# Challenges and possibilities in developing a programme-theory model through stakeholder engagement, dialogue, and reflection

**Aneta Cram, Rodney Hopson, Marvin Powell, Asia Williams, and Akashi Kaul**

Programme theory (PT) development can prove challenging, as power dynamics among stakeholders and/or between the evaluator and stakeholders can be hard to navigate. An important part of the PT-development process is navigating the points of knowledge of a programme and merging them to gain accurate insight on the programme and outcomes. This process typically involves inviting the perspectives of different stakeholder groups and consolidating these perspectives in a visual depiction. This article analyses the PT-development process for the STEM + Families programme of the National Parent Teacher Association (NPTA), based in Alexandria, Virginia (United States). This initiative seeks to increase access to, interest in, and understanding of STEM careers and education pathways for all children, and especially girls, children of colour, and children living in low-socioeconomic communities. The article explores how dialogue, challenges and questioning, and reflection on organisational culture within the NPTA were core components of

this process and the eventual PT-model development. In particular, this article will focus on how the PT revealed sources of tension and power dynamics, and illustrated the possibilities of and challenges to the evaluation process, from the beginning of the process to report writing. Finally, the article adds to the evaluation field, as it explores the key evaluation competencies and key learnings from the described process, which hold implications for PT-development processes more broadly. Explorations of situational analysis, reflective practice, and the political and contextual environment in which the evaluation is placed emerge as key considerations and essential skill sets in PT development.

## *Introduction to programme theory and article outline*

Programme-theory (PT) development is an integral part of the evaluation process. It encompasses representations of the activities, expected outcomes, programme mechanisms, and, in some instances, contextual environment in which a programme is situated (Funnell & Rogers, 2011). PTs are used in evaluation as a tool for understanding how a social programme is working as an intervention—in a given context or multiple contexts—to address one or more social problems (Astbury & Leeuw, 2010; Funnell & Rogers, 2011; Quinn-Patton, 2008; Renger & Bartel, 2013).

In this article, we explore the process of developing a PT model as part of a larger one-year evaluation for the STEM (science, technology, engineering, and mathematics) + Families programme of the National Parent Teacher Association (NPTA). We also identify challenges through an analysis of the stakeholder-engagement process, established to inform the PT model, and reflect on how these challenges arose.

We begin with an overview of dialogue and culturally responsive practice, and a brief introduction to the PT-development project and

the wider evaluation in which it was situated. Next, we explore the background of the larger evaluation of the NPTA programme and the development of the PT processes. The penultimate section highlights the challenges, tensions, and learnings from the PT-development process. A conclusion and discussion section summarises key learnings and implications for capability building in the field of evaluation.

## Overview of dialogue and its relational and culturally responsive use in evaluation practice

The practice of evaluation is inherently dialogic, and the importance of dialogue has been well documented (Greene, 2001; House & Howe, 1999, 2000; Mark, 2000; Rallis & Rossman, 2000; Ryan & DeStefano, 2000; Schwandt, 2001; Widdershoven, 2001). Additionally, there are multiple conceptions of dialogue that extend beyond evaluation practice (Bakhtin, 1981; Bohm, 1996; Buber, 1961). In this section, the use of dialogue in evaluation practice is explored, initially by highlighting the importance of dialogue or the dialogic in evaluation based on foundational references and considerations. Next, the section discusses the importance of dialogue to characteristics that define its relational and culturally responsive use in evaluation practice.

> A key characteristic of dialogue, the constant test and interchange of ideas and opinions, resembles closely the process of learning about evaluation: the test of practical experiences against theoretical concepts and vice versa. And there is, of course, the fact that the act of evaluation itself is profoundly dialogic. Evaluators have to communicate well, listen well, reflect back what they hear, and anticipate what they may come to understand (Christie & Rose, 2003).

Abma (2006) highlights the importance of dialogue and relationship building in reference to good practice in evaluation. Abma states that dialogue is a tool that can be used to broker and minimise power

dynamics, and that communication around the evaluation and what is involved can alleviate tension between groups to support more open and trusting relationships. Such communications can also allow all parties involved in the evaluation to feel that they are included and supported, and that their voices are being heard (Abma, Leyerzapf, & Landeweer, 2017; Habermas, 1984–7).

Habermas (2003) states that shared understanding and communication is reached when multiple parties are willing to understand others' points of view. Dialogue can change and move perspectives from viewing from a place of hierarchy to accessing and acknowledging a diversity of experiences with, and input on, the programme—and, by extension, strengthen the evaluation (Abma et al., 2017). Dialogue is required throughout the evaluation process, as evaluators are aware of and responsive to the environment within which, and the people with whom, they are working (Hood, Hopson, & Frierson, 2015).

Being an evaluator who practices reflection, or who practices evaluation dialogically, requires skill sets that draw from practical, real-world experience and integrate theoretical concepts, as noted above by Christie and Rose (2003). This reflective and dialogic skill set further includes being able to recognise evidentiary patterns, frame context, understand conflict and tension among stakeholder groups, and act in response to the contextual environment (Archibald, Neubauer, & Brookfield, 2018).

Dialogue inherently introduces multiple perspectives—based on either group or individual identity—to the evaluation process, the interaction of which is often dictated by power relationships. Hood, Hopson, and Kirkhart (2015) consider culturally responsive evaluation (CRE) a paradigm that merges theories of culturally responsive assessment and responsive evaluation, with the particular intention of giving attention to marginalised groups. Hood, Hopson, and

Kirkhart (2015) cite Frierson, Hood, Hughes & Thomas (2010) definition of "culture" as a "cumulative body of learned and shared behavior, values, customs and beliefs common to a particular group or society," while "being responsive" refers to acting meaningfully and politically in relation to issues of culture and race (Hood, 2001). Cultural competence is at the core of the CRE practice (Hood, Hopson, & Kirkhart, 2015), and it is+ integrated into the evaluation process, from engaging stakeholders, to framing the question(s), to disseminating the results.

In this dialogic approach to CRE, we propose that culturally responsive evaluators find ways to create discussion about critical elements of processes and procedures—in this case, regarding the PT. For instance, where Hood, Hopson, & Kirkhart (2015) identify engaging stakeholders as a key set of initial steps in their CRE framework, they highlight the need to engage stakeholders to establish and model climates of trust, respect, and participation. Engaging stakeholders with questions and dialogue is fundamental to this process and builds relationships and understanding of the points of view of the stakeholders—in this case, the NPTA staff most involved in scoping and engaging with the evaluation team. The intersection between a culturally responsive and a dialogic approach is illustrated by this engaging-stakeholders process and will be elucidated in the sections that follow.

In the next section, the evaluation project that birthed the PT will be detailed. We lay out the purpose of the evaluation, and the key players involved, and describe the programme. We will then detail the strategy for engagement with stakeholders and the importance of dialogue through communication and relationships over the course of this project.

## *Evaluation of the NPTA's STEM + Families programme*

In 2017, two of the authors of the present article were requested by the NPTA to i) support its efforts to evaluate programme effectiveness around the organisation's key pillars—advocacy, membership, and diversity—and ii) support the development of its internal capacity to evaluate its programmes. The organisation's programme leadership requested that the focus of the evaluation be placed on two programmes in particular, of the many different programmes the NPTA offers: Reflections1 and STEM + Families. STEM + Families seeks to increase access to, interest in, and understanding of STEM careers and education pathways for all children, and especially those most under-represented in the STEM field: girls, children of colour, and children living in low-socioeconomic communities. The programme aims to work toward this goal by providing STEM events at participating schools, exploring ways to engage families, and working with partner organisations to improve STEM learning environments within schools and communities. In our writing, we focus on the PT development that took place during the STEM + Families programme evaluation process, as it was through that process that challenges primarily arose.

The evaluation team, made up of the two authors (who were university faculty members) and three graduate students, developed a memorandum of understanding with the NPTA. In it the evaluation team outlined several activities that were initially expected for a 2-year process: framing and developing a project timeline; developing a PT model for the two programmes that made up the NPTA Partnerships

---

1    For nearly 50 years, the National PTA Reflections Program is a competition that has encouraged students of all ages to explore and be involved in the arts. Through the programme, the National PTA and 4,000 PTAs across the country support all students with all abilities to create and submit original works of art in the medium of their choice—dance choreography, film production, literature, music composition, photography or visual arts—that reflect on the annual theme. There also is a special artist division to provide all students—including those with disability—the opportunity to participate in the programme.

and Programs Directorate; identifying and assessing strategic needs for the evaluation; training staff, and developing staff evaluation-training capacity; and developing actionable and utilisation-focused recommendations. In practice, some of the ideas developed at proposal were not realised fully in the process of conducting the evaluation. Namely, plans to develop evaluation capacity were not realised among the NPTA team, as timelines grew tighter and data collection became the primary focus..

While efforts were made by the evaluation team to share resources and measures with staff, the purpose of this article is not to present the key learnings from working with staff. Nor does it describe the specific processes employed with the NPTA Reflections Program PT. Rather, this article explores the ways the evaluation team developed dialogue through its culturally responsive PT change processes. The article reveals the key challenges and learnings for the evaluation team, and presents possibilities for thinking through how evaluators might effectively and thoughtfully integrate CRE processes within PT, and how evaluators might think through PT in culturally responsive ways.

## *PT-model development*

The design of the model drew on programmatic expertise from the NPTA leadership. The evaluation team sought to include key parties at every stage of the evaluation, including the PT-development stage. To develop the PT, we used several cycles of a four-step process (see Figure 1):

1. Engage with stakeholders.
2. Reflect on engagement.
3. Translate reflections into model content.
4. Present model to stakeholders.
5. This section describes each of the steps, including discussions of how they were implemented in practice. Key relational challenges that emerged during this process will be detailed and learnings shared.
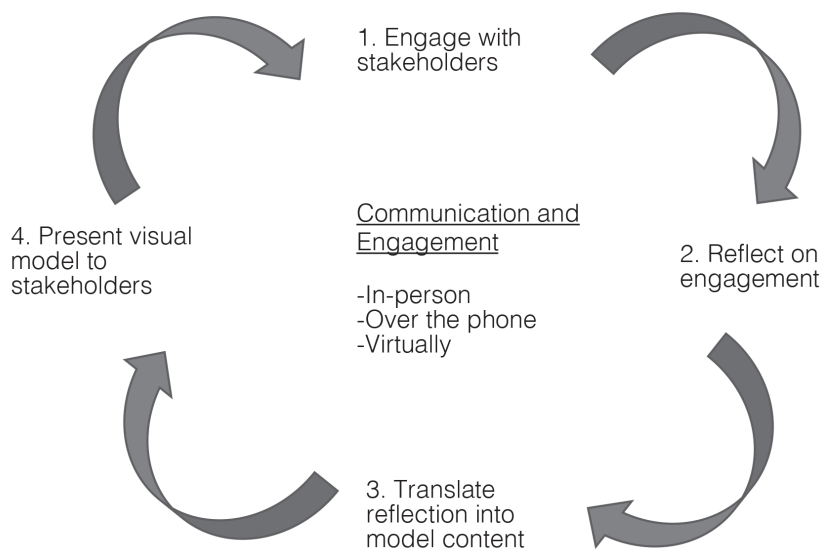
1. Engage with stakeholders

Communication and Engagement

-In-person
-Over the phone
-Virtually

2. Reflect on engagement

3. Translate reflection into model content

4. Present visual model to stakeholders

Figure 1. PT stakeholder engagement strategy cycle

## 1. Engage with stakeholders

The PT development was one of the initial tasks in the evaluation contract. Early in the planning process, the evaluation team presented at the NPTA directorate retreat to address key questions on the evaluation process—including those regarding the options for PT models, such as logic models and theories of action or change—and gauge the NPTA's familiarity with evaluative thinking, programme theories, and other potential ways to build collaborative capacity among the NPTA team.

Stakeholder engagement and dialogue were a priority in the PT-development process. Understanding NPTA leadership relationships and the environment in which they worked was an important aspect of this process. It was through a CRE lens that this task was

undertaken, ensuring that diverse viewpoints were respected and represented through purposeful engagement strategies. The stakeholders who played a critical role in the stakeholder-engagement process were the director and staff responsible for the two PTA programmes being evaluated.

Engagement between the evaluation and stakeholder groups involved frequent communication and dialogue about aspects of the evaluation. Communication was conducted virtually, in person, and over the phone (see Table 1 for further details on each cycle of the engagement strategy).

Throughout the evaluation, regular conversations emerged about timelines for carrying out procedures and processes. Meetings at times involved key NPTA staff who were not necessarily involved directly with the evaluation, but who were asked to participate to align key organisational strategic goals and to help clarify outputs and outcomes that were being articulated by key staff. The evaluation team asked questions and further explored the needs of these stakeholders. This communication was reciprocated, as the stakeholders also sought to understand certain aspects of the evaluation process.

## 2. Reflect on engagement

The evaluation team meetings and the anticipated data-collection processes created opportunities to reflect on the type of engagement that was developing, and its role in shaping the PT model. The notes and discussion that were part of the reflection process were used to develop key components of a visual representation of the PT model (see Figure 2).

We were learning very quickly, at least in the case of the STEM + Families programme, that few or no measures or data describing how the programme was being delivered were being collected at national, state, or local levels. We faced a deep challenge in making

sense of what did not exist and how existing measures aligned with our understanding of the PT, and in developing measures in a short time frame.

Additionally, and in the case of the STEM + Families programme, there were corporate interests that were not well integrated into the programme design and its implementation. One of our key tasks lay in ensuring that we knew which audiences were most interested in the overall programme, and had a stake in its evaluation, both within the national office and among the multiple variations of state and other local contributory programmes being implemented.

A key part of the process in the beginning was engaging in exploratory dialogue, in which the evaluation and stakeholder teams asked questions, engaged, reflected, and brought forth ideas. This was part of developing an initial understanding of the key components that underlie the two programmes, to make sense of and advance evaluative thinking and understanding. Schwandt (2001) recognises the practical (more than technical) aspects of evaluation, in terms of making sense and reconstructing details and aspects of the programme: "[A]ction-oriented self-understanding of what it means to do evaluation well is informed by evidence and reason but it is not governed (or governable) by technical rules and methodological prescriptions" (p. 230).

Figure 2 was one of the first iterations of the STEM + Families PT model. The initial iterations of the PT model had evident gaps in how different stakeholder groups moved through the PT logic (see Figure 2). These gaps were discussed and filled through discussion and feedback with the evaluation team and between the evaluation team and the NPTA leadership team.
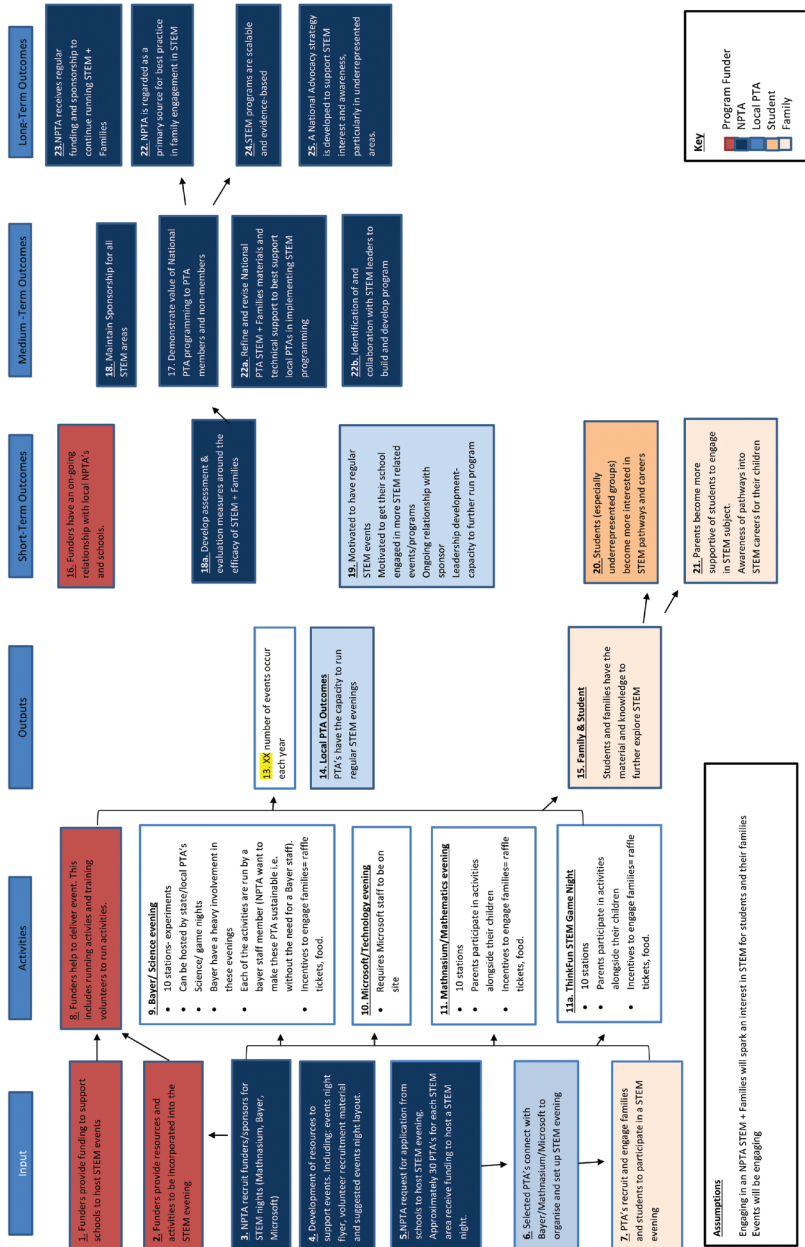
Figure 2. Third iteration of STEM + Families PT

The reflection process initially involved the first author of this article leading the development of the model, the larger evaluation, and NPTA programme teams through dialogue and reflection to aid in understanding of the programme's PT. Reflection involved assessing the data provided by the NPTA leadership, which included feedback, written or verbal, as well as observations made by the evaluation team.

Oftentimes, our efforts at stakeholder engagement included navigating the practical realities the NPTA programme leadership team faced when trying to integrate evaluation into existing programmes and strategic realities, such as programme timelines, funder expectations, and high workloads. These tensions are explored more in the "Challenges, Questioning, and Reflection" section.

## 3. Using reflections from stakeholder engagement to build the PT

Using stakeholder knowledge and feedback to develop the PT involved reflection from the evaluators and dialogic skills. Stakeholder groups were specifically asked to clarify sections of the PT, where necessary. This clarification occurred throughout the PT development, until the model was presented in person in the fourth cycle of the stakeholder-engagement strategy (see Figure 3).

# NPTA STEM+Families Program Theory of Change

**Input**

Program Sponsors provide:
- Personnel to support the delivery of each event
- STEM focused activities and resources
- Funding

NPTA recruit funders/sponsors for STEM nights (Mathasium, Bayer, Microsoft)

NPTA promote program:
- Communicating with state PTA's
- Use of media platforms

PTA's apply for funding to host STEM event.

NPTA select a limited number of PTAs to receive grants. Selection criteria: % are Title 1 schools, potential to add value to the community

NPTA communicate with PTAs:
- Contract and requirements made explicit
- 75% of funds before event
- 25% after event subject to meeting requirements
- NPTA connect PTA with local STEM sponsor organisation to co-ordinate planning and support
- NPTA develop and provide PTA's with resources to support organisation and delivery of events

PTAs identify additional school and community resources to include in program

**Activities**

Sponsors support the delivery of event:
- Running activities and/or training volunteers to run activities.

STEM+ Families event:
- Content and activities are determined in partnership between the local PTA and the sponsor organisation.
- Activities are parent, student and family inclusive.
- Incentives are provided to engage families

A₂

PTAs recruit and engage families and students to participate in a STEM event

PTAs collaborate with Sponsor and school community to organise local event

A₁

**Outputs**

Students, Families, schools and the community are exposed to the sponsor organisations

Evaluation data is collected during and after the STEM event

XX number of events occur each year

All grant recipients have the necessary resources and support to run host a STEM evening.

All grant recipients host a STEM evening.

Attending students are diverse, including female students and students from under-represented groups

All students and families receive materials that will aid in their further engagement with STEM

Students and Families attend and engage with the STEM evening. Demographic of families are representative of the demographic of students at the school

**Short-Term Outcomes**

Sponsors gain positive media coverage

NPTA gain positive media coverage from the events that position them in a positive light to deliver STEM programs

Develop assessment & evaluation measures around the efficacy of STEM + Families

Refine and revise National PTA STEM+Families materials and technical support to best support local PTAs in implementing STEM programming

PTAs have the capacity to run regular STEM evenings

Motivated to have regular STEM events & Motivated to get their school engaged in more STEM related events/programs

Students (including underrepresented groups) become more interested in STEM pathways and careers

Students and families have the material and knowledge to further explore STEM

**Medium-Term Outcomes**

Mathasium has an ongoing school-franchisee relationship

ThinkFun achieve an increase in sales.

Bayer receivers positive press around their involvement in the program

Maintain Sponsorship for all STEM areas

Demonstrate value of National PTA programming to PTA members and non-members

STEM programs are scalable and evidence based

PTAs and schools actively seek to STEM related programs/opportunities for students to engage with.

Students and families (including underrepresented groups) become more engaged in STEM pathways and careers

Parents become more supportive of students to engage in STEM subject. Awareness of pathways into STEM careers for their children

**Long-Term Outcomes**

NPTA receives regular funding and sponsorship to continue running STEM + Families

NPTA is regarded as a primary source for best practice in family engagement in STEM

A National Advocacy strategy is developed to support STEM interest and awareness, particularly in underrepresented areas.

Students seek careers and pathways in STEM, closing STEM gap

**Key**

Program Funder
NPTA
Local PTA
Student
Family

**Assumptions:**

A₁ There are other additional school and community resources that can be included in the program.

A₂ Engaging families in the program will support students to participate and be interested in STEM activities (White paper, pg. 3)

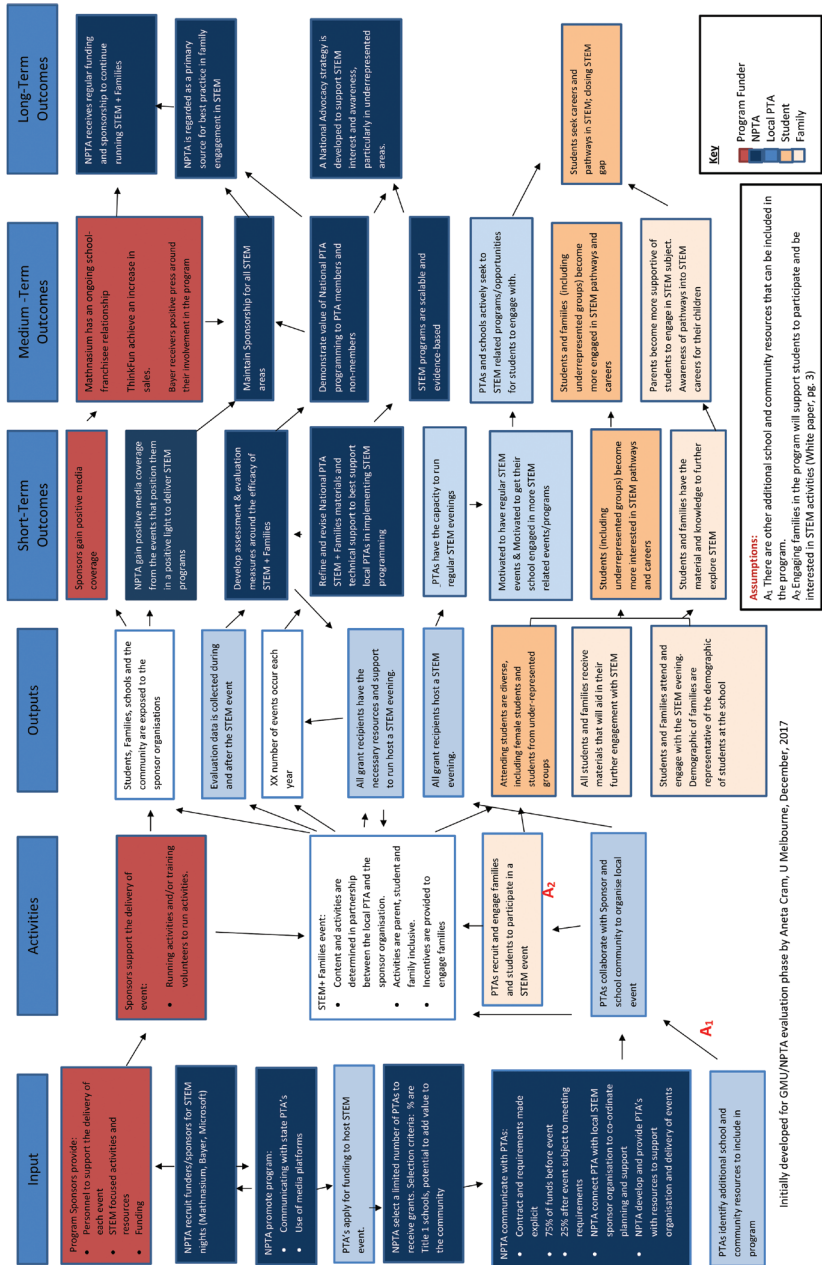Initially developed for GMU/NPTA evaluation phase by Aneta Cram, U Melbourne, December, 2017

Figure 3. Seventh iteration of the STEM + Families PT

Representation was key. The evaluation team specifically chose to represent the different stakeholder groups using colour on the diagram and to make links between model components explicit. This coding ensured clarity when components or links were missing in the logic of the PT (see Figure 2). For instance, the funder or corporate sponsors were represented by the colour red (see Figure 3). During an in-person meeting, gaps were identified. It was unclear what the programme was doing to achieve the funders' desired outcomes, which prompted further discussion on the level of corporate involvement and corporate interest in the programme.

Stakeholder feedback at times was communicated in person, over the phone, or in writing. Reflections and stakeholder feedback were incorporated into the next version of the PT model or were highlighted as areas in need of further discussion.

A step in this process initially tasked key programme stakeholders with the development of a document detailing the components or logic of the programme. From this, the evaluation team met and engaged with team leaders from the NPTA team. This activity centered the NPTA leadership team as experts on their programmes and promoted a relationship between the evaluation team and the key stakeholders that was based on communication and reciprocity (Hood, Hopson, & Kirkhart, 2015; Quinn-Patton, 2008).

## 4. Present models to stakeholders

Sharing versions of the PT as they were developed ensured that the key stakeholders were given ample opportunity to share their perspectives on the programme and recognise their understanding in each revision of the visual PT model. This strategy also enabled stakeholders to challenge and question the model, present different perspectives, and share their conception of the programme's PT (Christie & Rose, 2003; Hansen & Vedung, 2010).

Following the completion of the initial PT model, the model was sent to the stakeholder group and further feedback sought. The draft model was coupled with guiding questions about any uncertainty during the development of the model, aspects stakeholders should specifically focus on, and specific elements in the model, to support interpretation.

Early versions of the model were shared with key stakeholder groups virtually. The seventh version was shared in person on a poster-sized sheet (see Figure 3). The intention of presentation during the fifth cycle was to further examine the causal relationship between the elements in the model, as there was still ambiguity around the relationship between inputs and outcomes for certain key players and stakeholders in the programme.

Table 1 details how each stage of stakeholder engagement was implemented in practice over the course of model development by cycle.

Table 1. Stages of stakeholder engagement to develop the program theory

| Cycle | Intention | Implementation/In Practice |
|---|---|---|
| 1 | Initial presentation of ideas | • The evaluation team was invited to be involved in the NPTA leadership retreat to present overall ideas for evaluation, including an introduction to PT models.<br>• Examples of different forms of graphic conceptual models were shared, including a theory of action, theory of the problem, and outcomes model.<br>• Stakeholders were asked to develop a document that included their understanding of the inputs and outcomes of the program. The document was then shared with the evaluation team, who talked through each of the components. A recording was taken. This formed the basis for the initial draft of the PT model. |
| 2 | Reflecting on ideas | • The evaluation team was provided with the initial draft for their review.<br>• The draft was reflected on; feedback was shared in dialogue.<br>• The second draft was formed from this feedback. The second draft was shared with the NPTA leadership team, with a detailed supporting document explaining from where the content was derived. |
| 3 | Reflecting ideas | • Feedback was shared over the phone, during a regular team meeting.<br>• Feedback was received verbally from key stakeholders.<br>• Feedback was incorporated into the third version of model. |
| 4 | Finalizing model and documentation | • Version 3 was sent to the NPTA leadership team. A resource sheet of sources of the elements of the PT was shared along with the model. Guiding questions were included to support understanding and identify uncertainty in the interpretation of the model.<br>• Draft was presented in person to the team; opportunity was given to discuss the model and the pathways. The meeting took one hour, with the PT development process taking 40 minutes of this process. The meeting was audio recorded.<br>• Feedback from key stakeholders was also shared in the document via email.<br>• Feedback was taken and incorporated into the PT model |
| 5 | Finalizing model and documentation | • The final draft of the PT model was shared with the key stakeholder group via email. |

## Challenges, questioning, and reflection

In this section, the challenges that emerged during the PT-development process will be detailed. These challenges are not in order of occurrence; some of the challenges happened simultaneously or were

revisited when other challenges surfaced. Although communication and dialogue were made a priority by the evaluation team, there were breakdowns in understanding along the way. Ultimately, the evaluation team used dialogue and reflection to address challenges and questions from NPTA staff during the PT development. This dialogic method is featured in this section as we discuss three illustrative challenges that surfaced during PT development.

Three tensions or challenges emerged during the NPTA PT-building process: i) definition of "family engagement," ii) tensions regarding programme-sponsor expectations, and iii) questions and expectations about the role of the evaluator. An in-person meeting between the evaluation and key-stakeholder teams was recorded. Extracts from that recording have been used to support statements made in this section.

## Challenge 1: Definition of "family engagement"

Family engagement was a core component of the STEM + Families programme. However, during the first three cycles of the four-step engagement-strategy process, this was not drawn out in any of the conversations on core components or captured in how the NPTA leadership team described the programme. During the reflective dialogue, NPTA staff statements showed they were perplexed to notice that the visual presentation of the programme lacked any inputs, outputs, or outcomes related to family engagement (see Figure 2). This was a key learning area for them, as key staff reflected on the strategies that they were using to engage families, recognising that they needed to revisit this component of the programme. The NPTA staff address the initial concern in dialogue:

> I'm trying to see if I can capture this somewhere in this theory-of-change document . . . The family-engagement element was really how this programme was designed in theory and that might be a little bit light in what I'm looking at . . . For every National PTA

programme focused on family engagement, we could use them as a model. And so, if we are using them as a model you would want your [PT] to reflect that in its model design. And I just feel like it's a little bit light.

As we would learn through the same meeting, NPTA staff were meeting with the director of family engagement to both define the term and figure out what it looks like for the STEM + Families programme. While they realised the evolving nature of the family-engagement notion in practice, the PT-development process brought to light the need to define and operationalise key terms relevant to the evaluation and programme-development process. Additionally, this process highlighted the need to be cognisant of the academic–practice gap in relation to definitions of concepts. As the theory was developing, there were opportunities for the evaluation team to probe further into the gaps present in the model, which, in hindsight, should have meant that the absence of the family component to the programme be addressed by the evaluation team earlier in the PT-development process.

This lack of a definition highlighted engagement challenges, and, for one of the authors, it highlighted the absence of research-informed practice, as there was a lack of clarity on how families were being engaged on the ground and what the purpose of family engagement was for the overall delivery of the programme.

## Challenge 2: Tensions regarding programme sponsors' expectations

STEM + Families is a programme sponsored by a number of notable STEM companies in the United States. One of the tensions that arose centered around corporate interests and how they were represented in the PT. As the PT evolved, colour codes marked in the legend of the model identified which stakeholders had interest in which outcomes. For instance, as we learned more about the anticipated outcomes, we

realised that corporate stakeholders were interested in specific measures. For example, regarding who attended STEM nights and why, there was a lack of clarity around how funders contributed to STEM nights and what outcomes they wanted to achieve. One person stated:

> I just think that it is really interesting that there is no connection between any of these, as you pointed out. I mean it is interesting that we throw outcomes to the wall. It sounds really good. But we have no idea if these are being realised, so it's just a general observation about how a lot of it is based on what our sponsors want and not based on what is realistic and what is the real change. Like what's the real change that happens because of all this great stuff?

Miller (2018) notes (in the context of vulnerable and disenfranchised groups in the work of community-based programmes) that the role of evaluators is to facilitate tensions that exist in the complex ecologies of our work. Although the focus of the programmes Miller refers to is different, she too explores the complexity of multi-site programmes—in her case, around evidence-based HIV behavioral interventions—at multiple levels between states and localities, and the high-stakes nature of these interventions' ability to influence real social changes on a broad scale.

Even in the case of the NPTA evaluation, the statement quoted here and subsequent dialogue revealed tensions regarding sponsor expectations and what staff did or did not know, including the programme staff's own goals for the programme. The dialogue around the PT development triggered discussion of future data-collection efforts, including ways to ensure NPTA's and current (or future) corporate sponsors' needs were being met. Providing the key staff with a visual document that identified gaps in the programme logic drew out challenges that they faced about their funders' role in setting the direction of the programme. Having an in-person meeting allowed for questions and challenges to flow between the stakeholders and evaluation teams.

## Challenge 3: Challenges and questioning about evaluator role and expectation

> That's one of the biggest questions that I have looking at this model. I didn't know where to locate [STEM + Families as a model for best practices in providing engagement]. So maybe I need your expertise in doing it or the evaluation team. 'Cause *we don't know what your role is* [emphasis in original recording].

This quote reflects the curiosity and frustration displayed by one NPTA staff member who responded to the seventh iteration of the PT, when the PT was shared in a large group with evaluators and NPTA staff for the first time. Although the staff member refers to the location of an outcome on the PT, the quote largely reflects a questioning of the evaluator role and a need to clarify what our dialogic, culturally responsive process was intended to document. This role was not always welcome and sometimes sparked frustration from staff members about the details and levels of discussion needed during the PT development.

Further, the staff member's emphatic statement, "we don't know what your role is," was made in the group after a lengthy conversation between both parties about assessing and confirming whether the family-engagement strategies that the NPTA was employing could be confirmed as best practice. It was stated in a way that suggested the staff member was speaking on behalf of the team. Although the conversation was seemingly about best practices being documented, there were more fundamental questions about our role as evaluators, and the expectations that accompanied that role.

Ryan and Schwandt (2002) document how the role of the evaluator, central to both theory and practice, is ubiquitous, contentious, and value laden. Straddling a social-constructivist and critical social-scientist perspective, we recognised the inherent tendency of NPTA staff to attempt to understand the meanings of the evaluation,

best practices, and theories they were building to craft interventions and innovations (social constructivist), while also trying to understand the larger power dynamics underlying our evaluation exercise and the strategic priorities of the organisation, as well as the subtle power issues among staff members at the table.

We had to negotiate our evaluator role and expectations of that role not only at the outset of the project, when we established a memorandum of understanding as good evaluation practice (Yarbrough, Shulha, Hopson, & Carutehrs, 2011), but also during the development of the PT.

## *Key learnings and takeaways*

In this final section, we focus on the key learnings and takeaways from our reflective practice of applying culturally responsive evaluation processes, through dialogue and engagement, in the development of a PT. This article not only drew out the challenges that came with the development of a PT for a specific project, but also detailed how we developed this specific model, the theoretical frameworks that we were aligned with, and our reflections on that process. The key learnings and takeaways are shared with the intention of providing insight on a real-world example of PT development, the challenges faced, and, most importantly, the implications for the development of evaluators and evaluation skill sets.

### PT as integral to the evaluation process

As authors, we realise that PT development is an integral part of the evaluation process, and navigating stakeholder relationships is part of that process. Relational challenges can arise. Reflecting on these and using dialogic skills to recognise and mitigate them is, at times, part of the evaluation process, and central to the role of an evaluator. Recognising and highlighting challenges in evaluation can also provide opportunities for reflection and improvement. Our experiences

and learnings with this evaluation project reinforce the importance of establishing clear evaluation expectations and clear evaluator–client roles, and the importance of interpersonal competences (Garcia & Stevahn, 2019).

The development of a PT adds value to evaluation. It is a tool to elucidate the constructs around a programme and provide a focal point to draw stakeholders into the evaluation process. From there, programme components can be explored, developed, or assessed, depending on the framing of the evaluation. Convening stakeholder groups can provide an opportunity to draw out stakeholders' implicit understandings of a programme (Rossi, Freeman, & Lipsey, 1999) and can support stakeholder buy-in (Clinton, 2014).

When done well, PT can produce many benefits. The process of developing the theory can build a shared understanding among diverse stakeholder groups, or help to identify legitimate conflicting perspectives that require further exploration as part of the evaluation process (Funnell & Rogers, 2011). PT, when done poorly, can misrepresent programme activities and intentions, affecting the validity of wider evaluation findings. Our learnings and takeaways reinforce the thinking that the developmental nature of these PT processes requires navigating different sources of knowledge and merging them to gain an accurate, or at least near accurate, representation of a programme and a programme's intended outcomes (Funnell & Rogers, 2011).

## Benefits of the CRE lens in PT development

We aligned our thinking with, and realised the benefits of learning from, evaluation theorists who promote the importance of including multiple perspectives, and who highlight the value of involving key programme stakeholders in the design of a PT (Greene, 2007; Stake, 1975). These authors recommend that diverse stakeholders also be

engaged in general data collection to support buy-in, understand their informational needs, and develop strong stakeholder–evaluator relationships. Furthermore, since the introduction of multiple perspectives could introduce a myriad of political, religious, ethnic, linguistic, and racial groups, it is important for evaluators to be culturally responsive (Frierson, Hood, Hughes, & Thomas, 2010).

Hood, Hopson, and Kirkhart (2015) identify the engagement of stakeholders as a key set of initial steps in their CRE framework. They highlight the need to engage stakeholders to establish and model climates of trust, respect, and participation. We found that conversation and dialogue were integral to unearthing stakeholder concepts and understandings, which were sometimes fragmented and required translation and re-articulation for better understanding for the purpose of the evaluation (Hansen & Vedung, 2010; Rossi, Freeman, & Lipsey, 1999). We relied on Greene's (2001) notion of dialogue as a tool to support this PT process. That notion highlights a relational perspective of dialogue that attends to i) who is or is not participating in evaluative communications, ii) social relations, and iii) the democratising potential of evaluative communications. The fundamentally relational and communicative aspects of dialogue are inextricably linked to "moral-ethical and politicised power relationships among stakeholders—about power and voice, agency and moral purpose, caring and empathy, understanding and acceptance of difference" (Greene, 2001, p. 182). Essentially, dialogue in evaluative communications reveals important relationships.

## Importance of recognising and navigating power dynamics in PT development

Additionally, one of our key takeaways related to recognising and navigating tension and power dynamics, which requires interpersonal, culturally responsive, and reflective skills. Being able to

identify power dynamics, and the cause of these dynamics, and to navigate, reflect, and question is imperative to quality evaluation practice. In relying on notions of relationships as an important part of stakeholder engagement, dialogue, and reflection, this article posited that relationships are fundamental to understanding the role of evaluation in complex ecologies (Hopson & Cram, 2018).

During PT workshops, differing viewpoints and understandings of the PT can arise. We found that relational tensions arose during the evaluation process and were influenced by a variety of contextual and interpersonal factors. In our case, these tensions were due to a need to be more intentional and specific in our dialogic process throughout the evaluation. This was clear with Challenge 3, when there was uncertainty around the roles and responsibility of the evaluator. Being clear and employing a culturally responsive skill set are valuable strategies for evaluators in ensuring that all parties have a clear understanding of evaluation and the parameters and expectations of the evaluator–client/stakeholder relationship (Stickl Haugen & Chouinard, 2019).

We found that there were power dynamics being played out in the NPTA leadership team that inhibited the participation of all team members. It was also apparent in the ties that the NPTA leadership had to the programme funders. Stakeholders felt that this dynamic needed to change. More developed skills in group facilitation and interpersonal engagement would have enabled the evaluation team to tap into diverse stakeholder perspectives and ensure all perspectives were heard.

## Key aspects of evaluator competencies in PT development

A final key learning relates to our understanding of the role of key competencies and our experiences with this project.

Stevahn, King, Ghere, and Minnema (2004) brought wide interest in the competencies needed to be a competent evaluator with the development of the Essential Competencies for Program Evaluators Self-assessment text. This text outlined the skills and personal qualities necessary to be effective in practicing evaluation. Interpersonal competence and situational awareness were focused on, with recent research being conducted by Garcia and Stevahn on its importance (2019). Various studies have been conducted since, by a variety of respected authors and scholars, exploring different facets of evaluator competencies and what is needed for the development of the field and the emerging evaluators that come into it (Galport & Azzam, 2017). Our experiences would suggest that problem-solving skills, interpersonal skills, and strong dialogic skills are necessary to build rapport and trust with stakeholder groups and to be able to navigate relational challenges.

Cram and Mertens (2015) noted that to have power is to own the decision making about whose knowledge or what knowledge counts. They stated that the challenge for researchers and evaluators working in the space of social justice and equity is to be attentive to whose knowledge is being privileged within the evaluation context, how that plays out between stakeholder groups, and when and how that is recognised and traversed. This may prove challenging, particularly with a large and multi-site programme (Hansen & Vedung, 2010).

Karen Kirkhart states (as cited in Cram and Mertens, 2015) that culture shapes evaluation theory and is also affected by it. Our experience suggests that the key competencies required of evaluators when developing PTs are the ability to recognise the contextual environment and recognise the culture of that environment. Drawing out aspects that might be a point of tension requires skills in management and exploration, especially if the aspect of tension is imperative to understanding the PT.

Although a culturally responsive approach was used in implementing the evaluation, there were still barriers to understanding what caused challenges among the stakeholder groups and between the evaluation and stakeholder groups. Building interpersonal and problem-solving skills that allow challenges to be brought forth during a workshop, worked through, and resolved, on reflection, could improve this process.

## References

Abma, T. A. (2006). The practice and politics of responsive evaluation. *American Journal of Evaluation*, *27*(1), 31–43. https://doi.org/10.1177/1098214005283189

Abma, T. A., Leyerzapf, H., & Landeweer, E. (2017). Responsive evaluation in the interference zone between system and lifeworld. *American Journal of Evaluation*, *38*(4), 507–520. https://doi.org/10.1177/1098214016667211

Archibald, T., Neubauer, L. C., & Brookfield, S. D. (2018). The critically reflective evaluator: Adult education's contributions to evaluation for social justice. *New Directions for Evaluation*, 2018(158), 109–123. https://doi.org/10.1002/ev.20319

Astbury, B., & Leeuw, F. L. (2010). Unpacking black boxes: Mechanisms and theory building in evaluation. *American Journal of Evaluation*, *31*(3), 363–381. https://doi.org/10.1177/1098214010371972

Bakhtin, M. M. (1981). *The dialogic imagination: Four essays* (M. Holquist, Ed., C. Emerson & M. Holquist, Trans.). Austin, TX: University of Texas Press.

Bohm, D. (1996). *On dialogue* (L. Nichol, Ed.). London, UK: Routledge.

Buber, M. (1961). *Between man and man* (R. Gregor Smith, Trans.). London, UK: Collins.

Christie, C. A., & Rose, M. (2003). Learning about evaluation through dialogue: Lessons from an informal discussion group. *American Journal of*

*Evaluation*, 24(2), 235–243. https://doi.org/10.1177/109821400302400207

Clinton, J. (2014). The true impact of evaluation: Motivation for ECB. *American Journal of Evaluation*, 35(1), 120–127. https://doi.org/10.1177/1098214013499602

Cram, F., & Mertens, D. (2015). Transformative and indigenous frameworks for multimethod and mixed methods research. In S. Nagy Hesse-Biber & R. Burke Johnson (Eds.), *The Oxford handbook of multimethod and mixed methods research inquiry* (pp. 99–109). Oxford, UK: Oxford University Press. https://doi.org/10.1093/oxfordhb/9780199933624.013.7

Funnell, S. C., & Rogers, P. J. (2011). *Purposeful program theory: Effective use of theories of change and logic models*. San Francisco, CA: Jossey-Bass.

Frierson, H. T., Hood, S., & Hughes, G. (2002). Strategies that address culturally responsive evaluation. In *The 2002 user-friendly handbook for project evaluation* (pp. 63–73). Alexandria, VA: National Science Foundation. https://www.nsf.gov/pubs/2002/nsf02057/nsf02057.pdf

Frierson, H. T., Hood, S., Hughes, G. B., & Thomas, V. G. (2010). A guide to conducting culturally responsive evaluations. In *The 2010 user-friendly handbook for project evaluation* (pp. 75–96). Alexandria, VA: National Science Foundation. https://www.purdue.edu/research/docs/pdf/2010NSFuser-friendlyhandbookforprojectevaluation.pdf

Galport, N., & Azzam, T. (2017). Evaluator training needs and competencies: A gap analysis. *American Journal of Evaluation*, *38*(1), 80–100. https://doi.org/10.1177/1098214016643183

Garcia, G. L., & Stevahn, L. (2019). Situational awareness and interpersonal competence as evaluator competencies. *American Journal of Evaluation*. Advance online publication. https://doi.org/10.1177/1098214018814941

Greene, J. C. (2001). Dialogue in evaluation: A relational perspective. *Evaluation*, *7*(2), 181–187. https://doi.org/10.1177/135638900100700203

Greene, J. C. (2007). *Mixed methods in social inquiry*. San Francisco, CA: Jossey-Bass.

Habermas, J. (1984–7). *The theory of communicative action* (2 vols). London, UK: Heinemann.

Habermas, J. (2003). *Truth and justification*. Cambridge, MA: The MIT Press.

Hansen, M. B., & Vedung, E. (2010). Theory-based stakeholder evaluation. *American Journal of Evaluation*, *31*(3), 295–313. https://doi.org/10.1177/1098214010366174

Hood, S. (2001). Nobody knows my name: In praise of African American evaluators who were responsive. *New Directions for Evaluation*, 2001(92), 31–44. https://doi.org/10.1002/ev.33

Hood, S., Hopson, R., & Frierson, H. (Eds.). (2015). *Continuing the journey to reposition culture and cultural context in evaluation theory and practice.* Charlotte, NC: Information Age Publishing.

Hood, S., Hopson, R. K., & Kirkhart, K. E. (2015). *Culturally responsive evaluation*. In K. E. Newcomer, H. P. Hatry, & J. S. Wholey (Eds.), *Handbook of practical programme evaluation* (pp. 281–317). http://doi.org/10.1002/9781119171386.ch12

Hopson, R., & Cram, F. (Eds.). (2018). *Tackling wicked programs in complex ecologies: The role of evaluation.* Palo Alto, CA: Stanford University Press.

House, E. R., & Howe, K. R. (1999). *Values in evaluation and social research.* Thousand Oaks, CA: Sage.

House, E. R., & Howe, K. R. (2000). Deliberative democratic evaluation. In K. E. Ryan & L. DeStefano (Eds.), *Evaluation as a democratic process: Promoting inclusion, dialogue, and deliberation* (pp. 3–12). San Francisco, CA: Jossey-Bass. https://doi.org/10.1002/ev.1157

Mark, M. M. (2000). On the role of dialog in evaluation. *American Journal of Evaluation*, *21*(2), 205–206. https://doi.org/10.1177/109821400002100207

McKinnon, M. C., & Hole, D. G. (2015). Exploring program theory to enhance monitoring and evaluation in ecosystem-based adaptation

projects. *New Directions for Evaluation*, *147*, 49–60. https://doi.org/10.1002/ev.20130

Miller, R. M. (2018). Evaluating HIV practices and evidence-supported programs in AIDS community-based organizations. In R. Hopson, & F. Cram (Eds.), *Tackling wicked programs in complex ecologies: The role of evaluation* (pp. 69–102). Palo Alto, CA: Stanford University Press.

Quinn-Patton, M. (2008). *Utilization-focused evaluation*. Thousand Oaks, CA: Sage.

Rallis, S. F., & Rossman, G. B. (2000). Dialogue for learning: Evaluator as critical friend. In R. Hopson (Ed.), *How and why language matters in evaluation* (pp. 81–92). San Francisco, CA: Jossey-Bass. https://doi.org/10.1002/ev.1174

Renger, R., & Bartel, G. (2013). The reciprocal relationship between implementation theory and program theory in assisting program design and decision-making. *The Canadian Journal of Program Evaluation*, *28*(1), 27.

Rossi, P. H., Freeman, H. E., & Lipsey, M. (1999). *Evaluation: A systematic approach* (6th ed). Thousand Oaks, CA: Sage.

Ryan, K. E., & DeStefano, L. (Eds.). (2000). *Evaluation as a democratic process: Promoting inclusion, dialogue, and deliberation*. San Francisco, CA: Jossey-Bass.

Ryan, K. E., & Schwandt, T. A. (2002). *Exploring evaluator role and identity*. Charlotte, NC: Information Age Publishing.

Schwandt, T. A. (2001). Understanding dialogue as practice. *Evaluation, 7*(2), 228–237. https://doi.org/10.1177/13563890122209658

Stake, R. E. (1975). An interview with Robert Stake on responsive evaluation. In R. E. Stake (Ed.), *Evaluating the arts in education: A responsive approach* (pp. 33–38). Columbus, OH: Merrill.

Stevahn, L., King, J., Ghere, G., & Minnema, J. (2004). *Essential competencies for program evaluators self-assessment*. Minneapolis, MN: Minnesota Evaluation Studies Institute. Retrieved from https://www.cehd.umn.edu/olpd/MESI/resources/ECPESelfAssessmentInstrument709.pf

Stickl Haugen, J., & Chouinard, J. A. (2019). Transparent, translucent, opaque: Exploring the dimensions of power in culturally responsive evaluation contexts. *American Journal of Evaluation*, *40*(3), 376–394. https://doi.org/10.1177/1098214018796342

Widdershoven, G. A. (2001). Dialogue in evaluation: A hermeneutic perspective. *Evaluation*, *7*(2), 253–263. https://doi.org/10.1177/13563890122209676

Yarbrough, D. B., Shulha, L. M., Hopson, R. K., & Carutehrs, F. A. (2011). *The program evaluation standards: A guide for evaluators and evaluation users*. Thousand Oaks, CA: Sage.

## The authors

**Aneta Cram**, MA, Ngāti Kahungunu, Ngāti Pahauwera, is a sole contractor and PhD candidate at Victoria University of Wellington, Aotearoa. Her Phd focus is on the place that kaupapa Māori and indigenous evaluation holds in informing public policy and current government program funding structures.

**Email:** cramaneta@gmail.com

**Rodney Hopson**, PhD is Professor of Evaluation in the Department of Educational Psychology, College of Education, University of Illinois-Urbana Champaign, USA. His primary research areas lie in comparative and international education policy and politics with a focus on official, indigenous, and medium of instruction language issues and critical qualitative and ethnographic research and evaluation methods.

**Email:** hopson@illinois.edu

**Marvin Powell**, PhD, is an assistant professor in the College of Education and Human Development at George Mason University, USA. Powell focuses on measurement creation and validation, applying quantitative methods, and is interested in research using a critical lens.

**Email:** mpowel11@gmu.edu

**Asia Williams** is a PhD candidate at George Mason University, USA.

**Email:** awillid@gmu.edu

**Akashi Kaul** is a PhD candidate in the Education Policy program at George Mason University, USA. Her research interests include international development, evaluation, social justice research and research methodologies.

**Email:** akashikaul@gmail.com