose to “improve the quality of impact and evaluation across the CRIs by sharing best practices, developing shared language and paradigms; and promoting quality within each organisation” (iPEN, 2015, p. P1). This network was created as all seven CRIs in New Zealand recognised the benefit of working together to build evaluation capacity as pressure grew to prove and improve their science and science impact.

Most members of the A&PC project team were also members of the larger People and Agriculture Social Science team. The People and Agriculture team has broad social science membership including geography, psychology, and environmental economics who use a

range of methodological and theoretical approaches to enhance A&PC within the agricultural sector. To bolster the capability of both teams, a training and mentoring programme was put in place to support other AgResearch colleagues when working with evaluation. This collaboration between the programme team and social research team (Figure 2), in which some staff had dual membership, was essential to combat the challenges faced due to the company's size, diversity of science, and multiple locations (Cousins & Bourgeois, 2014).

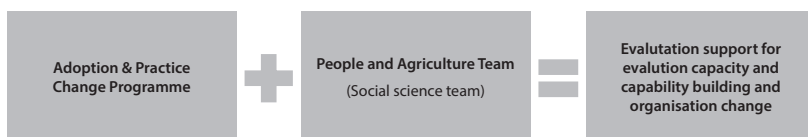


Figure 2. AgResearch's internal capacity to build evaluation capability

The second key focus area of building an evaluative culture and internal ECB was on the AgResearch science project managers and their research teams. The approach used reflects many of the ECB strategies and activities identified by Preskill and Boyle (2008) in their multidisciplinary model of ECB. Within AgResearch, these strategies included training opportunities to encourage familiarisation with processes, running research team workshops to engage whole teams in evaluation planning, providing technical assistance in response to queries, as well as access to tools, resources, and written materials including a website with guiding information (for a more comprehensive list, see Table 2). Labin et al. (2012) indicate that employing multiple approaches in this way helps to achieve optimal outcomes. The team identified the importance of not only ensuring that expertise is imparted when undertaking ECB activities but also trying to ensure that empowerment for learning occurs for everyone involved (Bakken et al., 2014). For this reason, a coaching approach was integrated into the organisation's ECB efforts. To achieve this, the evaluators worked with research project teams,

often in a participatory workshop, to facilitate the development of a programme logic, key evaluative questions, and measures to align with their science programmes.

Within AgResearch, the A&PC team identified the need for research project teams to develop and learn about evaluation, to build their own skills, and implement new knowledge into their programmes in order to achieve more effective science outcomes (Ensminger et al., 2015). It was important to empower the research project teams to be able to collect much of the data themselves and, being scientists, they were accustomed to data collection. Where capacity was required to engage in particular methodologies, such as interviews or network analysis, then capability from the wider People and Agriculture team was sought. Further, queries about approaches, methods, or support required was available by contacting and working with either the A&PC or People and Agriculture teams.

	AgResearch Management	Adoption & Practice Change (A&PC) Team	People and Agriculture Team	Research project managers & science teams
External & internal drivers	Expectations from Pastoral Industry science users Statement of Corporate Intent (SCI) Adoption & Practice Change Roadmap People and Agriculture Team Strategy (which has a key focus on evaluation) Science contracts			
Organisation support	Provides leadership and managerial support at different levels. Makes resource available. Enables A&PC Team to support new company processes. Endorses requirement for new funding proposals to include evaluation.	Provides adoption and practice change as well as evaluation leadership. Allocates funds toward improved evaluation. Makes staff time available. Supports evaluation coaching.	Provides support to research project managers and science teams.	Receive support in the form of evaluation advice and research capability provided by the A&PC Team and the People and Agriculture Team.

Table 1. Drivers and organisation support within AgResearch for ECB

Strategies and activities	Team or group of people			Research project managers & science teams
	AgResearch Management	Adoption & Practice Change (A&PC) Team	People and Agriculture Team	
Strategy	Learn about, value and support ECB in company.	Build and use team capability to support ECB	Build and use team capability, particularly coaching for improved ECB	Learn about and apply evaluation in science programs, receive individual and team coaching
Resources and activities	Provided feedback to A&PC team. Guided ECB through company hierarchy. Designed ECB friendly policies.	Provided training, written materials to other teams and groups. Defined steps for evaluation. Designed and made ECB support tools available. Provided technical support to other teams and groups. Gained and provided access to science teams. Designed and implemented a communication plan. Branding of evaluation programme to increase recognition. Attended science team meetings and run ECB workshops with them. Provided website with guiding information to other teams/ groups. Write and submit publications about evaluation and ECB. Co-designed and implemented iPEN (CRI evaluation network).		Attended ECB workshops. Received and provided technical support. Accessed written materials and use support tools. Utilised the Wwebsite. Learned how to participate in, plan for and collect evaluation data within ECB project. Utilised coaches.
Reflecting	Received ECB progress reports, reflected upon them and provided feedback and guidance.	Designed and implemented internal ECB programme evaluation. Received internal staff feedback, analysed and reported. Engaged and utilised critical external programme reviewer.		Participated in internal programme evaluation, e.g., surveys, feedback from participating in events etc.

Table 2
AgResearch strategies, activities, and resources that support ECB (based on Preskill & Boyle, 2008)

Challenges for science organisations and science teams to engage in new practices of research programme monitoring and evaluation

Organisational-level challenges

The challenge of introducing ECB into an organisation such as AgResearch, that has diverse research disciplines, teams, and research topics, is not one that can be tackled easily. There was a need to be selective in the evaluation opportunities that were available with science teams in order to use the available resources to best advantage. In the first year of activity, the A&PC team focused on getting voluntary participation in evaluation to make progress using the early adopters in the company. Following on from this, science teams participating in new funding rounds were required by the organisation to engage in evaluation processes, which in turn provided a much greater demand for ECB from science teams. This requirement caused a mixture of voluntary and mandated adoption of evaluative processes and practice. A much greater uptake was realised with the larger science programmes once requirements for them to build in evaluative practices was mandated. Having experienced these processes, programme leaders often then made requests for assistance to align evaluative practices with programmes of work despite not being mandatory at the time.

Tasked with ECB, the A&PC and People and Agriculture teams faced challenges associated with the size of the organisational challenge, the huge variety and sheer number of science projects and teams, as well as the logistics of, and managing, ECB processes across four geographically spread out campuses (two in the North Island and two in the South Island of New Zealand). Access and complexity issues such as these were also identified by Cousins and Bourgeois (2014) in their analysis of eight different organisations’

ECB activities. This highlights that, although organisations' ECB efforts are likely to differ in approach, there are common obstacles that organisations are encountering as they seek to build evaluation capacity. Organisations need to be cautious and aware of these kinds of challenges when attempting to initiate and embed a culture of evaluation.

Challenges for science teams with evaluation and ECB

Due to the very recent introduction of evaluation and ECB to the field of biophysical science, it is unlikely that this journey into ECB will be issue free at either the organisational level or at the individual programme level. This poses a question about what challenges await our science organisations and project teams as they engage in the new practices of ECB and programme monitoring and evaluation. Labin et al. (2012) in their synthesis of the ECB literature, and Cousins and Bourgeois (2014) in their cross-case analysis of multiple organisations' ECB efforts, have explored a number of the barriers to successful ECB. Contributing to this, we believe it would be useful to explore and document the challenges, barriers, and opportunities to introducing ECB into an organisation such as AgResearch which has had little previous evaluation exposure. This begins to establish a local evidence base and provides transferable learnings that may further benefit other organisations attempting to create an evaluation culture.

Within 2 years of instigating the ECB programme, AgResearch has had good organisational support. However, it is timely to look to the future to try to identify potential challenges to current progress that may occur at the organisational level (for example, changes in key staff or management, and shifts in resources and support). Another potential issue is the risk of disengagement with an evaluation culture in the lag phase between initiating an ECB programme within the organisation, and actual results coming through from

programmes of sizeable duration (both of these issues were experienced by DFID (2014b)). Ensminger et al. (2015) have shown that it can take up to 5 years for an evaluative culture to begin to embed. There are time lags for partial and full adoption of new processes, and AgResearch, like other organisations, should be prepared to maintain its investment in ECB for a sustained duration.

Through reflection on the literature and our experiences, a number of challenges have been identified as the A&PC team has sought to introduce evaluative thinking and learning with science project teams. These include the following:

- Additional investment (financial as well as time and effort) is required to collect data for monitoring and evaluation purposes over and above core science activities.
- Shifting output-minded scientists towards an outcome mindset; that is, linking what they intend to do, what they intend to achieve, and how they intend to show their impact.
- Scientists' understanding and learning new terminology and language from the evaluation field.
- Ensuring evaluative capacity is maintained within project teams and within the core A&PC team; that is, considerations of succession and sustainability.
- A value proposition is required to encourage science teams to adopt an evaluative culture:
 - Project teams need to be able to see the value of evaluation findings. This can be difficult when they are early in programme development.
 - Diversifying the type of outputs and outcomes and seeing value in this diversity; that is, not just focusing on scientific publications or numbers attending events.

- The importance of the soft skills, such as relationship building and facilitation, that are required to make ECB work is often undervalued.
- Forethought must be given to the value of effective evaluative reporting to ensure data analysis and utilisation is also a focus rather than just data collection to meet requirements (also identified by Labin et al. (2012)).
- Evaluation can be seen to be an obligation rather than an opportunity (also identified by Rennekamp and Engle (2008)).

The challenges identified here provide an opportunity to reflect on the dynamics of introducing groups to the practice of evaluation. Although science has its own unique characteristics, many of the challenges highlighted are not science specific and thus the reflections outlined above provide useful insights for other organisations as they develop their own evaluative capacity-building strategies and approaches.

Opportunities gained

A number of opportunities have been realised by AgResearch as we have undertaken an ECB journey on our way to trying to develop a culture of evaluation within our science. Among others, these include:

- the ability to engage with other science organisations to build ECB within the New Zealand CRI sector
- to increase understanding among scientists of the value of monitoring and evaluation
- to utilise evaluation processes to achieve not only better programme outcomes, but also better programme management
- to start to utilise new learning opportunities to improve and prove the impact of our science.

Progress so far

After 4+ years, a culture of evaluation is still far from being realised. However, internal surveys show staff are increasingly becoming familiar with evaluation tools and processes, and requests for evaluation services and support are ongoing, even when no mandatory requirement exists. Examples exist of some programme leaders developing confidence with evaluation processes building to the point where they are comfortable instigating these processes on their own and without assistance from the A&PC team.

The A&PC initiatives have continued to be supported by management, including support for Evaluation Champions—science staff who are trained to facilitate participatory workshops to help science teams undertake stakeholder assessments, theory of change development, and overarching evaluation plans. An ongoing challenge is to motivate science teams to get optimal value from these workshops and resources developed through transitioning from the “doing” phase of evaluation to the “using” phase.

Conclusion

It is evident from the AgResearch experience and the wider literature (Naccarella et al., 2007) that a number of lessons can be extrapolated to contribute to the field of ECB practice. Firstly, leadership is important at all levels of the organisation, including management, team leaders, evaluators, and project team leaders. Every leader’s commitment must be earned with a robust value proposition in order to maintain the demand and justify continued time, effort, and financial investments. Once earned, these leaders also become active champions for the organisation’s ECB efforts and for evaluation. Active, continued engagement to keep evaluation top-of-mind is needed throughout an organisation for this reason. Secondly, multiple communication avenues must be developed within organisations in order to achieve buy-in for valuing and using the findings from evaluation. Thirdly, it

is essential to recognise that embedding an evaluation culture in any organisation is an evolutionary process with time lags; that is, the delay that occurs between the activity(s) occurring and the resulting outcome or impact of that activity. Fourthly, organisations will need to create multiple avenues for evaluative engagement, activity, and learning in order to maximise the adoption of these practices, particularly at the programme level. This includes both voluntary and mandated adoption as part of the organisation's policies and practices. Tailoring ECB processes to specific teams' or funding proposal requirements will help maintain engagement at all levels. Fifthly, it is necessary to work to teams' strengths. For example, if data collection is a familiar activity, engaging the team in the evaluative data collection process can optimise resource use, increase buy-in, and encourage the use of results, not only to improve impact but also for continued research practice improvements. Finally, the issues of sustainability and succession of ECB should not be overlooked. Relying heavily on a few leaders or evaluators is a risk that can be mitigated with a strategic approach to ECB development within an organisation—an area that AgResearch continues to actively explore; and by having a team to help shepherd the development of evaluation capacity through the organisation.

In the opinion of the authors, the demand for ECB within New Zealand science will continue to increase. This will leave science organisations with little option but to embrace the opportunity to monitor and evaluate their work, and the impacts thereof. Due to the low exposure to an evaluation culture in the past, it will require a concerted effort and investment from the managerial level right down to individual project teams. By sharing experiences, and taking lessons learnt from other organisations, the findings as discussed in the article may help facilitate greater understandings of the challenges faced and potentially ease the transition to building and sustaining strong evaluative cultures within science organisations.

References

- AgResearch. (2013). *Adoption and practice change roadmap* (p. 48). Hamilton: Author.
- AgResearch. (2015). *Annual report 2014/15. Delivering science and technology to benefit New Zealand's pastoral sector* (pp. 1–68). Hamilton: Author.
- AgResearch. (2016a). *Annual report* (p. 70). Hamilton: Author.
- AgResearch. (2016b). *Statement of corporate intent 2016–2021* (p. 12). Hamilton: Author.
- Bakken, L. L., Núñez, J., & Couture, C. (2014). A course model for building evaluation capacity through a university–community partnership. *American Journal of Evaluation, 35*(4), 579–593. <https://doi.org/10.1177/1098214014523671>
- Bourgeois, I., & Cousins, J. B. (2013). Understanding dimensions of organizational evaluation capacity. *American Journal of Evaluation, 34*(3), 299–319. <https://doi.org/10.1177/1098214013477235>
- Bourgeois, I., Whynot, J., & Thériault, É. (2015). Application of an organizational evaluation capacity self-assessment instrument to different organizations: Similarities and lessons learned. *Evaluation and Program Planning, 50*(0), 47–55. <https://doi.org/10.1016/j.evalprogplan.2015.01.004>
- Brown, R. (2011). *Evaluation use at Environment Canada*. Masters project, University of Victoria.
- Clark, W. C., Mitchell, R. B., & Cash, D. W. (2006). Evaluating the influence of global environmental assessments. In R. B. Mitchel, W. C. Clark, D. W. Cash, & M. Dickson (Eds.), *Global environmental assessments: Information and influence* (pp. 1–28). Cambridge: MIT Press.
- Cousins, J., & Bourgeois, I. (2014). Cross-case analysis and implications for research, theory, and practice. *New Directions for Evaluation, 141*, 101–119. <https://doi.org/10.1002/ev.20078>
- Cousins, J., Bourgeois, I., & Associates. (2014). Multiple case study methods and findings. *New Directions for Evaluation, 141*, 25–99. <https://doi.org/10.1002/ev.20077>

- Cousins, J., Goh, S. C., Elliott, C. J., Aubry, T., & Gilbert, N. (2014). Government and voluntary sector differences in organizational capacity to do and use evaluation. *Evaluation and Program Planning*, *44*, 1–13. <https://doi.org/10.1016/j.evalprogplan.2013.12.001>
- Cousins, J., Goh, S. C., Elliott, C. J., & Bourgeois, I. (2014). Framing the capacity to do and use evaluation. *New Directions for Evaluation*, *141*, 7–23. <https://doi.org/10.1002/ev.20076>
- de Laat, B. (2014). Evaluator, evaluand, evaluation commissioner. A tricky triangle. In M. Laubli Loud & J. Mayne (Eds.), *Enhancing evaluation use. Insights from internal evaluation units* (pp. 15–36). Thousand Oaks, CA: Sage. <https://doi.org/10.4135/9781506335124.n2>
- Department for International Development (DFID). (2014a). *DFID evaluation strategy 2014–2019* (pp. 1–17). UK: Author.
- Department for International Development (DFID). (2014b). *Rapid review of embedding evaluation in UK Department for International Development* (pp. 1–77). UK: Author.
- Douthwaite, B., Kuby, T., van de Fliert, E., & Schulz, S. (2003). Impact pathway evaluation: An approach for achieving and attributing impact in complex systems. *Agricultural Systems*, *78*(2), 243–265. [https://doi.org/10.1016/S0308-521X\(03\)00128-8](https://doi.org/10.1016/S0308-521X(03)00128-8)
- Duignan, P. (2002). Building social policy evaluation capacity. *Social Policy Journal of New Zealand*, *19*, 179–194.
- Ekboir, J. (2003). Why impact analysis should not be used for research evaluation and what the alternatives are. *Agricultural Systems*, *78*(2), 166–184. [https://doi.org/10.1016/S0308-521X\(03\)00125-2](https://doi.org/10.1016/S0308-521X(03)00125-2)
- Ensminger, D. C., Kallemeyn, L. M., Rempert, T., Wade, J., & Polanin, M. (2015). Case study of an evaluation coaching model: Exploring the role of the evaluator. *Evaluation and Program Planning*, *49*, 124–136. <https://doi.org/10.1016/j.evalprogplan.2015.01.002>

- European Initiative for Agricultural Research for Development (EIARD). (2003). Impact assessment and evaluation, European Initiative for Agricultural Research for Development. *Agricultural Systems*, 78, 329–336.
- Impact, Planning and Evaluation Network (iPEN). (2015). *Terms of reference for Impact Planning and Evaluation Network (iPEN)*. Christchurch: Author.
- Kelley, T., Ryan, J., & Gregersen, H. (2008). Enhancing ex post impact assessment of agricultural research: The CGIAR experience. *Research Evaluation*, 17(3), 201–212. <https://doi.org/10.3152/095820208X331711>
- Labin, S. N., Duffy, J. L., Meyers, D. C., Wandersman, A., & Lesesne, C. A. (2012). A research synthesis of the evaluation capacity building literature. *American Journal of Evaluation*, 33(3), 307–338. <https://doi.org/10.1177/1098214011434608>
- Leviton, L. C. (2014). Some underexamined aspects of evaluation capacity building. *American Journal of Evaluation*, 35(1), 90–94. <https://doi.org/10.1177/1098214013502844>
- Mackay, R., & Horton, D. (2003). Expanding the use of impact assessment and evaluation in agricultural research and development. *Agricultural Systems*, 78(2), 143–165. [https://doi.org/10.1016/S0308-521X\(03\)00124-0](https://doi.org/10.1016/S0308-521X(03)00124-0)
- Midmore, P. (2017). The science of impact and the impact of agricultural science. *Journal of Agricultural Economics*, 68(3), 611–631. <https://doi.org/10.1111/1477-9552.12242>
- Naccarella, L., Pirkis, J., Kohn, F., Morley, B., Burgess, P., & Blashki, G. (2007). Building evaluation capacity: Definitional and practical implications from an Australian case study. *Evaluation and Program Planning*, 30(3), 231–236. <https://doi.org/10.1016/j.evalprogplan.2007.05.001>

- New Zealand Government. (2015). *Government response to the New Zealand Productivity Commission report on regulatory institutions and practices*. Wellington: Author. Retrieved from <http://www.treasury.govt.nz/regulation/nzpcresponse/response-nzpc-jul15.pdf>
- Norton, G. W. (2015). *Evaluating economic impacts of agricultural research: What have we learned? Seminar paper* (pp. 1–18). Department of Agricultural and Applied Economics, Virginia Tech, Blacksburg.
- Patton, M. Q. (2012). *Essentials of utilization-focused evaluation*. Thousand Oaks, CA: Sage.
- Preskill, H., & Boyle, S. (2008). A multidisciplinary model of evaluation capacity building. *American Journal of Evaluation*, 29(4), 443–459. <https://doi.org/10.1177/1098214008324182>
- Rennekamp, R. A., & Engle, M. (2008). A case study in organizational change: Evaluation in cooperative extension. *New Directions for Evaluation*, 120, 15–26.
- Stockdill, S. H., Baizerman, M., & Compton, D. W. (2002). Toward a definition of the ECB process: A conversation with the ECB literature. *New Directions for Evaluation*, 93, 7–25. <https://doi.org/10.1002/ev.39>
- White, T., & Sheath, G. (2009). *Tairāwhiti Land Development Trust program evaluation 2009*. Hamilton: AgResearch.

The authors

Toni White, Helen Percy, and Bruce Small

AgResearch Ltd, Private Bag 3123, Hamilton 3240

Email: Toni.White@agresearch.co.nz