ORIGINAL ARTICLE

Communities of practice as an effective ECB approach for nonevaluation specialists: A case study in an Australian state government department

Florent Gomez[1] 🔘 | George Argyrous[2] 🔘

[1]New South Wales Agency for Clinical Innovation, St Leonards, Australia

[2]Paul Ramsay Foundation, Sydney, Australia

Correspondence
Florent Gomez, New South Wales Agency for Clinical Innovation, St Leonards, Australia.
Email: florent.gomez@gmail.com

Abstract
Among the various evaluation capacity building strategies, there is an important opportunity to enhance the evidence on the effectiveness of communities of practice (CoPs). We argue that evaluation CoPs have great potential to develop evaluative thinking among nonevaluation specialists and build the foundations of a sustainable organizational evaluation culture. This article presents a case study of an evaluation community of practice in an Australian state government agency, with evidence collected over 4 years against a comprehensive evaluation framework. It outlines in detail an empirically tested evaluation-CoP good practice from the initial design and logic model to implementation and measurement against individual, group, and organizational outcomes.

CoPs AS ONE OF THE LESS RESEARCHED ECB STRATEGIES

Evaluation capacity building (ECB) is defined by Stockdill, Baizerman, and Compton as the "intentional work to continuously create and sustain overall organizational processes that make quality evaluation and its uses routine" (2002, p. 14). The seminal work by Preskill and Boyle identifies 10 different ECB strategies to help individuals, groups, and organizations "develop the knowledge, skills and attitudes to think evaluatively and to engage in evaluation practice" (2008, p. 444). These strategies reflect various teaching and learning approaches, ranging from internships to coaching or mentoring, and include communities of practice (CoPs). The most commonly used definition, by Wenger et al., defines CoPs

as groups of people "who share a concern, a set of problems, or a passion about a topic, and who deepen their knowledge and expertise on this area by interacting on an ongoing basis" (2002, p. 28).

The ultimate goal of ECB as per the Preskill and Boyle model is sustainable evaluation practice, "where members continually ask questions that matter; collect, analyze, and interpret data; and use evaluation findings for decision-making and action" (Preskill & Boyle, 2008, p. 444). Evaluative thinking has become an increased focus of ECB efforts, especially in the context of targeting nonevaluators. Evaluative thinking is essentially understood as critical thinking or a systematic inquiry mindset. Education researchers refer to evaluativist-level thinking, which is likened to higher-order thinking skills in the cognitive science tradition (Buckley et al., 2015). Buckley et al. argue that

> Not everyone in an organization or on a program team needs to be an evaluator or to do evaluation work. However, if everyone involved in planning, implementing, and evaluating a program is an evaluative thinker, the program and its evaluation have the best chance for success. (2015, p. 377)

Of the 10 different ECB strategies identified by scholars (Preskill & Boyle, 2008), not all have received equal attention in the research literature. One notable example, which is the focus of this paper, is the role of CoPs in ECB efforts.

As one of the ECB strategies identified by Preskill and Boyle (2008), CoPs have great potential to promote evaluative thinking to nonevaluators. CoPs sit within a long tradition of expertise-based knowledge-sharing peer forums—a tradition that includes "corporations" of craftsmen in classical Greece and guilds in the Middle Ages (Wenger & Snyder, 2000). They have been conceptualized as a specific teaching and learning strategy in the knowledge management literature, where they are positioned as forms of situated learning among peers (Lave & Wenger, 1991).

CoPs have been mostly studied in the private sector, in particular in engineering, business management, and health care services, and more recently in the public sector (Jørgensen et al., 2021). CoPs vary in their purpose, membership, and activities or methods of interaction—Wenger and Snyder argue that "CoPs are as diverse as situations that give rise to them" (2000, p. 140); a CoP may take the form of a group of functional experts from various teams or regions within the same organization; a group of new employees (Hatmaker et al., 2011); or a group of experts across multiple organizations focusing on addressing specific problems (Agranoff, 2008). However, CoPs share some common characteristics: They aim to facilitate knowledge sharing and learning, and often originate from a "gut-level conviction in [their] value" (McDermott, 2002, p. 26); they are not part of organizational formal structures, and they have a degree of informality—which brings the opportunity to break down organizational barriers (Ranmuthugala et al., 2011); and they have self-selected membership, with some form of give-and-take mechanism among members (Wenger & Snyder, 2000).

While they were originally viewed as organic entities, CoPs have been progressively considered as forums that could be purposefully designed by organizations (Harvey et al., 2013), in both business (Aljuwaiber, 2016) and public (Amayah, 2013) contexts. CoPs provide access to knowledge that can be explicit (know-that) or tacit (know-how), either directly via the exchanges taking place during CoP activities (Hatmaker et al., 2011) or by allowing members to identify those who have the expertise (know-who).

The literature, however, regularly points to a lack of empirical evidence on the effectiveness of CoPs (Bertone et al., 2013; Harvey et al., 2013; Li et al., 2009), in particular in the public sector (Jørgensen et al., 2021; Smith, 2016). Some of this evidence may be captured

in the grey literature, as pointed out by Ranmuthugala et al. (2011), which is why this article aims to bring practitioner-led evidence to the public domain.

Evaluating and reporting on the effectiveness of CoPs carries a number of measurement challenges, including the multiple levels of measurement (individual and organizational), the long timelines of impact, the hard-to-measure outcomes, the challenge of attributing impact to the CoP (McKellar et al., 2014), and the related issue of identifying appropriate counterfactuals at the organizational level. These challenges may explain, in part, the lack of published evidence including the dynamic nature of CoPs. This article contributes to the evidence base and addresses some of these challenges based on empirical evidence collected from an evaluation CoP in a government organization over 4 years.

THE CASE STUDY OF AN EVALUATION COMMUNITY OF PRACTICE IN AN AUSTRALIAN STATE GOVERNMENT DEPARTMENT OVER 4 YEARS

We explore the case study of an evaluation community of practice in an Australian state government department, the New South Wales Department of Customer Service, with evidence collected over 4 years, between November 2017 and January 2022. As per Yin (2009), case study is a form of social research inquiry where examining the context is particularly critical to understanding the case. We follow an inductive approach in exploring the effectiveness of this particular CoP within its specific context to identify a potential pattern for the use of CoPs to promote evaluative thinking in public organizations.

The New South Wales Department of Customer Service is a government department with around 10,000 staff at the time of this study and a large remit, including government service centers, regulatory agencies, particularly those related to work, health and safety, and consumer rights, as well as tax collection administration.

The evaluation CoP was launched in November 2017, and the time frame considered for the case study spans over 4 years, until early 2022. This article's lead author, Gomez, was instrumental in launching and leading the evaluation CoP. The initial drive for the CoP was to connect small pockets of evaluation capacity across the department. Its stated purpose was to provide an open an informal forum to share good evaluation practices across the organization; opportunities for ECB, for example, an introduction to program logic and/or how to draft a good evaluation brief to contract an external evaluator; and opportunities for networking and identifying evaluation practitioners across the organization. The CoP targeted evaluation as well as nonevaluation staff; however, the department had fewer than 10 dedicated evaluation staff at any one time. Participants in the CoP were mostly in project officer, senior project officer, and manager roles. Executives were a secondary audience, in line with the intended outcome of increasing leaders' awareness of and support for evaluation. The CoP was also open to participants from other government departments, in particular from other evaluation staff.

As of January 2022, the community of practice had 343 active members, that is, people who posted, read, or reacted to messages, and 45% of members had attended more than one event since November 2017 (Table 1). As with most CoPs, regular events were the cornerstone of the CoP learning strategy. Events were organized every quarter, face to face or online; each lasted 2 h, with a mix of presentations, table group discussions, panels, and interactive exercises. The CoP included two additional components: a corporate social media group (Yammer) used as an alternative to a blog and to offer a platform for members to engage in evaluation-related discussions between events; and intranet pages to provide easy access to key evaluation resources, including presentations from the CoP events. Between November 2017 and August 2021, 14 quarterly CoP events were organized, with an average of 65 participants at each event. Participation in individual events ranged

TABLE 1 Reach of the CoP.

		Apr 2019	Jan 2020	Jan 2021	Jan 2022
Quarterly events	Average number of participants per event since the start of the CoP in Nov 2017	56	56	58	65
	Number of active members having attended one or more events since the start of the CoP in Nov 2017	154	204	248	343
Intranet pages	Average number of unique users to the CoP page per month over the past year	23	32	14	46
	Average number of unique users to the program evaluation page per month over the past year	24	28	19	30
Yammer group	Number of members	112	300	332	470[a]
	Number of active people over the past 12 months	308	445	527	851[a]

Abbreviation: CoP, communities of practice.

[a] As of April 2022.

from 46 in September 2019 to 103 in August 2021, with a substantial increase in attendance in 2020–2021 with the move to online events. Participation was voluntary, and events were advertised via internal communication avenues such as newsletters and Yammer posts, allowing new staff members to be regularly added to the CoP. The CoP organizing team organized the events in coordination with a rotating chair identified from among various units across the organizations. The team and chair set event agendas and topics based on topics of interest identified by the chair and feedback from previous events, with the goal of ensuring a mix of presentations and interactive formats.

The theory of action (Funnell & Rogers, 2011, p. 34) of the CoP relied on these three components to activate a situated learning theory of change (Lave & Wenger, 1991), in which the CoP would offer an authentic context and a safe space for practitioners to learn from each other (Hatmaker et al., 2011). A program logic model was developed to outline the intended outcomes of the CoP at both individual and group levels; this logic framed evidence gathering and annual self-evaluations (Figure 1).

In the CoP literature, the effectiveness of CoPs is assessed against outcomes defined as changes in behavior or work practices influenced by participating in CoP activities or accessing CoP resources (Ranmuthugala et al., 2011). In our case, the CoP is an ECB strategy, and its success is being judged against ECB outcomes. Typical ECB outcomes at the individual level include increased knowledge and understanding of evaluation concepts, terms, and approaches; behavioral outcomes, such as increased evaluation skills, ability to develop logic models and design data collection instruments, ability to collect valid and reliable data, and ability to teach others about evaluation; and affective outcomes, including an increased commitment to evaluation, stronger positive beliefs about data and evaluation, and decreased evaluation anxiety and fear (Preskill & Boyle, 2009). These outcomes were also purposefully applied to nonevaluators to promote evaluative thinking, outcome framing, and some foundational measurement knowledge, which would form the basis of an evaluation culture across the organization. While these were the initially anticipated outcomes, it was not feasible nor intended for the evaluation of the CoP to capture all of them as part of its formative purpose; it focused primarily on individual-level and short-term outcomes, with a view to inform continuous improvement of the CoP.
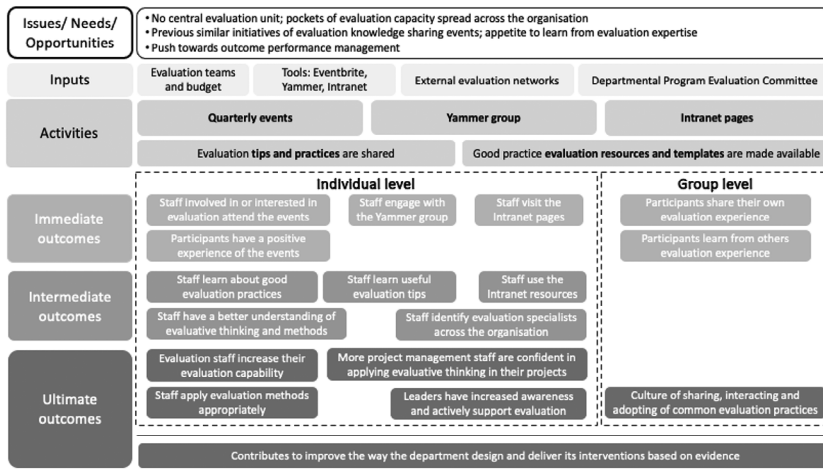
FIGURE 1   Program logic model. CoP, communities of practice.

A MIXED-METHODS APPROACH TO EVALUATE EVALUATION CoPs AT MULTIPLE LEVELS

In the case of the New South Wales Department of Customer Service Evaluation community of practice, the team coordinating the CoP conducted annual evaluations of the CoP in April 2019, January 2020, and January 2021 for formative purposes. This form of self-evaluation is deemed appropriate where people want to learn and progress from the analysis of their achievements against set objectives rather than for accountability (de Laat, 2013).

The objectives of these evaluations were to take stock of what had been achieved, assess the performance of the CoP, enable continuous improvements, and inform discussions about the future of the CoP. The intended users of the evaluation were the managers of the teams involved in coordinating the CoP. These evaluations brought together evidence collected from various methods and data sources against intended CoP outcomes and form the basis of the data considered in this article (Table 2). To inform continuous improvement of the CoP, data was collected since November 2017 about registration and attendance records and feedback about individual events. Annual surveys of CoP members were also conducted in April 2019 ($n = 45$), November 2020 ($n = 43$), and December 2021 ($n = 55$) to regularly take stock of the CoP achievements and measure impact. These surveys were administered to members of the CoP who had attended one or more events over the year; however, they do not constitute a longitudinal study, as the membership of the CoP continuously changed. On top of broad feedback about the CoP events and other components, these annual surveys included questions about reasons for attending, application of CoP learning in participants' workplaces, and involvement in evaluation activities, along with retrospective pretests (also referred as post–pretests). In these questions, individual participants were asked to assess their own understanding and knowledge of three key evaluation areas before and after attending the CoP over a year: what evaluation is and is not, program logic, and survey design. The annual survey did not, however, collect data about the number of events attended by respondents nor did it allow for linkage with participation data.

Retrospective pretests are appropriate instruments, in the context of measuring changes in knowledge and skills, to counter the threat to internal validity caused by response shift bias (wherein respondents rate their prior knowledge differently in hindsight—"You don't

TABLE 2 Evaluation methods.

Method	Time frame	Focus
Contact database	Nov 2017–Jan 2022	CoP participants by organization and position
Event participation data (Eventbrite)	Nov 2017–Aug 2021	Events attendance
Event agendas (document review)	Nov 2017–Aug 2021	Types of agenda items
Budget spreadsheet	Nov 2017–Dec 2021	Cost of the CoP; cost per event
Event surveys	Nov 2017–Aug 2021	Participants' satisfaction with events organization; feedback on agenda items
Annual surveys	April 2019, Nov 2020, and Dec 2021	Application of learning, change in evaluation understanding and skills, involvement in evaluation activities, evaluation maturity, organizational maturity assessment
Yammer stats	Nov 2017–Apr 2022	Level of engagement
Intranet stats	Nov 2017–Apr 2022	Access to resources
Staff interviews	Apr 2021	Examples of changes

Abbreviation: CoP, communities of practice.

know what you don't know"). However, these questions also come with increased risks of bias, such as social-desirability bias, compared to traditional pre- and posttests. The choice of the appropriate data collection methodology depends on what is being measured and the context for measurement and must strike the right balance between feasibility, in particular in terms of impact on participants, and rigor. Hill and Betz (2005) argue that retrospective ratings are appropriate where the objective is to describe changes as experienced subjectively by participants (compared to providing an estimate of mean program effects, an objective for which prospective ratings are more appropriate). Retrospective pretests are also better suited to the context of a CoP targeting nonspecialists where the group of participants evolves over time; they offer the opportunity for participants to reflect on how much they have learned as part of the intervention (Hill & Betz, 2005), which is particularly well suited to situated learning mechanisms such as CoPs.

As pointed out by Brown and Reed (2002), frameworks for evaluating training effectiveness tend to focus on change at either the individual or the organizational level, but rarely conceptualize the interrelationships between them. As part of evaluating the impact of the CoP, we collected feedback from CoP members about their perception of where the organization was sitting on key dimensions of organizational-level evaluation maturity, as part of the 2020 and 2021 annual survey. The purpose was to first identify areas for future ECB efforts at the organizational level, and second to identify changes over time, some of which the CoP may have contributed to.

The instrument included a series of questions across six dimensions to capture organizational evaluation maturity: evaluation culture, ECB, evaluation policies and procedures, data and IT systems to support evaluation, use of evaluation of program design, and approach to evaluation planning. The instrument was designed by one of the authors (Argyrous) using an analytic rubric approach that decomposes organizational evaluation maturity into dimensions (Davidson, 2005, pp. 137–142) where each of the five response options for each dimension was a statement reflecting a different level at which an organization might sit, ranging from 1 (*opposed*) to 5 (*leading*). This instrument (Figure 1A) had been developed and refined through previous applications in other organizational capacity
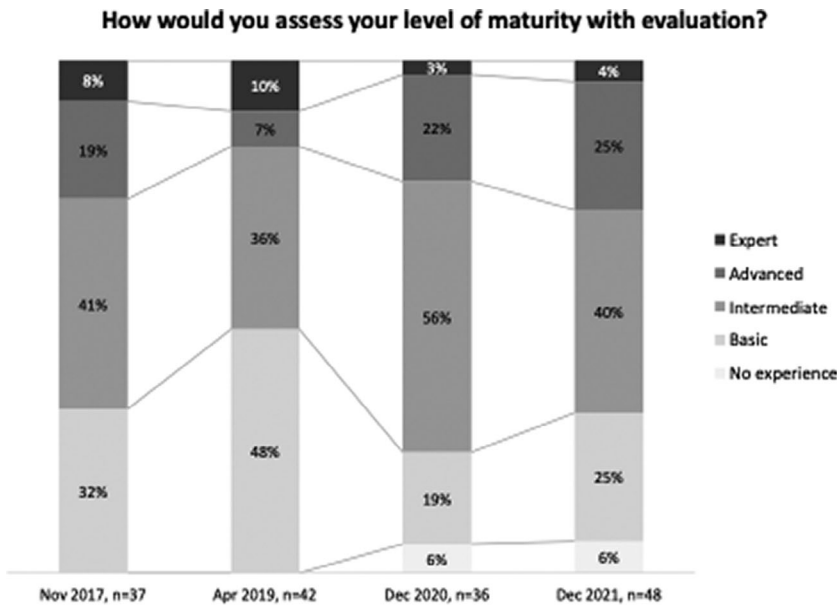
How would you assess your level of maturity with evaluation?

FIGURE 2 Self-reported evaluation maturity of CoP members. CoP, communities of practice. CoP, communities of practice.

building contexts and through extensive consultations with the Australian evaluation community.

RESULTS SHOW THAT CoPs CAN BE A COST-EFFECTIVE ECB STRATEGY TO PROMOTE EVALUATION SKILLS AND EVALUATIVE THINKING AMONG NONEVALUATORS

Participants' satisfaction with the CoP events was very high, with an average rating of their overall experience of 4.3 out of 5 across 10 events where such feedback was collected. Out of the 64 agenda items covered across the 14 quarterly events since November 2017, 50% were presentations—in particular of evaluation case studies. Other formats included executive addresses, group discussions, panels, and interactive exercises. In event feedback surveys, participants repeatedly indicated their preference for case studies. When asked about what they liked most in the CoP more broadly via annual surveys, the most frequent theme was around "sharing and learning from others in a collegiate, supportive environment," in the words of one CoP member.

As part of the annual survey, members of the CoP were asked to self-assess their level of evaluation maturity (Figure 2). This both provided a picture of the diversity of level of evaluation maturity across members at those different points in time, and gave an indication of overall progress made in terms of lifting evaluation maturity—and, more specifically, confidence—across the organization. The very first survey was conducted as a baseline following the first CoP event at the end of 2017. Results of the second annual survey in April 2019 may be interpreted, in line with anecdotal comments made, as reflecting response shift bias; some participants may have realized, for instance, that they had not understand as well as they thought what evaluation was. This interpretation is in line with one of the common characteristics of CoPs: offering a safe space for members to admit lack of knowledge (Smith, 2016), and it supported the choice of retrospective pretests as appropriate measurement instrument.
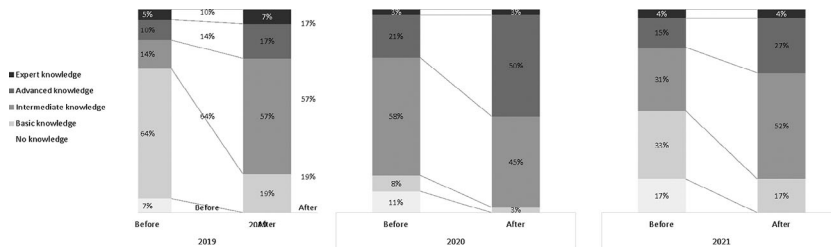
FIGURE 3 Retrospective pretests on CoP members' understanding of what evaluation is and is not. CoP, communities of practice.
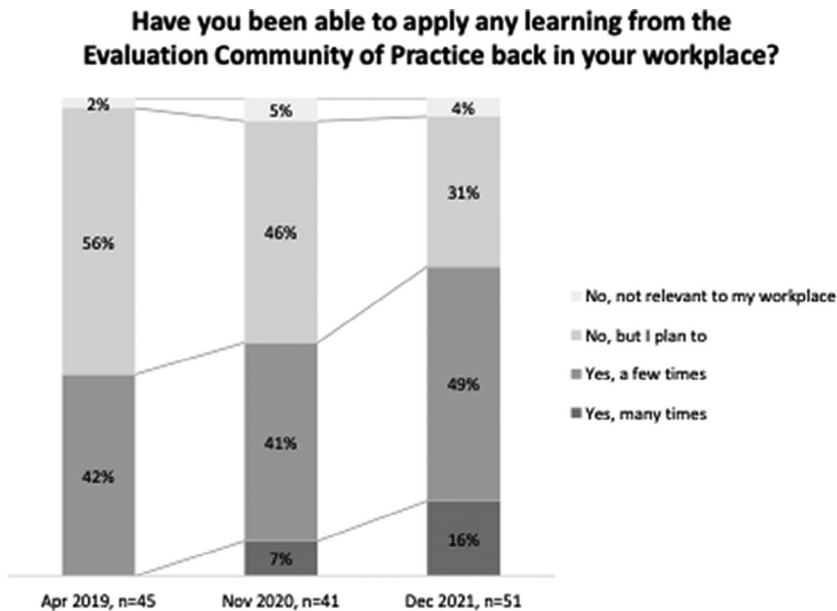

FIGURE 4 CoP members' reported ability to apply learning from the evaluation community of practice back to their workplaces. CoP, communities of practice.

Results of the annual survey retrospective pretest questions show a substantial shift in participants' reported knowledge since joining the CoP toward intermediate and advanced levels of knowledge, especially with regard to understanding what evaluation is and is not. As the intent was to measure change experienced subjectively by individual CoP members rather than estimate a mean effect of the CoP (Hill & Betz, 2005), we report here detailed results with the breakdown by reported level of understanding of what evaluation is and is not before and after joining the CoP, as reported in each of the annual surveys (Figure 3). In 2020 and 2021, CoP members reported a higher previous level of knowledge compared to 2019. These results indicate progress toward ECB outcomes within the context of the dynamic nature of the CoP, where some members have been part of the CoP over several years, while others joined later.

To get a sense of how participation in the CoP generated actionable and actioned learning, CoP members were asked in the annual survey about the extent to which they had been able to apply any learning from the CoP back to their workplace, with response options also capturing intention (Figure 4). This is in line with the theory of planned behavior (Ajzen, 1985), where behavioral intention is assumed to be an immediate antecedent of behavior and is based on attitude toward the behavior, subjective norm, and perceived
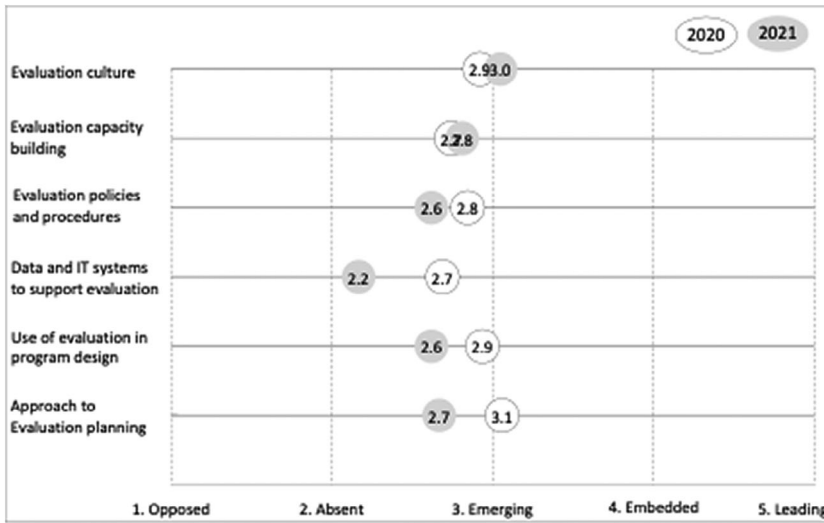
FIGURE 5 Organizational evaluation maturity assessment, 2020 and 2021. CoP, communities of practice.

behavioral control (or self-efficacy). Results show an increase in CoP members' ability to apply learning back in their workplace with intention. While 42% of respondents reported actual application of learning in April 2019, 65% did so in December 2021. Examples of how participants applied what they learned from the CoP include budgeting and procuring an external evaluation as well as designing new programs.

The CoP also proved to be a highly cost-effective ECB approach for internal decision-makers, with an average cost per event (in Australian dollars) of $423, or $9.60 per participant over 14 events (costs substantially decreased with the move to online format following the start of the COVID-pandemic). These costs do not include the costs of staff time to coordinate the CoP, which are already funded by the organization and are not considered as additional expenses. As a comparison, formal external training usually involves a much higher cost (such as AU$2,000 per participant for four modules, based on external evaluation training contracted at the time) and does not reach such a large audience. It is acknowledged, however, that the cost-effectiveness of the CoP was mainly driven by the fact that it was coordinated internally by evaluation staff on top of their project work. Efforts were made to limit imposition on these staff members by developing streamlined processes. These evaluation staff would also benefit from this investment in time through their own learning, connections across the organization, and most importantly by contributing to getting greater engagement and inputs from nonevaluation staff during projects.

The results of the organizational maturity assessment via the 2020 and 2021 annual surveys showed slight progress on the evaluation culture and ECB dimensions, while other dimensions (data and IT systems to support evaluation, approach to evaluation planning, use of evaluation in program design, and evaluation policies and procedures) went backward (Figure 5). This may show that the CoP had a positive impact on the dimensions it more directly contributes to as per the intended group-level outcomes and that ECB is part of a broader complex system. Evaluation policies and procedures in the department are determined by whole-of-government circulars and guidelines, which were under review in 2021—new evaluation guidelines at the whole-of-government level were expected to trigger the development of new policies and procedures at the departmental level when released. Another hypothesis is that CoP participants become more aware and critical of these organizational dimensions as they increase their evaluation maturity.

EVALUATION CoPs AS A KEY STRATEGY IN THE ECB MIX TO LAY THE FOUNDATIONS OF A SUSTAINABLE ORGANIZATIONAL EVALUATION CULTURE

When it comes to ECB, public administrations often put the focus on setting up overarching evaluation frameworks and/or individual training programs. However, what is frequently lacking is appropriate strategies to support the development of a sustainable evaluation culture by developing foundational evaluation skills among nonevaluation staff. In this regard, as demonstrated through the case study examined over 4 years, CoPs offer great potential, in particular by promoting outcome framing and evaluative thinking skills among nonevaluators. Results show progress on key ECB behavioral outcomes at the individual level, as well as some progress on a couple of organizational dimensions the community of practice contributed to (ECB and evaluation culture).

Evaluation CoPs have the potential to lay the foundations of an organizational evaluation culture that formal evaluation policies and training can build on to successfully develop sustained evaluation practice in public administrations. They also provide a platform to learn tacit knowledge (know-how) and where to go for deep expertise (know-who). We argue that prescriptive evaluation policies and formal evaluation training are much less likely to develop sustained evaluation practice in the absence of such foundations. Considering the cost-effectiveness of CoPs, such a strategy should be considered in the ECB mix developed by public administrations to promote sustainable evaluation culture and practice, on top of overarching evaluation frameworks and individual training programs.

Within the context of the considered CoP case study, ongoing collection of feedback from participants allowed the organizing team to continuously adapt its offerings to expectations from a dynamic member base. The COVID-19 pandemic also forced a move to online events, which substantially increased attendance at CoP events and subsequent engagement on the online platform but reduced opportunities for networking (one of the intended outcomes of the CoP). Building on these strong foundations, future endeavors could consider some cocreation mini-projects where a subset of CoP members would work together over several months to develop particular evaluation guidance or tools, in particular following the release of new government-wide evaluation guidelines. This cocreation function is common among CoPs (Wenger & Snyder, 2000), where members discuss common challenges they face and either identify or build on an existing good practice that some of the members may have already developed (Jørgensen et al., 2021), or else develop together a new way to address the common problem via some form of collaborative action research process.

CONCLUSION

The CoP presented in this case study, led by specialist evaluation practitioners for nonevaluators, proved particularly effective in breaking down the "expert wall" and laid the foundations of an evaluation culture beyond the confines of evaluation teams. It is important to note, however, the characteristics particular to this case study that may be different in other contexts. The public sector has lower staff turnover compared to other sectors, which allows for such interventions and longitudinal design over longer periods of time. This is particularly critical for such organizational capacity building interventions: This CoP was evaluated over 4 years, and in that time shifts in the organizational culture only started to be observed. The program being delivered and evaluated by an internal evaluation team also allowed stability in the resources allocated to coordinating the CoP, while benefiting the team itself by expanding the evaluation capacity across the organization.

Such an approach may be more challenging for external ECB practitioners or in smaller organizations.

This case study also provides a robust framework for the evaluation of evaluation CoPs against ECB outcomes and could be adapted to other types of CoPs. The evidence and insights generated by the annual evaluations were successful not only in informing continuous improvement of the CoP processes, but also in securing continued support for the CoP, including the allocation of a dedicated budget, by demonstrating its cost-effectiveness. Participation in CoP events and engagement online kept strong over time across the whole organization, that is, not just the core evaluation teams, while other CoPs may find it challenging to maintain momentum after a year and a few successful events, as observed in other evaluation CoPs launched around the same time in other government departments. Engagement was maintained over time and across the organization, by shaping agendas with the respective rotating chairs and presenting case studies, as requested by participants in events feedback. Improvements informed by the annual evaluations included streamlining events logistics with the use of templates and tools such as Eventbrite; establishing a rotating chair system; shaping event agendas according to capability needs and participants' feedback, especially requesting case studies; improving engagement in the corporate social media CoP group; and improving data collection processes. Areas identified for further exploration in the evaluation related to the optimal format for delivery of CoP events post-COVID-19 restrictions, and organizational evaluation maturity areas of concern that the CoP could help improve.

Internal evaluators planning to set up evaluation CoPs can build on the successful strategy and comprehensive measurement approach used for this CoP. This measurement approach could also be replicated for any type of CoP, in particular by considering the sequence of intended outcomes at both individual and group levels. The limitations identified with the evaluation strategy relate to its formative purpose; it focused on continuous improvement and did not collect comprehensive evidence to assess the effectiveness of the CoP. In particular, there was missing evidence on the whole organization staff mix to compare CoP participants and reach with. Importantly, the data collected did not allow linking outcome data with level of participation, which would have allowed analysis of whether the level of participation was reflected in the level of outcomes. Evaluation practitioners evaluating CoPs should consider asking for consent to link participation data with survey data for research and evaluation purposes to help assess different responses to the intervention.

In terms of methods, it would be recommended to complement the assessment approach used here with some more regular interviews with CoP members, in line with common methods used in the CoP literature (Ranmuthugala et al., 2011), to unpack the influence of CoPs in practice changes, among other factors. Measurement tools specifically focusing on evaluative thinking, such as those developed by Archibald et al. (2011), would also help capture the impact of the CoP in that regard at the individual level—these tools can take the form of self-reporting (evaluative thinking inventory), observation, or focus groups. Future research may further explore the intersection among individual, group-level, and organizational outcomes to further enhance the understanding of how to build a sustainable evaluation culture.

ORCID

Florent Gomez 🔟 https://orcid.org/0000-0003-1637-6961
George Argyrous 🔟 https://orcid.org/0000-0003-3121-2092

REFERENCES

Agranoff, R. (2008). Enhancing performance through public sector networks: Mobilizing human capital in communities of practice. *Public Performance & Management Review, 31*(3), 320–347.

Ajzen, I. (1985). From intentions to actions: A theory of planned behavior. In J. Kuhl & J. Beckmann (Eds.), *Action control: From cognition to behavior* (pp. 11–39). Springer.

Aljuwaiber, A. (2016). Communities of practice as an initiative for knowledge sharing in business organisations: A literature review. *Journal of Knowledge Management, 20*(4), 731–748. https://doi.org/10.1108/JKM-12-2015-0494

Amayah, A. (2013). Determinants of knowledge sharing in a public sector organization. *Journal of Knowledge Management, 17*, 454–471. https://doi.org/10.1108/JKM-11-2012-0369

Archibald, T., Buckley, J., & Trochim, W. (2011). *Evaluative thinking: What is it? Why does it matter? How can we measure it.* Paper presented at the annual conference of the American Evaluation Association, Anaheim, CA, United States. (3 November 2011).

Bertone, M. P., Meessen, B., Clarysse, G., Hercot, D., Kelley, A., Kafando, Y., Lange, I., Pfaffmann, J., Ridde, V., Sieleunou, I., & Witter, S. (2013). Assessing communities of practice in health policy: A conceptual framework as a first step towards empirical research. *Health Research Policy and Systems, 11*, 1–13. https://doi.org/10.1186/1478-4505-11-39

Brown, R. E., & Reed, C. S. (2002). An integral approach to evaluating outcome evaluation training. *The American Journal of Evaluation, 23*(1), 1–17. https://doi.org/10.1177/109821400202300102

Buckley, J., Archibald, T., Hargraves, M., & Trochim, W. M. (2015). Defining and teaching evaluative thinking: Insights from research on critical thinking. *American Journal of Evaluation, 36*(3), 375–388. https://doi.org/10.1177/1098214015581706

Davidson, E. J. (2005). *Evaluation methodology basics: The nuts and bolts of sound evaluation.* SAGE Publications Inc.

de Laat, B. (2013). Evaluator, evaluand, evaluation commissioner: A tricky triangle. In M. L. Loud & J. Mayne (Eds.), *Enhancing evaluation use: Insights from internal evaluation units* (pp. 15–26). SAGE Publications. https://doi.org/10.4135/9781506335124

Funnell, S. C., & Rogers, P. J. (2011). *Purposeful program theory: Effective use of theories of change and logic models.* John Wiley & Sons.

Harvey, J. F., Cohendet, P., Simon, L., & Dubois, L. E. (2013). Another cog in the machine: Designing communities of practice in professional bureaucracies. *European Management Journal, 31*(1), 27–40. https://doi.org/10.1016/j.emj.2012.07.008

Hatmaker, D. M., Park, H. H., & Rethemeyer, R. K. (2011). Learning the ropes: Communities of practice and social networks in the public sector. *International Public Management Journal, 14*(4), 395–419. https://doi.org/10.1080/10967494.2011.656051

Hill, L. G., & Betz, D. L. (2005). Revisiting the retrospective pretest. *American Journal of Evaluation, 26*(4), 501–517. https://doi.org/10.1177/1098214005281356

Jørgensen, R., Edwards, K., Scarso, E., & Ipsen, C. (2021). Improving public sector knowledge sharing through communities of practice. *VINE Journal of Information and Knowledge Management Systems, 51*(2), 318–332. https://doi.org/10.1108/vjikms-08-2019-0115

Lave, J., & Wenger, E. (1991). *Situated learning: Legitimate peripheral participation.* Cambridge University Press. https://doi.org/10.1017/CBO9780511815355

Li, L. C., Grimshaw, J. M., Nielsen, C., Judd, M., Coyte, P. C., & Graham, I. D. (2009). Evolution of Wenger's concept of community of practice. *Implementation Science, 4*, 1–8. https://doi.org/10.1186/1748-5908-4-11

McDermott, R. (2002). Measuring the impact of communities: How to draw meaning from measures of communities of practice. *Knowledge Management Review, 5*(2), 26–29.

McKellar, K. A., Pitzul, K. B., Yi, J. Y., & Cole, D. C. (2014). Evaluating communities of practice and knowledge networks: A systematic scoping review of evaluation frameworks. *EcoHealth*, *11*, 383–399. https://doi.org/10.1007/s10393-014-0958-3

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

Preskill, H., & Boyle, S. (2009). Insights into evaluation capacity building: Motivations, strategies, outcomes, and lessons learned. *Canadian Journal of Program Evaluation*, *23*(3), 147–174. https://doi.org/10.3138/cjpe.0023.008

Ranmuthugala, G., Plumb, J. J., Cunningham, F. C., Georgiou, A., Westbrook, J. I., & Braithwaite, J. (2011). How and why are communities of practice established in the healthcare sector? A systematic review of the literature. *BMC Health Services Research*, *11*(1), 1–16. https://doi.org/10.1186/1472-6963-11-273

Smith, A. E. (2016). Knowledge by association: Communities of practice in public management. *Public Administration Quarterly*, *40*(3), 655–689. https://www.jstor.org/stable/24772885

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*, 1–25. https://doi.org/10.1002/ev.39

Wenger, E., McDermott, R. A., & Snyder, W. (2002). *Cultivating communities of practice: A guide to managing knowledge*. Harvard Business School Press.

Wenger, E. C., & Snyder, W. M. (2000). Communities of practice: The organizational frontier. *Harvard Business Review*, *78*(1), 139–146. https://hbr.org/2000/01/communities-of-practice-the-organizational-frontier

Yin, R. K. (2009). *Case study research: Design and methods* (Vol. 5). SAGE Publications Inc.

APPENDIX

Scale point	Evaluation culture	Evaluation and research capability building	Evaluation policies and procedures	Data and IT systems to support research/evaluation	Use of research and evaluation for program design	Evaluation planning and practice
Opposed	People are scared to use evaluations; they actively oppose it or are hostile to it	Developing skills in evaluation and research is seen as a waste of time or discouraged	1. Our policies and procedures work against the use of evaluation in decision-making	Barriers and roadblocks are put up to prevent data and IT systems from supporting research and evaluation	Staff are discouraged from using research and evaluation to design and implement projects and programs	Staff plan to avoid assessing the effectiveness of their programs or are discouraged from assessing the effectiveness of their programs
Absent	I don't feel there is any culture of evaluation; it is not understood or else it is seen as unnecessary	No effort is put into building evaluation and research capability among staff	There are no formal evaluation policies and procedures in place	Data and IT systems are not designed to support research and evaluation	Research and evaluation are not taken into account when projects/ programs are designed and implemented	Staff think that evaluation planning is not needed. Programs either lapse or keep rolling on without any thought of evaluating them
Limited/ Emerging	There is some awareness of the benefits of evaluation and the need for evidence-informed decision-making	There is awareness that staff need skills in evaluation and research, and some training and support is provided	There are some formal evaluation policies and procedures in place	Relevant data are sometimes available to support research and evaluation but systems could be better integrated	Research and evaluation inform some projects and programs, but in an ad hoc way	Some projects or programs have evaluation plans in place. When evaluation happens it is infrequent and ad hoc
Embedded	There is widespread awareness of the benefits of evaluation, and evidence-informed decision-making is highly valued	Evaluation and research skills are actively supported through targeted training and recruitment	Evaluation policies and procedures are extensive, robust, integrated and of proven effectiveness	Data and IT systems fully support the use of data for research and evaluation	Research and evaluation inform most projects and programs from start to finish	Evaluation plans are in place for most projects and programs. Evaluations prioritized and scaled according to importance and risk
Leading	The organization regularly comes up with new ways to show staff the value it places on evaluation	Our evaluation and research skills are recognized by others as 'cutting edge'	Evaluation policies and procedures are regularly updated to reflect best professional practice	Data and IT systems are used in new and innovative ways to inform research and evaluation	Our projects and programs are innovative because of the way we draw on research and evaluation	We use innovative approaches to evaluating projects or programs. We learn and innovate from past practice
	Don't know/ can't say	Don't know/ can't say	Don't know/ can't say	Don't know/ can't say	Don't know/ can't say	Don't know/ can't say

FIGURE 1A Organizational maturity capability matrix. CoP, communities of practice.

AUTHOR BIOGRAPHIES

Florent Gomez is associate director for evidence generation and dissemination at the New South Wales Agency for Clinical Innovation and has over 15 years' experience as an evaluation practitioner across consulting and government in Europe and Australia.

George Argyrous is head of measurement, evaluation, research, and learning for the Paul Ramsay Foundation and has extensive teaching experience in evaluation, political economy, research methods, and statistics.