PART 5

LEADERSHIP AND SCHOLARSHIP PROGRAMMES ON THE AFRICAN CONTINENT

15 Walking the talk: Shaping a transformative approach to evaluation of leadership fellowship programmes

Barbara Klugman

Overview

Drawing on 16 leadership fellowship programmes based in South Africa, which undertook a peer learning initiative, this chapter explores how transformative leadership programmes can measure their influence on leadership capabilities and social change. It explores whether evaluation approaches are consistent in their values and assumptions with those associated with leadership development programmes aimed at system transformation. Evaluation approaches must take into account that leadership development is a process with unpredictable outcomes. Hence, evaluation cannot focus on the set, predictable indicators of a traditional results-based management approach, focused on upward accountability to management and funders, but requires social learning capabilities to understand what is changing and why, and to adapt on that basis. Evaluation approaches must also consider if and how fellowship programmes model transformative leadership, or whether they manifest 'power over' rather than 'power with' fellows. The chapter describes ways that programmes are trying to address these challenges.

Introduction: Challenges in evaluating leadership fellowship programmes

De Kock and Swartz (Chapter 2, this volume, p. 24) argue that transformative leadership is leadership committed to 'unearthing, problematising, and dismantling those structures of power and privilege that act against equity and freedom'. The concept has 'a common core: progressive, flexible, and responsive leadership committed to driving social change where leaders are located, while creating a strong collegial culture that engenders critical agency, awareness, and action' (De Kock and Swartz, Chapter 2, this volume, p. 34).

There is growing recognition that influencing social change requires operating in systems which are characterised by multiple players and relationships, multiple perspectives, and unclear boundaries (Caffrey & Munro, 2017; Williams & Van 't Hof, 2016). There are no recipes for catalysing, strengthening, or sustaining individuals as transformative leaders in this context; indeed, the concept itself is still in development. Richardson and Patton (2021, p. 142) argue that '[b]ecause complex adaptive systems operate in a frequently unpredictable

context – not merely unknown but unknowable – governance and leadership involve navigation skills, not command-and-control authority'. This has significant implications for what it means to be a transformative leader and for a programme for transformative leadership. 'When complexity of the environment reduces certainty because of turbulence and lack of definitive knowledge about how to achieve desired results, strategic approaches must be more emergent and flexible' (Patton & Patrizi, 2010). In addition, according to Atkinson, Lasbennes, and Nabarro (2021, p. 133):

> In a world where they are held accountable for their actions and outcomes, and constantly evaluated against these, the realisation that control is illusory and that contexts matter a great deal is deeply disturbing. It raises profound questions about what leaders should be doing and how to determine what can be or has been effective.

Moreover, leadership for social change is increasingly understood as a collective and dialogical process (Kliewer & Priest, 2019; Orians, Chew, & Rowland, 2018).

Programmes are similarly caught between the results-based paradigm that has shaped the management of development initiatives for the last 30 years (Lomofsky & Grout-Smith, 2020) and a transformative evaluation approach (Ligero Lasa, Espinosa Fajardo, Mormeneo Cortes, & Bustelo Ruesta, 2014), characterised by 'inclusion, addressing of power, and adaptive learning approaches that better mirror contemporary understandings of social change as a process rather than a final, pre-determined result' (Rayner & Bonnici, 2021, p. 21). The former has come to be associated with inflexibility and upward accountability, and the latter with trust, enabled by more equitable power relations between funders and those funded, and an accountability to the mission of the organisation and its beneficiaries (Guijt, 2010). This approach enables trust-based accountability (Bovens, Goodin, Schillemans, & Mansbridge, 2014) in which leaders 'navigate by judgment' rather than by pre-set targets (Honig, 2020).

The evaluation challenge is exacerbated where a leadership fellowship programme aims to enable individuals to lead wherever they are, as opposed to leadership training being directed towards people who already have a social change goal focused on a particular place or issue. In the latter context, training effectiveness can be judged in relation to the goals of that initiative (Meehan & Reinelt, 2012a). But by what criteria and process does one assess the effectiveness of leadership development of people working on different issues in different contexts?

A related tension is that just as leadership programmes are increasingly needing to take on board the questions of 'if' and 'how well' they demonstrate the types of leadership they aim to build, so the evaluation process has to consider what values and principles it demonstrates.

A note on research questions and methodology

The first set of questions this chapter explores are what a transformative leadership programme might need to know in order to assess if, and how well, it is contributing to influencing both individual fellows and social change; why this is the case; and how programmes can improve their offerings and processes. The second question is what kind of evaluation approaches themselves demonstrate a transformative approach.

The issues and findings presented in the chapter are the product of a peer learning initiative among 16 leadership fellowship programmes based in South Africa, facilitated by the author. Fellowship programmes with an interest in evaluation were identified and contacted. Interviews were conducted with each of the 14 programmes that responded, and these interviews identified the programme's core monitoring, evaluation, learning values and processes, what was working well, and where challenges were experienced. A survey captured descriptive information from 11 of the programmes.[1]

Findings from this engagement were shared with the programmes, and in this process topics were identified for a peer learning agenda that covered six topics in three online sessions during October and November 2020. Some organisations with nascent fellowship programmes joined the conversations, and one fellowship programme that heard about the process joined mid-way. For this reason, these programmes were not included in the descriptive data of the programmes.

A report was produced (Klugman, 2021) covering the themes that emerged from the deliberations as well as illustrative examples of how programmes were addressing the issues. Programmes added to the report and indicated what information they were happy to share in the public realm.

The chapter also draws on the author's learning from involvement as a practising developmental evaluator during the 'incubation' of Tekano: Atlantic Fellows for Health Equity (South Africa 2016–2019) and on a 2021 review of the perspectives and experiences of alumni of the Asia Safe Abortion Partnership's Youth Advocacy Institutes. The process was not a research initiative shaped by the literature, but rather a grounded exploration of the experiences of evaluators and programme staff with a utilisation focus – a peer learning initiative that aimed to strengthen participants' own thinking and confidence about and actual evaluation practices.

---

1   One of these, the Kellogg Fellows Alumni Network, has long since stopped running fellowships, but is running a very active alumni programme.

The individual capabilities and social connections which leadership programmes aim to catalyse

One cannot consider what to evaluate and learn about without articulating what a programme aims to achieve. Almost all programmes under consideration aimed to influence three dimensions: the head, heart, and hands. The 'head' refers to fellows' knowledge and conceptual capabilities, including their understanding of context, of diversity of structural drivers of inequities, and of interests supporting and opposing their objectives. Within this, a few aimed to inculcate fellows' ability to recognise and make sense of multiple ways of knowing around an issue, and to form their own perspective based on this analysis. The 'heart' implies fellows' sense of self or emotional intelligence, resilience, reflexivity, and what Klugman (2021) describes as their 'relational competence'. Feminists would call this building 'power within' (VeneKlasen & Miller, 2007) and psychotherapists might call it 'internal authority' or agency (Parsons, 2021). The 'hands' denote fellows' ability to apply a multiplicity of skills, from working in teams, to organising others, to influencing public and political opinion, to bringing a complexity and innovation lens to whatever they take on. The programmes hypothesise that all three dimensions are key to developing or strengthening the transformative leadership capabilities of fellows.

Each one of these areas of leadership requires a journey. It is not a matter of a fellowship programme pouring something into a person so that they walk out a leader. It is about programmes offering tools and information along with processes that engage fellows' experience and expertise, which each person will synthesise in their own way and apply in their own time. Another underlying assumption that came to light is that relationships are important. Most fellowship programmes in this initiative recruited individuals rather than groups. Both the literature and these programmes recognise that leadership is relational and that for programmes to achieve their social change goals, they need to 'foster a network mindset' among participants (Meehan & Reinelt, 2012a, p. 7). As a consequence, transformational leadership has to operationalise the features of complex systems by navigating networked relationships and eliciting and valuing multiple perspectives (Meehan & Reinelt, 2012b; Richardson & Patton, 2021). The wider one's networks of influence, or social capital, the greater one's chances of being able to influence the diversity of stakeholders that can make change happen.

Some take this further, arguing that 'leadership is collective, emergent, and co-constructed through dialogue' (Kliewer & Priest, 2019, p. 2), thereby reinforcing the shift from individual leadership development towards leadership learning as a collective systems level endeavour. Morse (2002, p. 10) notes that '[t]he collaborative style of leadership has become increasingly more important [in order] to look at issues from the myriad of perspectives now represented'. The theories of change of the fellowship programmes envisage the fellows in each

cohort and across cohorts as potential social capital (Hodgson, 2020) for each other. Some shape their recruitment to facilitate this. The assumption, supported by evidence in some programmes, is that fellows will build relationships during the fellowship programme that they then leverage in support of system change. In practice, the programmes had learned that they have to deliberately incorporate relationship building into their pedagogy and processes if they are to enable the building of such social capital. Thus, the process orientation and relationship-building facets of leadership programmes need to be foregrounded as part of the explicit aims of a programme, and this will assist in developing monitoring methods to track their success.

What do leadership fellowship programmes need to learn, and how to do this?

When deciding on evaluation strategies, leadership fellowship programmes must be clear about what needs to be learned, and only then decide how this can be measured.

Evaluating shifts in individual capabilities

Programmes have to recognise that not all fellows will come in with the same capabilities, and that each fellow will process the information and experiences in different ways and apply them differently at different times. Much of the application will manifest well after the programme is over.

In the short term, evaluation will be looking to see how fellows react to the experience, and what new ideas, ways of seeing, and ways of being resonate with them and motivate them to participate actively in the programme. The Kirkpatrick model (Kirkpatrick & Kirkpatrick, 2006), developed for evaluating corporate training programmes, but now more widely used (Paull, Whitsed, & Girardi, 2016), describes this as level 1 of an evaluation: reaction. In the evaluation literature, it is generally recognised that in the short term, training programmes need to assess whether the training experience itself excites and motivates participants (Garred & Refai, 2020). They also need to assess whether fellows internalise the new thinking, information, values, or tools – learning: Kirkpatrick level 2. This is an essential step if a programme is going to influence participants' actions (Kirkpatrick level 3), and if those actions are in turn going to make change happen (level 4). That said, success in the first two levels does not guarantee shifts in fellows' behaviour.

Most programmes in this peer learning initiative, as with training programmes generally, focus on the first two levels, because they can be captured during and at the immediate end of the training, whereas shifts in fellows' behaviour and influence are much harder to identify (Meehan & Reinelt, 2012a; Reinelt, Foster, & Sullivan, 2006; TCC Group, 2018).

This is where the tension emerges between traditional evaluation approaches and those needed for evaluating fellowships building transformative leadership capabilities. The Kirkpatrick model was built for evaluating corporate training programmes. For these, behaviour change – Kirkpatrick's level 3 – means behaviours on the job, which can be relatively easily identified in routine workplace management processes. The Kirkpatrick model's level 4 (results) measures 'the impact that the training has on broader organizational or institutional outcomes' (Steele et al., 2016, p. 324), or in Kirkpatrick Partners' (n.d.) terms, 'the degree to which targeted outcomes occur as a result of the training and the support and accountability package'. This too can be relatively easily assessed within a workplace. The methodologies for data gathering and the evaluation framework can be pre-planned, with clear indicators related to whatever changes the company intended the training to deliver, as per the results-based management approach.

However, where a programme's theory of change is that fellows will draw on their increased self-insights, expertise, and relationships to do things differently and thereby influence unspecified actors to make changes that are unspecified (over which they have no control) but that contribute towards challenging or undercutting the drivers of inequity or oppression, there are no specifiable indicators. Actual outcomes cannot be predicted. 'The inherent characteristics of advocacy make it resistant to control, predictability, and certainty' (Coe & Schlangen, 2019, p. 111). This requires the fellowship programme, and its funders or host institution, to live uncomfortably in the terrain of the unknown. It requires an evaluation approach that focuses on what fellows actually do differently (level 3), what changes fellows actually 'influence' others to make (a reframing of Kirkpatrick's level 4 'results' as 'influence'), and a retrospective reflection on if and how the fellowship programme played any role in this. Systems theorists characterise this as an emergent learning approach (Darling, Guber, Smith, & Stiles, 2016; Darling & Parry, n.d.). While fellows may make some changes during the leadership fellowship – something that is more likely the longer the programme runs – they are most likely to synthesise their learning, start doing things differently, and influence others after the programme is over. Evaluation at these levels requires resources, management support, and expertise that many institutions lack (Kennedy, Chyung, Winiecki, & Brinkerhoff, 2013).

While Ebrahim (2019) argues that an organisation may well deploy multiple evaluation strategies depending on the nature of different strategies for change, the evaluation approaches of programmes in this initiative demonstrated different understandings of systems change and accountability. This was often a source of stress for staff. Rayner and Bonnici (2021, p. xxxiv) distinguish between the dominant approach of 'technical systems change' and the emergent approach of 'transformational systems change'. The former approach brings a set framework with advance prediction of outputs and outcomes, and upward accountability for

delivery of these. The latter brings questions, curiosity, and engagement in order to understand, from the ground up, how change is happening. This adaptive learning approach explores whether the training plausibly contributed in any way, whether directly or indirectly, intentionally or unintentionally, and asks what the programme can learn from this in order to adjust its theory of change or improve its programme delivery.

In a cogent critique of the Organisation for Economic Cooperation and Development's Development Assistance Committee criteria – the standard evaluation criteria used by the development community globally to evaluate delivery using set indicators – a leading figure in global evaluation, Patton (2021, p. 53), posits the need for a change in evaluation criteria: '[T]he greatest danger for evaluators in times of turbulence is not the turbulence – it is to act with yesterday's criteria'. He proposes that interventions that aim for transformation need to ensure that their engagement itself is transformative, and that programmes that are focusing on systems change apply systems concepts in their evaluations – in this case taking on questions of complexity and unpredictability. He puts forward criteria that resonate strongly with the understanding of transformative leadership, as discussed above, such as evaluating the extent of deepening of interconnectedness and the energy this catalyses, and the degree and influence of increased diversity/equity/inclusion.

A number of fellowship programmes in this initiative monitored reaction and learning (levels 1 and 2) and had set indicators for longer-term outcomes such as changes in fellows' position in their workplaces or organisations, changes in career direction or jobs, as well as fellows' publications, presentations, or receiving awards. Beyond this, they let go of evaluating levels 3 and 4 – that is, changes in fellows' behaviour and their influence on others, and thereby on social change in communities, fields, and systems. Other programmes modelled a systems approach by using the Outcomes Harvesting methodology which, rather than asking 'Did the fellows do what we predicted they would do?', explores what fellows actually did differently and what changes actually took place, and then asks whether the fellowship programme plausibly contributed to each change (Wilson-Grau, 2019). This approach is well suited to the uncertainty and unpredictability of systems change.

To harvest the (unknown) outcomes and the role of their programmes in influencing these, these programmes engaged their fellows outside of and after the fellowships ended. For example, one does site visits to experience fellows training community members or their organisations on a topic from the fellowship, or working with them on campaigns or other actions. One programme commits fellows' organisations to partner with fellows in assessing progress in addressing the organisational change issue identified by the fellow, both during and after the programme.

All the programmes recognise the need to support alumni programmes to continue in catalysing their and the group's capacity to influence social change. In addition to this fulfilling the fellowship programmes' commitment to influencing social change, the programmes' experience is that through this participation, fellows will be more inclined to share what they are doing and what changes they are influencing over time. This enables programmes to learn in what ways the programme experience had had its intended longer-term influence.

Evaluating the building and leveraging of connections

One of the issues most programmes in this initiative found hard to gauge was whether fellows built and strengthened their bonds with each other – personal bonds as well as collaborative bonds – and then, over time, whether, and in what ways, fellows used each other as resources to take forward their social change agendas, including using each other to access decision-makers or institutions they wanted to draw into their social change efforts. Two programmes used social network analysis, providing data on the degree and depth of connectedness before, during, and after the fellowship programme (Klugman, 2021; Plastrik & Taylor, 2006; Taylor, Whatley, & Coffman, 2015). One programme attempted an analysis of the nature of engagement among fellows on the organisation's WhatsApp list (Klugman, 2021).

Does the programme walk its talk? The politics of decolonial evaluation

A number of issues seem key when looking for synergy between evaluation practice and the goal of transformative leadership. These include issues of process, power, and mutual learning. Each one will be considered in turn.

The evaluation process as the focus of learning

The first issue is that evaluation has to be built into the programming process rather than seen as a technical and separate role. Most programmes in this initiative invested substantial time in the technical side – that is, shaping their recruitment criteria and processes – and the content and pedagogy of programmes. In addition, where they had separate evaluators, programme staff perceived evaluation and the evaluators as having separate roles, not embedded within programming, something also documented in the literature (Taylor & Liadsky, 2018; TCC Group, 2018). The evaluator in one of the leadership programmes in this peer learning initiative notes, 'To programme staff their main concern is about implementation, implementation, implementation. As an evaluator you're looking at the why of this programme – what kind of impact are we trying to have?' (Klugman, 2021, p. 11).

Yet evaluators of social change interventions note that programme development and implementation should not be separated from evaluation (Symonette, 2007). An organisational learning culture should be built, in other words, 'when an organization uses reflection, feedback, and sharing of knowledge as part of its day-to-day operations…and applying that learning to improve' (Center for Nonprofit Excellence, 2016, para. 1).

What was striking among programmes that did not have a highly elaborated data-collection system in line with results-based monitoring was that programme staff often articulated a sense of disempowerment in relation to the idea of what 'M&E' should be, given their understanding of monitoring and evaluation as the quantitative measurement of set indicators. They had the impression that they were not good enough. Yet when actually describing their processes, they were indeed engaged in emergent learning without the separation between programming and evaluation described above. For example, Naledi Maite of the Zanele Mbeki Development Trust Fellowship stated:

> The thing is that we haven't yet developed an approach to monitoring and evaluation because we're new and after the first year we put our energy into improving the programme based on what we learnt in the first year. (Klugman, 2021, p. 9)

This approach of ongoing reflection and improvement is entirely appropriate when outcomes are unpredictable, and it is aligned with a transformative leadership approach.

Dealing with power

The second challenge regarding evaluating a pre-planned programme with set indicators is that both the programme and the evaluation processes constitute forms of 'power over' (VeneKlasen & Miller, 2007), where the programme knows best in relation to programme content and pedagogy, and the evaluator and funder, in turn, judge the programme staff's and fellows' performance. This creates a culture of fear rather than one that enables learning through exploration and risk-taking, both of which are essential for systems change (Williams, Dickman, & Smurthwaite, 2021). One of the defining dimensions of an emergent learning approach is the need to recognise and learn from failure, something that the results-based management approach disincentivises. 'Devoting organizational resources to creating an environment where it is okay to admit failures, express fears, ask questions, and engage in constructive debate, is critical to achieving tangible results during and after the program' (Spasser, n.d., p. 7).

In some cases, the challenge is not the staff, but the institutions who run the programmes or their funders, who assume a top-down form of accountability based on reporting on agreed-upon metrics. In this way of operating,

institutional and programme processes do not align with the very principles of transformative leadership practice that they aim to inculcate among fellows. Such principles include commitment to social justice and ethical behaviour; recognising and reconfiguring divergent levels of power; working collaboratively to foster connection, social inclusion, and equity (Rayner & Bonnici, 2021; De Kock & Swartz, Chapter 2, this volume); and building trust across diversity (Odora Hoppers, 2014). Programming itself would require staff '[giving] up some control in favour of encouraging the self-organisation and initiative of participants as co-designers' (Meehan & Reinelt, 2012b, p. 12). In relation to both programming and evaluation:

> Accountability would be about understanding, cultivating, supporting, and sharpening judgment while earnestly and honestly addressing successes and failures. It would be aimed towards accompanying and supporting staff, not monitoring and controlling them. It would also attempt where practicable to draw on the judgment of recipients. (Honig, 2020, p. 22)

Ligero Lasa et al. (2014, p. 72) describe this orientation as a 'critical change-driven or transformative paradigm' of evaluation which centres on identifying power dynamics and addressing these not only in relation to the programme being evaluated, but also in relation to the participation of those with less power in the evaluation process.

To live by these principles, programme staff would need to invite fellows into the processes of conceptualisation, running, and evaluation of the programme. In extremis, this would mean starting from scratch each year, but even within a broad institutional framework, processes of real participation that draw on fellows' expertise are likely to strengthen programme content and the fellows' experience. To really operationalise systems thinking would mean practising the principles associated with transformative leadership. Programme staff and evaluators, *with* fellows, would have to invest in the process of building social learning capability. Wenger-Trayner and Wenger-Trayner (2021) describe this as a process of building legitimacy, bringing in people with multiple perspectives, and building everyone's identification with the goal of improving the programme. They note that 'without a bottom-up component in which participants experience agency, there is no social learning capability' (Wenger-Trayner & Wenger-Trayner, 2021, p. 103). They argue that attention to power dynamics, particularly to who is at the table, and whose voices are listened to or are silenced, is critical for success. This thinking resonates with the notion of 'ubuntu leadership' or decision-making through consultation and dialogue (Chedondo, 2019). Ultimately, everyone has to want to be part of a collective learning journey. Staff, facilitators, and fellows all have to experience it as being valuable for themselves and for the overall purpose of the process. This emergent learning approach asks: 'What did we

hope to achieve?', 'What actually happened?', and 'Why – was it because of what we did, or because of actions of others or the broader context?' Interrogating change, or lack of change, in this way, allows an institution to then look at what it should continue to do, and which parts of its programming it should adapt and reassess (Darling et al., 2016).

Some programmes that participated in the peer learning initiative identified that without building in this kind of collective reflection process, fellows experienced evaluation as extractive and were less inclined to respond to surveys or other data-collection efforts. Their experience indicates that evaluation has to focus on what is of interest to fellows and what they consider meaningful. As Kellogg Fellows Southern Africa Liaison Mary Hlalele said:

> If they are not experiencing value from the programme, then they are less likely to put in the time to respond. Unless people are committed to the programme, they won't respond truthfully or will respond to get you out of the way. So we've been very mindful of that – how to keep alumni actively interested – activities keep people totally connected and interested so when you send out an evaluation question they respond from the core of their being. They will only respond if you're relevant to how their lives are shaping. (Klugman, 2021, p. 30)

Programmes in this initiative found that to the extent that they had recruited people who had demonstrated some leadership potential, these fellows were likely to be people who saw their role as challenging the status quo, to 'dismantle structures of power and privilege' (De Kock & Swartz, Chapter 2, this volume), and the first thing they challenged, on entering the fellowship programme, was the power and privilege undergirding the institution running the fellowship programme, its staff, facilitators, and evaluators. It took highly adaptive staff, facilitators, and evaluators to balance having clear objectives and ways of engaging fellows to buy into these as part of the recruitment and orientation process, with enabling fellows to contribute their expertise in shaping and evaluating the programme and outcomes in a continual cycle of learning and improvement. This exemplifies 'adaptive management' in that it is not about changing the overall goals of an initiative, but about making changes in programming to achieve these goals, based on learning (Vähämäki & Verger, 2019, p. 30). Indeed, staff and facilitators need to exemplify transformative leadership and emergent learning and foster a learning community by breaking down power differentials and building the trust needed to learn from the diversity and expertise of all participants. They need to be extremely self-reflexive and able to recognise and amplify diverse ways of knowing and doing among the fellows – one of the cornerstones of a decolonising approach. They need to be able to build a culture of 'power with' rather than exercising 'power over' (VeneKlasen & Miller, 2007). Haugen and Chouinard (2019, p. 389) note,

'If power is not a central focus in all aspects of the evaluation design, process, and outcome, results may be invalid.'

None of this precludes documentation – that is, the monitoring of the basic metrics of the programme to inform learning and accountability. As Ebrahim (2019) notes, an organisation may well deploy multiple evaluation strategies depending on the nature of the programme's different strategies for change. Hence, the programme can and should keep a record of outputs (what it has done) and of participant outputs, and use these for learning. For example, programmes likely need to monitor the level of participation or drop-out, and whether these differ by gender, age, education level, or another descriptor, as this will provide information that could feed into programming decisions and recruitment decisions. This kind of data would not require a complex process of social learning. They are also easy to obtain, and are therefore not where the evaluation conundrums for transformative programmes lie.

Evaluation for learning

A last rationale for an emergent learning approach is that it is an essential capability for transformative leadership. For this reason, evaluation for learning needs to be built into leadership programming, not only for the programme staff to learn and improve programme pedagogy and content, but for fellows to strengthen their own skills in methods and processes that enable learning based on evidence and collective, inclusive deliberation. The better this is done, the more fellows will embed reflective practice and collective learning into their ways of operating, and bring this into both their alumni activities and their social change efforts in the world. In this way, they will also see value in the fellowship and alumni programmes, and want to continually assess if and how well these programmes are working, including if and in what ways they personally are contributing towards making them work. They will also want to know if and in what ways their fellow alumni are influencing social change. The programme will not need to battle to gather information; fellows will not experience it as extractive. Rather they will consider the programme and the learning theirs, in a form of distributed 'power with'.

Rather than requiring fellows to fill in endless surveys looking for shifts in knowledge, attitudes, and practice, many programmes used forms of evaluation that also contributed to achieving programming objectives, for example storytelling, which helped other participants learn from each fellow's development and the initiatives they had taken during or after the programme. It also strengthened a sense of connection and solidarity among fellows, while providing evidence of the degree to which each fellow had internalised or strengthened certain ways of being, seeing, and doing from their programme experience. Programmes found this kind of evaluation approach extremely

valuable in enabling their learning and adaptation, even while it did not provide the metrics expected in their logical frameworks. The Outcome Harvesting approach helped somewhat in that by categorising whatever changes fellows described in their behaviour or their influence over others, the programme could identify patterns of change to which it had contributed, whether directly or indirectly, such as strengthening of community organisations, alliances, or movements; influence on media or decision-makers' discourses; and influence on policy and practice. They could also use the stories of change as case studies to strengthen their understanding of the different processes of leadership development and influence, and the ways in which the fellowship had programme contributed towards these.

Conclusion

In answering questions around what a transformative leadership programme might need to know in order to assess if, and how well, it is contributing to influencing both individual fellows and social change; why this is the case; and how programmes can improve their offerings and processes, this chapter showed that it is more conventional and easier to monitor fellowship outputs, fellows' reaction, and any immediate shifts in knowledge or attitudes, which can be measured in relation to pre-planned indicators. However, the need and the challenge is to learn about the ways in which fellows have changed their own behaviour and influenced the behaviour of others towards social change over time – outcomes that cannot be predicted, yet are key to assessing whether a programme is contributing towards building transformative leadership.

In terms of what kind of evaluation approaches are appropriate for a transformative leadership development programme, the chapter argues that the approach needs to engage with the main dimensions of complex systems – that is, recognising and inviting multiple perspectives and ways of knowing, challenging power relations to create space for building relationships of trust across diversity, setting boundaries beyond the immediate moment of the fellowship, and being not only open to but inviting the unexpected and unknown. It argues that the more that evaluation is merged into programming, fostering a seamless culture of collective and emergent learning, the more likely it is to be able to respond to the needs of the moment, to a diversity of fellows, and to social change efforts both during the programme and with alumni over time.

#TransformativeEvaluation
It is time to evaluate programmes aiming to foster transformative leadership with processes and methods that themselves strengthen participation, critical thinking, and collective action.

Questions for discussion

1. What criteria would you suggest for evaluating whether a leadership fellowship programme is successful?

2. Do you think it is fair to expect people who go through fellowship programmes to stay in touch and keep sharing their achievements and challenges? Why? What might make it more fair?

3. How can we make the process of evaluation more transformative in nature?

References

Atkinson, J., Lasbennes, F., & Nabarro, D. (2021). Reflecting on our times: Valuing transformative leadership in real-world 'living systems'. *American Journal of Evaluation*, *42*(1), 130–138.

Bovens, M., Goodin, R. E., Schillemans, T., & Mansbridge, J. (2014). A contingency theory of accountability. In M. Bovens, R. Goodin, & T. Schillemans (Eds) *The Oxford handbook of public accountability* (pp. 55–68). Oxford, United Kingdom: Oxford University Press.

Caffrey, L. & Munro, E. (2017). A systems approach to policy evaluation. *Evaluation*, *23*(4), 463–478.

Center for Nonprofit Excellence. (2016, 11 May). *What's a learning culture and why does it matter to your nonprofit?* Groundworks New Mexico.
Retrieved from https://www.groundworksnm.org/news/whats-learning-culture-why-does-it-matter-your-nonprofit/2016-5-11.

Chedondo, M. (2019). *Ethics and leadership: A critical analysis of cultural resources for ethical leadership in Southern Africa*. PhD thesis, University of KwaZulu-Natal, Pietermaritzburg, South Africa. Retrieved from https://ukzn-dspace.ukzn.ac.za/handle/10413/16818.

Coe, J. & Schlangen, R. (2019). *No Royal Road: Findings and following the natural pathways in advocacy evaluation*. Center for Evaluation Innovation.
Retrieved from https://evaluationinnovation.org/publication/no-royal-road-finding-and-following-the-natural-pathways-in-advocacy-evaluation/

Darling, M., Guber, H., Smith, J., & Stiles, J. (2016). Emergent learning: A framework for whole-system strategy, learning, and adaptation. *The Foundation Review*, *8*(1), 59–73.

Darling, M. & Parry, C. (n.d.). *Emergent learning: Taking learning from experience to a new level*. The Systems Thinker. Retrieved from https://thesystemsthinker.com/emergent-learning-taking-learning-from-experience-to-a-new-level/.

Ebrahim, A. (2019). *Measuring social change: Performance and accountability in a complex world*. Redwood City, CA: Stanford University Press.

Garred, M. & Refai, M. (2020). *Change inside and out: An outcome harvesting plus attitude change toolkit for peacebuilders and other changemakers – Trial version*. Joint Learning Initiative on Faith & Local Communities.
Retrieved from https://jliflc.com/wp-content/uploads/2021/04/ohac_toolkit.pdf.

Guijt, I. (2010). Accountability and learning: Exploding the myth of incompatibility between accountability and learning. In A. Fowler & C. Malunga (Eds). *NGO management: The Earthscan companion* (pp. 339–352). Abingdon-on-Thames, United Kingdom: Routledge.

Haugen, J. S. & 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.

Hodgson, J. (2020). Disrupting and democratising development: Community philanthropy as theory and practice. *Gender & Development*, *28*(1), 99–116.

Honig, D. (2020, 5 February). *Actually navigating by judgment: Towards a new paradigm of donor accountability where the current system doesn't work*. CGD Policy Paper 169. Centre for Global Development. Retrieved from https://www.cgdev.org/sites/default/files/actually-navigating-by-judgement-paper.pdf.

Kennedy, P. E., Chyung, Y. S. Y., Winiecki, D., & Brinkerhoff, R. (2013). Training professionals' usage and understanding of Kirkpatrick's level 3 and level 4 evaluations. *International Journal of Training and Development*, *18*(1), 1–21.

Kirkpatrick, D. L. & Kirkpatrick, J. D. (2006). *Evaluating training programmes: The four levels* (3rd ed.). Oakland, CA: Berrett-Koehler.

Kirkpatrick Partners (n.d.). *The Kirkpatrick Model*.
Retrieved from https://www.kirkpatrickpartners.com/the-kirkpatrick-model/.

Kliewer, B. W. & Priest, K. L. (2019). Building collective leadership capacity: Lessons learned from a university–community partnership. *Collaborations: A Journal of Community-Based Research*, *2*(1), 1–10.

Klugman, B. (2021). *Evaluating leadership fellowship programmes: Report on challenges and lessons shared among South Africa-based initiatives*. Tekano.
Retrieved from https://tekano.org.za/2021/09/06/evaluating-leadership-fellowship-programmes-by-barbara-klugman/.

Ligero Lasa, J. A., Espinosa Fajardo, J., Mormeneo Cortes, C., & Bustelo Ruesta, M. (2014). *Making evaluation sensitive to gender and human rights: Different approaches*. Madrid, Spain: Spanish Ministry of Foreign Affairs and Cooperation.

Lomofsky, D. & Grout-Smith, J. (2020). Learning to learn: The experience of learning facilitation for grantees of Comic-Relief-funded projects. *African Evaluation Journal*, *8*(1), 1–10.

Meehan, D. & Reinelt, C. (2012a). *Leadership and collective impact: A guide for strengthening the impact of your leadership development work*. Leadership Learning Community. Retrieved from https://leadershiplearning.org/publications/leadership-collective-impact/

Meehan, D. & Reinelt, C. (2012b). *Leadership and networks: New ways of developing leadership in a highly connected world*. Leadership Learning Community.
Retrieved from https://leadershiplearning.org/publications/leadership-networks-new-ways-of-developing-leadership-in-a-highly-connected-world/

Morse, S. W. (2002). *Crafting a new design for civil leadership: A Pew Civic entrepreneur initiative report*. Pew Partnership for Civic Change.
Retrieved from http://www.civicchange.org/pdf/crafting_new_design.pdf.

Odora Hoppers, C. (2014). Wounded healers and transformative leadership: Towards revolutionary ethics. In K. Kondlo (Ed.). *Perspectives on thought leadership for Africa's renewal* (pp. 23–36). Pretoria, South Africa: Africa Institute of South Africa.

Orians, C., Chew, A., & Rowland, P. (2018). *Evaluating leadership for social change*. ORS Impact.
Retrieved from https://www.orsimpact.com/directory/evaluating-leadership-for-social-change.htm

Parsons, M. (2021). Authority and freedom. *Journal of the American Psychoanalytic Association*, *69*(1), 1163–1190.

Patton, M. Q. (2021). Evaluation criteria for evaluating transformation: Implications for the coronavirus pandemic and the global climate emergency. *American Journal of Evaluation*, *42*(1), 53–89.

Patton, M. Q. & Patrizi, P. A. (2010). Strategy as the focus for evaluation. In P. A. Patrizi & M. Q. Patton (Eds). *Evaluating strategy: New directions for evaluation* (Number 128)*, (pp. 5–28). Hoboken, NJ: Wiley.

Paull, M., Whitsed, C., & Girardi, A. (2016). Applying the Kirkpatrick model: Evaluating an 'interaction for learning framework' curriculum intervention. *Issues in Educational Research*, *26*(3), 490–507.

Plastrik, P. & Taylor, M. (2006). *Net Gains: A handbook for network builders seeking social change*. Network Impact.
Retrieved from https://www.networkimpact.org/resources/net-gains-a-handbook-for-network-builders-seeking-social-change.

Rayner, C. & Bonnici, F. (2021). *The systems work of social change*. Oxford, United Kingdom: Oxford University Press.

Reinelt, C., Foster, P., & Sullivan, S. (2006). *Evaluating outcomes and impacts: A scan of 55 leadership development programs*. Leadership Learning Community.
https://ctserc.org/documents/misc/ptfg-2017-06-16-wk-kellogg-evaluating-outcome.pdf.

Richardson, R. & Patton, M. Q. (2021). Leadership-evaluation partnership: Infusing systems principles complexity concepts for a transformational alliance. *New Directions for Evaluation*, *2021*(170), 139–147.

Spasser, A. (n.d.). *Empower, change, transform: A guide to building a successful leadership development programme*. Schusterman Family Foundation. Retrieved from https://schusterman.widen.net/s/skvbhlmx5n/leadershipdevelopmentguide.

Steele, L. M., Mulhearn, T. J., Medeiros, K. E., Watts, L. L., Connelly, S., & Mumford, M. D. (2016). How do we know what works? A review and critique of current practices in ethics training evaluation. *Accountability in Research*, *23*(6), 319–350.

Symonette, H. (2007). Making evaluation work for the greater good: Supporting provocative possibility and responsive praxis in leadership development. In K. Hannum, J. Martineau, & C. Reinelt (Eds). *The handbook of leadership development evaluation* (pp. 111–136). Hoboken, NJ: Wiley.

Taylor, A. & Liadsky, B. (2018). *Achieving greater impact by starting with learning*. Ontario Nonprofit Network.
Retrieved from https://theonn.ca/wp-content/uploads/2018/09/Achieving-Greater-Impact-by-Starting-with-Learning-September-2018.pdf.

Taylor, M., Whatley, A., & Coffman, J. (2015). Network evaluation in practice: Approaches and applications. *The Foundation Review*, *7*(2), 22–37.

TCC Group. (2018). *Summary report of the leadership development evaluation learning lab*. New York, NY: TCC Group.

Vähämäki, J. & Verger, C. (2019). *Learning from results-based management evaluations and reviews*. OECD Development Co-operation Working Paper 53. OECD. Retrieved from https://www.oecd-ilibrary.org/development/learning-from-results-based-management-evaluations-and-reviews_3fda0081-en.

VeneKlasen, L. & Miller, V. (2007). *A new weave of power, people and politics: The action guide for advocacy and citizen participation*. Warwickshire, United Kingdom: Practical Action Publishing.

Wenger-Trayner, E. & Wenger-Trayner, B. (2021). *Systems convening: A crucial form of leadership for the 21st century*. Wenger-Trayner.com.
Retrieved from https://www.wenger-trayner.com/wp-content/uploads/2021/09/Systems-Convening.pdf.

Williams, A., Dickman, J., & Smurthwaite, R. (2021). Advancing evaluation and learning on transformational change: Lessons from the climate investment funds' transformational change learning partnership. *American Journal of Evaluation*, *42*(1), 90–109.

Williams, B. & Van 't Hof, S. (2016). *Wicked solutions: A systems approach to complex problems* (2nd ed.). Research Triangle, NC: Lulu Press.

Wilson-Grau, R. (2019). *Outcome harvesting: Principles, steps and evaluation applications*. Charlotte, NC: Information Age Publishing.

Acknowledgements