

Leadership Development Outcomes Research and the need for a Time Sensitive Approach

Abstract

Many leadership development studies consider developing leadership as a dynamic process that takes time. However, few evaluative inquiries examine the effects of *time* on leadership development outcomes. As the concept of time has begun to receive the attention it deserves in leadership research, we present a case for including temporal dimensions in leadership development outcomes research. We review conceptual evaluation frameworks and published empirical evaluations in order to highlight the fact that scholars have paid scant attention to time-related considerations in programme evaluation. By using a goal-free evaluation of healthcare leadership development programme as a case example, we illustrate six types of outcomes such as a symbol, rejuvenation, discovery, change, engagement and transformation and reveal their different temporal dimensions. Based on the findings, we argue that, for evaluations to be rigorous and more meaningful to key stakeholders, adopting a time-sensitive approach may be critical.

Keywords

Evaluating Leadership Development, Time, Evaluation Frameworks, Goal-free Evaluation, temporal dimensions

Introduction

The purpose of this paper is to argue for the use of time sensitive evaluation frameworks in appraising LDPs. Developing leadership (broadly defined to include the development of human capital and social capital) takes time because it is a dynamic and longitudinal process

(Day 2011). As organisations continue to spend a considerable amount of resources in leadership development programmes (LDPs), it is important to know *what* outcomes result from these LDPs (King and Nesbit 2015; Edwards and Turnbull 2013). It is also equally important to know *when* the programme outcomes emerge and what happens to those outcomes with the passage of time (Day 2011), because such an understanding could affect how an evaluation is designed, implemented and used. By knowing the timing of outcomes, evaluators can then make a more realistic estimation about the scope of change that could be observed in a time frame and create more-efficient learning and development investment strategies. However, our knowledge about ‘when’ LDP outcomes occur and the possible effect of time on programme outcomes is limited, as even in the broader leadership studies ‘time is an unexplored dimension’ (Shamir 2011 307), and consideration of time in the leadership literature “has been cursory at best” (Day 2014 864). Therefore, this paper aims to draw attention to the omission of time-related considerations in current leadership development evaluation frameworks and programme outcomes research, and it will discuss some of the implications of this neglect. Evaluation of high potential programmes’ effectiveness across the globe (Sinar et al. 2018) seems to be paramount, as “there is evident dissatisfaction with the current state of leadership and ... Leadership Development” (Spencer-Oatey, Dauber, Jordans 2020, 295). Our case for the inclusion of temporality, therefore, is of interest to international HRD scholars.

Considerable HRD evaluation work has been done over the years (see, Russ-Eft and Preskill 2009; Burke and Saks 2009; Saks and Burke 2012; Edwards and Turnbull 2013; Han and Boulay 2013; Kraiger 2014; Russ-Eft et al. 2014; Passmore and Velez 2015), and we aim to contribute to advance this conversation in this paper by developing a case for time-sensitive LDP evaluation research. We make at least three contributions in this paper. First, we provide

a critique that most of the LDP evaluation frameworks pay little attention to the temporal dimensions of programme outcomes. We also suggest that by not explicitly addressing the effect of time on programme outcomes, many evaluators tend to base their studies on at least three functionalist assumptions. These are (1) that outcomes emerge only after a LDP has a specific start date and end-date; (2) that outcomes are a-temporal and static; and (3) that outcomes can be identified and measured by undertaking an episodic evaluation activity, which is generally conducted at the end of a given LDP or immediately after the LDP. Second, by drawing on the literature on time, we reveal that most evaluators tend to treat time as linear ‘clock time’, and this treatment leads to a narrower view of programme outcomes (that programme outcomes emerge at specific time periods, and they tend to be static and are experienced equally by all programme participants). Third, we build and advance an argument that there is an urgent need to conduct evaluations that explicitly adopt a multi-temporal view of programme outcomes. While analysing the possible reasons for this neglect, we identify the benefits and potential challenges of including time in LDP evaluations. To further strengthen our argument, we use an evaluation study, which adopts a goal-free evaluation approach – a theory-light, user-friendly, under-utilised (particularly in HRD literature) approach - that has the unique potential to unearth a wide-range of outcomes and reveal their temporal dimensions more vividly. We conclude this paper by discussing the implications of temporal work for LDP evaluation theory, practice and policy.

Leadership development outcomes research – the state of play

Industry estimates put leadership development spending as high as 14 Billion USD annually (Bersin 2019). In a survey of 1500 Chief Learning Officers of global organisations, 94 percent of them said that they either plan to increase or maintain their level of investment

the same in leadership development (cited in Kruse 2020). These organisations tend to view formal LDPs as a protective space for introducing new ideas and complex models of leadership (Yost and Plunkett 2009). They use these programmes to speed up the learning processes, and to make learning more efficient, interactive and engaging through exploration, instruction, coaching, practice, feedback and reflection (Gill 2011). They tend to view LDPs as a great way to create more leaders faster (Howard and Wellins 2008). As recruiting leadership talent becomes increasingly harder (Chartered Management Institute 2014), these organisations tend to believe that investing in leadership development will create effective leaders and leadership, so that organisational results will be automatically produced (Collins et al. 2004).

Notwithstanding the many concerns on the effectiveness of these LDPs (Sinar 2014), and the lack of programme evaluation in organisations (West et al. 2015; DeRue et al. 2014), scholars have recognised that LDPs can potentially produce various types of outcomes on multiple levels, such as on individual, organisational and societal levels (Clarke 2012). Grove, Kibel and Hass (2007) further clarified, that LDPs that focus on developing social capital, along with human capital, can produce outcomes not only on multiple levels but also on multiple occasions, and in multiple rounds. They argued that, although the direct influences of a programme on its participants (known as, ‘upstream activities’) could be observed relatively quickly, the effects of what these participants do with the skills and insights they glean from the programme, as they engage as leaders in a given context (‘downstream results’), may take a long time. Further, LDPs that are based on certain leadership models, such as sustainable leadership (e.g. Hallinger and Suriyankietkaew 2018), shared or distributed leadership (e.g. Bolden 2011) and/or relational leadership (e.g. Cunliffe and Erikson 2011), see leadership learning occurring all the time, with feedback loops among individuals, their teams, organisations and communities further reinforcing learning. However, most LDP facilitators

and programme sponsors, despite having an idea of the types of outcomes expected to emerge from their programmes, do not specify the timing of those outcomes. Additionally, most LDP evaluation studies (e.g. Reyes et al. 2019; Scott 2017; Smith 2015; Tomlinson et al. 2013; McBain et al. 2012; McAlearney 2010; Wilson and Corral 2008; Dexter and Prince 2007; Ford and Harding 2007; Carden and Callahan 2007) do not generally indicate what happens to the reported outcomes, with the passage of time. Yet, the benefits of incorporating time in evaluations are many.

One of the main benefits of incorporating time in LDP evaluations is that we are better able to confidently assert that some outcomes are, indeed, caused by a LDP. Davidson (2005) cites many such potential scenarios when time-related information can be useful for evaluators. For example, consider the most widely used taxonomy of outcomes proposed by Kirkpatrick (1996) that includes *Reactions, Learning, Behaviour, and Results*, as targets of evaluation. If an outcome, such as *learning*, occurs *before* the LDP then few evaluators would confirm a causal connection between a LDP and that outcome. Similarly, if *behaviour* outcomes, such as *improved team morale* and *increased work motivation among team members*, suddenly occur *soon after* a team leader completes his/her LDP, then an evaluator may decide whether the observed outcomes are indeed influenced or caused by the LDP. In other circumstances, however, if an evaluator observes participants' performance improvement, for example, to occur *after* a LDP but has also coincided with the introduction of a new reward system in the organisation, then the timing of that outcome may prompt the evaluator to look at what exactly has caused or contributed to that performance improvement. Similarly, by including a delay between a programme and its evaluation, or through longitudinal evaluation-designs, we may also be able to understand how much knowledge was actually retained from the programme, and what happens to other outcomes of interest over time (e.g. Gentry and Martineau 2010;

Singer and Willett 2003; Baldwin and Ford 1988), evaluators can make a more accurate judgement on programme impact and value. If we know when outcomes occur, as Peters, Baum and Stephens (2011, 105) have observed, then evaluators need not wonder ‘how long’ they should wait to assess outcomes. Thus, time-related information assists evaluators to make appropriate decisions about when to conduct evaluations, and when to make causal claims, so that resources for effective evaluation can be deployed efficiently.

Leadership development outcomes research - some reasons for the omission of time

While recognising the importance of including the question of time in leadership and leadership development studies, scholars (Day 2011; Shamir 2011) have called for incorporating a temporal dimension in LDP research. Despite their calls to do this, the concept of time has received insufficient attention in the LDP evaluation literature (Day 2014). This may be due to four possible reasons. First, there is lack of consensus among leadership development scholars on what should be developed in both leader development programmes and leadership development programmes. For example, some argue that LDPs should develop ‘meta-capabilities’, such as big-picture sense making, the ability to deliver change and skills for inter-organisational representation in individuals (Storey 2011), and few others indicate that it is an individual’s own self-identity (Day and Harrison 2007) that should be developed through LDPs. In addition, some scholars have stated that LDPs should develop an organisation’s capacity in order to allow for the five different kinds of developments: the development of individuals, the development of relationships between individuals; the development of capacities of teams, groups, networks, and communities; the development of interconnections among those collectives; and the development of culture and systems in which individuals and collectives operate (Giber et al. 2009). As a consequence of these complex expectations, programme commissioners, designers and facilitators struggle to determine when

these multiple goals will be achieved as the outcomes of the LDPs that they had sponsored.

The second reason for not including time in a LDP evaluation is that the developmental approaches, prescribed in the literature and those that are used in practice, are extremely varied (See Reyes et al. 2019; Ardichvili et al. 2016) and they do not generally include considerations of time. On the one hand, for example, Day and his associates propose a multi-level, identity-based approach to leadership development (Day and Harrison 2007), whereas Mumford et al. (2007) recommend a WICS (Wisdom, Intelligence, Creativity Synthesised) approach to LDPs. Some others (Burgoyne and Turnbull-James 2001; Leskiw and Singh 2007) propose best practice approaches that aim to integrate theory into leadership developmental practice. These approaches and best practices, although useful in clarifying ‘how’ leaders can be developed, do not specify ‘when’ outcomes of interest are expected to emerge. On the other hand, organisations tend to use ‘innovative and unusual’ methods (Edwards et al. 2015, 4) to leadership development, with relatively less of an explicit reference to the developmental theories or conceptual approaches that are proposed by scholars (Ardichvili et al. 2016). Their methods vary from off-site, group-activity based formal sessions, which are organised sometimes “in mountain locales or in close proximity to the sea and small boats” (Storey 2011, 27), to coaching, multi-source feedback, or some form of experiential learning, which sometimes include novelty items, such as “walking on burning coals, contemplating cacti, pruning bonsai trees and paint-balling” (Burgoyne et al. 2009, 3). Problematically, these approaches and methods tend to be prolonged, comprising several modules and sessions that are facilitated by visiting experts who may not fully know the learners’ contexts and challenges. It is logical to assume that some aspects of a programme may not work equally well with certain learners who may have different preferences, needs, attitudes, values and gendered power relations (Stead 2014; Hirsh et al. 2011). In addition to these formal interventions, scholars

have long argued that learning from experience is the best way to developing leadership capabilities in many cases (Lacerenza et.al. 2017; Scott 2017; Hezlett 2016; McCauley et al. 2013; McCall 2004). Recently, Boak and Crabbe (2019) identify experiences such as facing challenges, contributing to major decisions, taking on new responsibilities as part of a project team or in a new job, representing one's team/department/organisation, learning from other people whom one meets through work, acting as a mentor/coach/supervisor to someone else, among others, have been perceived to contribute to developing a group of mature managers and professionals as leaders. Unfortunately, LDP evaluation methods and tools that usefully guide organisations to understand which of these methods best works for whom, and in what circumstances, are very limited (Burgoyne et al. 2004). As a result, time is virtually ignored in outcomes research.

Third, the neglect of time in LDP evaluation is to do with the developmental contexts. As leadership is “always contextual”, and “always cultural” (Turnbull 2005, 101), the programme context becomes crucial in understanding the timing of programme outcomes. Developmental experiences, along with the many formal LDPs, happen in hugely diverse contexts i.e. in public, private and voluntary sectors, in organisations of different sizes, in different industrial sectors, in various geographical locations, and in different and unique national cultures (Modisane 2018). Learners too vary in their age, ethnicity, gender, educational, occupational and professional background and management positions. All these factors can potentially influence how a LDP is perceived, delivered, used, and applied, and how outcomes are experienced. We simply do not have universally applicable theoretical tools that help us understand when programme outcomes could be expected, if any, in specific contexts, for a particular group of learners, and how external factors might restrict the evaluators' ability to assess change accurately over time. Theoretical ambiguity still remains

as to *when* LDPs might produce outcomes of interest in a given context.

Finally, many of the conceptual LDP evaluation frameworks do not clearly incorporate the dimension of time in their prescriptions. Taxonomies proposed by Kirkpatrick (1996) Kaufman and Keller (1994), Kearns and Miller (1997) and Birdi (2010), among others, have usefully identified a range of outcomes, but they too have not addressed the question of *when* do those prescribed outcomes occur and what would happen to those outcomes with the passage of time. Prescribing ‘reactions’ as a Level One outcome, these evaluation frameworks have inadvertently promoted the idea that evaluation begins only after the learning event (Bee and Bee 2003). It is not surprising then that “most of the evaluation efforts are focused on participants’ reactions to the developmental programme (i.e. smile sheets)” (Day 2011, 46), measured normally at the end of, or soon after, a programme. Such a prescription has also accidentally encouraged many time-poor evaluators to omit emerging programme outcomes that may occur, for example, during a programme and between data collection points.

In discussing the influence of evaluation frameworks on the negligence of time in outcome studies, the Return on Investment (ROI) frameworks (Cascio and Boudreau 2011; Avolio et al. 2010; Phillips and Phillips 2007; Wentz and Hodges 2005) and Theory of Change approaches (Gutierrez and Tesse 2007) deserve special mention. The ROI proponents tend to measure programme outcomes only at a single point in time, but not the benefits that accrue over a period of time, as Eagan (2011) contended. In the past, Mitchell and James (2001) have argued that, although a development programme can help developing a skill, performance enhancement might only be seen after considerable practice, experience and further learning; thus, a programme’s influence can only be fully understood over a period of time. However, quantifying the influence of time on programme outcomes is generally absent in ROI evaluation frameworks, in spite of an explicit acknowledgement of such an influence in their

narratives (Phillips, Phillips and Ray 2015). Interestingly, proponents of ROI frameworks base their calculations on many time-related assumptions in order to arrive at a monetary value of programme outcomes. Avolio et al. (2010, 636), for example, based calculations for their *Return on Development Investment* method on several questionable assumptions, such as “leaders and followers will be influenced by the intervention for eight weeks.” Arguably, these types of assumptions, although helpful in order to arrive at rounded ROI figures, do not fully account for the possible long-term, spillover effects of LDPs. With an inherent retrospectively oriented, snapshot-type design, ROI studies are not able to report whether certain outcomes are maintained or if they have decayed after the programme (Tyler 2004). Similarly, those who use a ‘theory of change’ evaluation approach, attempt to articulate a theory of how and why a programme works, often with the help of a diagrammatic representation of boxes and arrows that show assumed causal connections between activities and programme outcomes (Weiss 1997; Stein and Valters 2012). Although critics continue to argue about the nature of those connections (Davies 2018), due credit must be given to theory-of-change evaluators for specifying short-, medium-, and long-term outcomes of a programme they are evaluating. However, even they too fail to indicate *the timing, duration, speed and intensity* of those outcomes (for an example see, Watkins, Lyso and deMarrias 2011).

To sum up, with reasons such as those stated above, there is a lack of clarity on what should be developed by LDPs, there is the variation in developmental methods that generally lack a temporal orientation, there is the complex nature of developmental contexts that make any temporal predictions difficult, and finally there is the failure of most evaluation frameworks to specify the timing of outcomes that they themselves recommend as targets, which all contribute to the practice of evaluators not considering time in their LDP evaluations. Scholarly attempts to develop theoretical clarity and theory-informed tools are, therefore,

needed, as there are benefits to understanding the timings and specifications of when outcomes occur. In light of the potential benefits, we argue that there is a need for studies that specifically incorporate time in their leadership development programme evaluation so that evaluators can (1) more clearly connect the participants' leadership-learning experience to outcomes of interest, (2) be aware of when to conduct programme evaluations, and (3) understand the effect of time on programme outcomes in order to make more accurate value decisions of LDPs. In the final part of this section, we highlight some of the notable efforts taken to highlight the temporality of programme outcomes.

Leadership development outcomes research with an appreciation of time

Grove et al. (2007), in an attempt to providing a more comprehensive approach to LDP evaluation, proposed an EvaluLEAD framework in which they acknowledged the temporal dimensions of outcomes. They argued that LDPs can potentially trigger well-defined and time bound *episodic changes* that address the deficits in individuals, as well as the *developmental changes* (i.e. effects that support growth), and *transformative changes* (fundamental changes in values and perspectives) that mark their leader becoming in a wider context, across times. Therefore, they recommend that evaluators should not over-commit to any one epistemological stance, but they should use 'evidential inquiry forms' that capture the hard evidence that links learner experience with programme outcomes, along with the use of 'evocative inquiry forms' that capture the human dimension of programme experience with its emotional richness through stories and viewpoints. Although powerful, and more sensitive to the temporal dimension of outcomes, Grove et al.'s (2007) framework is not frequently used in LDP evaluations (except in Black and Earnest 2009).

It is only recently that some researchers have begun to incorporate temporality in evaluation designs, in addition to the limited number of longitudinal studies on leader

development. To cite an example, Day and Sin (2011) in their longitudinal investigation of over 1300 university students of a leader development programme established that individuals start from different initial points in their development as leaders. They showed that individuals develop in dissimilar ways corresponding to different personal change trajectories, with some even experiencing negative changes along the way, due to their internalisation of identity as leaders, various goal orientation, and adult development processes they experience. This temporally oriented exploratory study, despite the focus on young university students who may not represent mid-level or senior managers in organisations, and the inherent problem of attrition in its design, it helps us understand that individuals do not benefit from LDPs in identical ways.

Within LDP evaluation context, in particular, a few scholars have successfully addressed the issue of temporality. We recognise three of such papers. Firstly, Gentry and Martineau (2010) who illustrated the power of hierarchical linear modelling (HLM) as a multilevel methodological technique to examining change over time in school teams participating in a LDP. They demonstrate how HLM can accommodate missing data (and thus addressing the problem of attrition) and how it can help demonstrate the types and levels of changes a LDP is intended to create in its participants, their organizations, and beyond. Their methodological contribution is invaluable for measuring programme impact over time. Secondly, Packard and Jones (2015), when evaluating a seven-year long LDP within an educational sector, were able to recognise and report that their participants' self-efficacy improves over time. This temporal orientation helped them report the accrued value of programme investment. Finally, Getha-Taylor et al. (2015) have examined the effects of time on the outcomes of a local government sponsored LDP. By focusing on the development of conceptual and interpersonal leadership skills, they found out that participants' self-

assessments change over time. More specifically, because of their temporality-integrated study design, they were able to report that there was a clear decay in conceptual skill effectiveness, although no degradation was found in self-reported interpersonal leadership skills for the time-period included in their study.

Collectively, what these studies add to our understanding is that, first, there is an explicit recognition of the effect of time on programme outcomes; second, by incorporating time in evaluation-design, it becomes possible for the evaluators to avoid some biases related to the overestimation or undervaluing of programme outcomes, and it allows them to arrive at a more realistic presentation of what the LDP in question actually achieved in a given context. Despite the inherent limitations of these studies (such as reliance on individual self-reports only), similar studies are needed. Only then will organisations know the waxing and waning of programme outcomes, and that this knowledge, in turn, can help them strategically invest in tailor-made LDPs, which are delivered at specific times when leadership learners are required to learn, unlearn and re-learn. In the next section, by way of building on the above works, we present a unique evaluation case study that includes a set of outcomes that reveal a more nuanced set of temporal dimensions with greater clarity.

A CASE EXAMPLE

A goal-free evaluation study that examines programme impact with a temporal focus

Informed by the limitations of the LDP evaluation literature in integrating temporality in evaluation designs and the benefits of doing so, as we discussed above, we designed an interpretative qualitative design-based evaluation study with a specific temporal question. One of the largest National Health Service healthcare organisations in the United Kingdom gave us an opportunity to evaluate two similarly oriented LDPs called the ‘Living Leadership Programme.’ This programme was a strategically important and investment-heavy

intervention, that aimed at creating leaders at all levels (Trust, 2017). It was offered to all middle managers, but they were given a prioritised admission after a selection process based on a set of performance criteria. This six-day long, off-site LDP used information, demonstration and practice-based methods, and focused on leader and leadership development, with a particular emphasis on developing changing mindsets and developing networks. As typical with most healthcare learning contexts, evaluation commissioners invited us only after the programme was completed, to evaluate its impact for the participants. We decided to include a temporal question in our evaluation design. In consultation with the sponsors, we decided to ask two questions for participating middle managers: (1) What are the outcomes of the Living Leadership Programme? (2) When do the reported outcomes emerge? Since we were not able to engage in the design and delivery of these LDPs, we were able to utilise a ‘goal-free evaluation approach’, that enabled us to look for a wide-range of programme outcomes, beyond the all-too familiar categories commonly targeted in post-programme evaluations (i.e. reactions, learning, behaviour and results).

A goal-free evaluation approach refers to a process of collecting data on a range of actual effects, outcomes, and impacts of a programme, as opposed to collecting data on intended outcomes. Specifically, a goal-free evaluator collects data without any particular knowledge of, or reference to, stated or predetermined goals and objectives, and then compares the observed outcomes with the actual needs of the programme participants (Scriven 1999) with a view to making a judgement of the merit or worth of that programme. Scriven (1991) believes that the task of evaluation should be to determine exactly what effects a programme *actually* produced, and not to be too concerned with whether or not those effects were intended. Thus, without being cued up to what a programme is actually trying to do, a goal-free evaluator looks for what the programme is actually doing.

In line with the GFE technique, we avoided looking at the programme documentation, and contacting the programme staff; thus, we remained unaware of the programme objectives, intended outcomes, delivery modes and other contextual influences that participants might have had during and after the programme. After obtaining ethics approval from the relevant bodies, we asked an administrative team member, who did not have any programme-level connections, to invite all completers of this programme, to volunteer for this evaluation. From those who came forward, we purposively sampled 43 middle managers, using maximum variation sampling strategy (See Table S1). In order to capture a wide range of perspectives on outcome experiences of learners, we ensured that the sample exhibit variation in terms of their age (range 27 – 65 years), gender (Female – 74% and Male – 26%), managerial experience (range 3-35 years), and a team Size (range 6 - 52 members), from 37 different departments, representing the six hospitals belonging to the organisation. Participants belonged to 22 different cohorts, who completed the programme from six weeks to 2 years ago, at the time of the data collection. Semi-structured interviews were commonly used in phenomenographic studies (Han and Ellis 2019; Stenfors-Hayes, Hult and Dahlgren 2013) and in line with this practice, the first author conducted all 43 face-to-face, semi-structured interviews. Learners were asked (1) what is your conception of ‘the outcomes of the Living Leadership Programme’, and (2) ‘when’ did they see the reported outcomes emerge in their lives – a question important to this paper. Only these two central questions were used during the interviews, and probes were used only to uncover underlying meanings or additional details of lived experiences where necessary. On average, interviews were between 45–60 min in length. We transcribed the conversations verbatim and co-authors verified accuracy of the transcripts, against audio recordings.

We used phenomenographic data analytic techniques (Marton and Booth 1997) to

reveal the variation in participants' experience of LDP outcomes. To encourage greater awareness of alternative perspectives, as a way of improving the quality of our analysis (Trigwell 2000), we as a team engaged in the analysis. On reading and rereading of a preliminary sample of five transcripts before bringing in the full set of transcripts (Prosser 1994), we searched for the underlying foci and intentions expressed in them, to reveal the 'what' and 'how' aspects of how they experienced outcomes. We separated smaller quotes from the transcripts and combined them as one de-contextualised 'pool of meanings' (Marton 1986) with a view to interpret them within the context of all transcripts (so that we can identify how the group as a collective experienced the programme outcomes, while not losing ourselves in individual perspectives). After process of iterative analysis, that involved initial coding, continual sorting, resorting and within-group comparisons of data, we developed several 'categories of descriptions' based on the similarities found in learners' ways of experiencing outcomes. We looked for key structural relationships which related as well as distinguished them to and from each other. During this process, we cross-checked our judgements, and compared emerging categories until we felt that each category had a distinctive character, and that we had captured all ways of experiencing programme outcomes, found across all the transcripts. Then, we labelled the categories so that the critical aspects that distinguish one way of experiencing from the other, are visible. To enhance trustworthiness of the process, two of the co-authors applied the categories and analysed the data independently. They were able to achieve a high degree of intersubjective agreement on the variations found in the data. Disagreements were resolved in team discussions, to minimise researcher bias (Marton et al. 1992; Tight 2016). Besides testing of identified categories by a three-member team, rigorous debate, close examination of specific results, striving for coherent interpretations, and maintaining an interpretative awareness by explicitly dealing with our subjectivity throughout the research process (Sandberg 2005) were used to insure the reliability

and validity of the findings. Overall, we arrived at a consistent set of categories of description, representing six different ways of experiencing LDP outcomes.

FINDINGS

What are the outcomes of Living Leadership Programme?

In order to set the foundation for presenting the temporal dimensions of the outcomes, we briefly present these outcomes below. We present each category, using data excerpts that exemplify the defining features of the respective category. All identifiable information was removed and only transcript numbers (Tx) are used below to protect anonymity. We found that the programme participants experienced LDP outcomes in six different ways: namely, outcomes as

(1) **Symbols:** In this category, LDP is seen as symbols of the organisation's care, gratitude, its commitment to developing people's capability, and its way of recognising and acknowledging managerial contribution to the organisation. Participants ascribed the *felt* meanings to the programme outcomes. These meanings were mostly unrelated to the programme contents. (*The very fact I was invited has a meaning. I realised that the Trust cares for me-T43.*)

(2) **Rejuvenation:** These outcomes refer to participants' experience of a restoration of their inner energy that is often lost in a healthcare context. In this category, outcomes are seen not as the acquisition of something new from the programme, but rather as a *renewal* of knowledge, attitudes and other inner affective resources, such as beliefs and values that the participants already had within them. (*I 'refreshed' what I already knew; I was able to consolidate my thoughts-T 23, 24.*)

(3) **Discovery:** Participants experienced programme outcomes as instances of ‘discovering’, for instance, dimensions of self that had been inaccessible, of new networks that had not existed previously, and of managerial techniques that had been unknown. Participants reported that new friendships were formed, that respect for other people’s work was developed, and that potential projects for future collaboration were discovered. *(Here in this organisation, a lot of people I know only by their names, and I never met them before! I met a few of them now and we have built a social network group now. We are growing slowly as a group-T7.)*

(4) **Change:** The primary focus is on participants experiencing *changes* in various domains, such as knowledge, attitudes, values, beliefs, language, work patterns and / or an outlook on work and life. Importantly, this category does not include developing their skill-sets and improving their performance. Rather, participants see leadership learning as an acquisition of knowledge that results in changing mind-sets, in living the corporate values, in strengthening their beliefs and in revisiting their management and leadership styles. *(I had more confidence in taking certain issues forward. Only after attending the programmes, such as these, I now confidently reach out to others to proactively learn and change my actions, for the benefit of my patients-T18.)*

(5) **Engagement:** The focus in this way of experiencing a LDP outcome is managers experiencing a progressive engagement with the sponsoring organisation. *(I felt safe and felt very motivated to give the very best to the Trust -T43; The programme helped me to lead in the right context, in a way that is expected of me....That awareness is the outcome. I was able to lead my team differently because of that awareness -T6.)*

(6) **Transformation:** In this category, outcomes are experienced as the personal transformation of individuals. Leadership learning is experienced as transformative, and

participants report some type of self-realisation, shifts in action, slowing-down, and broadening their perspective. They were aware of a broader context of their lives and their awareness includes an unlearning of the old ways of seeing, doing and being and it is radically changing their lives. *(By nature, I would be dictatorial by nature. I would like to rule the world. As you grow old you realise that is not the best way...These sort of programmes help me get a perspective to understand why that way of managing creates tensions...I hope my leadership style has now become more inclusive...and I am trying to bring others up through this system. I am readjusting every day and finding newer dimensions of others and myself-T43.)*

Interestingly, not everyone in the sample reported all of the above six outcomes. The within group variation was very evident in the data, with some not experiencing specific categories of the outcomes that we identified. This experience of nil-impact or limited-impact, articulated by some participants, led to our research question on temporality (the timing of these outcomes) and provided us with the necessary explanation for these in-group differences.

When do the reported outcomes emerge?

Temporal dimensions of Symbolic outcomes

Participants' descriptions suggest that the symbolic outcomes occurred 'immediately' and 'at once' they learnt that they had been chosen to attend the event. A participant confirmed that she experienced this outcome 'instantaneously' (T30:3) with her place in the programme being confirmed, and even before attending the event. Some even said it was 'automatic'. For some participants, an invitation to the programme and/or acceptance in the programme at once created the feeling of a caring, thankful, committed and acknowledging organisation.

One of the outcomes from the leadership course was the fact that it made me think, Oh! They think we are leaders...we are classed as leaders. The leadership course emphasised my responsibility and role as a leader. Even if it did nothing else, it definitely does that. With me, personally, it resonated with me that I am a leader; that happens at once just by the course being there...! I am invited to participate... that itself brought that outcome.... (T39:4).

Temporal dimensions of Rejuvenation outcomes

The analysis revealed that the rejuvenation outcomes occurred ‘*during*’ the programme; this was in contrast to the symbolic outcomes that emerged *prior* to the programme. As seen in the quotes provided earlier, participants were able to refresh, refine, and renew their energies by listening to others, by being involved in programme activities, or by interacting with others in the programme. The programme appeared to have reminded some of their leader-responsibilities, updated their knowledge, questioned their anxieties, helped them shift their thoughts from unnecessary preoccupations, and thus raised their energy levels. These ‘updating’ and ‘refreshing’ outcomes occurred not continuously, but sporadically or incidentally during the programme, as and when participants got ‘hooked’ or inspired by specific moments *during* the programme.

When undertaking the programme, I happened to meet some of my peers in the wards, during the period that was in-between the formal sessions. We continued the conversations we had in the programme, and these unplanned meetings cemented my connections; I felt motivated again (T28:4)

The fact that outcomes emerged during the programme period (i.e. 3 months, in this case) implies that the programme duration may also be a potential time to search for programme

outcomes, in contrast to the traditional evaluation practice of searching for outcomes mostly on or after the final day of a programme.

There are so many things happening here... Look at the Trust website...there are at least a few hundred corporate documents...so many strategies, policies, protocols, and procedures. The programme helped me update the basics and told me what is important for the Trust (T36:4).

Temporal dimensions of Discovery outcomes

When compared with the above two categories, Discovery outcomes were described as emerging ‘during, and after the programme’. The discovery process was perceived to be gradual.

The only outcome I can think of is the networks I created during that programme (T16:4).

During the programme, we experienced a lot of collective group exercises. It is in those collective, visible actions [that] development happens. I was challenged by some; my assumptions and beliefs about people and the NHS were challenged. I felt uncomfortable; in those vulnerable moments. I felt that I was developing something in me... a sort of endurance, a sort of grit within me (T42:2).

I spotted an opportunity to collaborate with those who care about the same things that I care about. After that programme, we continued to meet. We are developing as a small group of people who has an increased sense of efficacy, ethics and commitment to the patients (T18:7).

Temporal dimensions of Change outcomes

The change outcomes were seen to emerge ‘gradually.’ It appears that they emerged after the programme, with some delay, as illustrated in the following extract.

Almost one year after the programme...I was at a meeting... and I was a bit agitated about a person. I saw him as very negative and disruptive.....It suddenly dawned on me ...This person is wearing a different type of HAT to me. It helps me not to be too judgmental...in terms of getting frustrated and agitated...and by not judging them: "Hey! You are a Blocker...!" I was able to stop and think: "What is the issue here? Why am I seeing him as obstructive? I recalled the whole MBTI stuff... I was struck by how quickly I was able to recollect what I learnt in that programme [as it] came to my mind. It calmed me down (T4: 2, 4).

In this extract, this change in mind-set was seen to be triggered by the LDP she attended. [After the data collection, and during the analysis, we learnt from the programme staff that this participant was referring to a 'Six Hats Thinking' exercise (De Bono, 2008), used in the LDP]. For this participant, change was seen to have emerged after almost a year. Although the one-year time lag may not be similar in other changes, it indicates that there is a time delay of some duration. This extract also indicates that the change reported was in relation to an isolated event, which happened in the past. The participant referred back to that incident at the time of the data collection.

Temporal dimensions of Engagement outcomes

Participants describe engagement outcomes as emerging 'during' and 'after' the programme. Several quotations were found in the data in relation to how the characteristics of the individual learner, of learner-groups, of facilitator, of line manager, and of senior manager contribute to either engagement or disengagement. These quotes indicate that engagement happens during

and after the programme. Consider this example of how the participant progressively experienced engagement:

After the programme, I met X, the Learning and Development Manager in our staff canteen. I instantly recognised her because she visited me during the leadership programme. She encouraged me to apply for the new Coaching programme. I am now a qualified coach, and my name has been placed on the Trust's Internal Register of Coaches. It is a feeling of being resourceful. It is a great place to be (T24:8).

Temporal dimensions of Transformation outcomes

Finally, transformation outcomes were experienced in various ways: sometimes sudden and abrupt, and other times in a more gradual process. For some, there is neither drama nor abrupt discontinuity in experiencing these outcomes. They were just ready to be surprised by effective learning and development.

The course started with an exploratory piece of work, the MBTI; the results of that came to me as a shock... This spark...this business of being poked, being wakened up, 'This is YOU...THIS is how you appear to people!' that gave me that spark...more like an electric shock....If I didn't know this, there may be other stuff I didn't know. This programme acted as a trigger, a catalyst..., and I was ready to be trained...Ever since, I am constantly aware of myself...continuously reinventing myself... (T30:4-5).

I am becoming a better leader day by day... (T11:7).

It is a constant struggle...every day brings its own challenges...no two days are the same for me (T42:2).

Our data suggests that some participants in this category experience a sudden shift of attitudes and start a new journey to newer spaces where they have not gone before. They report

a sense of excitement and joy. Some of the others had a tendency to experience these outcomes in a slow, smooth, and seamless process of awakening. They make subtle shifts in their daily managerial and leadership practice, and sustain their on-going renewal with the help of their inner motivation, empowering networks, personal coaching support and meditative practices. These distinct temporal characteristics distinguish this category from others.

Taken together, and because of our focus on temporality, we were able to make three conclusions about the nature of the programme outcomes: First, the participants experience outcomes emerging at various points on a continuum of time. Symbolic outcomes emerge even before the programme; whereas rejuvenation, discovery, change, engagement and transformation outcomes emerge during and after the programme, and are following the varying levels of sequence, frequency, duration, rhythm, and speed. Second, it was possible to make causal claims that the programme did produce these outcomes, because participants explicitly connected these outcomes to the specific LDP. Although there may be other contributing factors in the transfer context, participants demonstrated a high-degree of certainty in asserting the programme's influence – a positive aspect of time-sensitive evaluation. Third, not all participants experienced the six outcomes reported here. Some managers who reported Transformation outcome (i.e. a more expansive level) were aware of a wide range of experiencing outcomes such as symbolic, rejuvenation, discovery, change and engagement outcomes, while those that report symbolic outcome (i.e. a less complex level) do not have access to and awareness of the same range. Some reported nil outcomes or minimal outcomes on several occasions. At first, this may appear unusual, but they seem to confirm what Schriesheim (2003) observed. He cautioned LDP facilitators and evaluators not to assume that once managers are instructed about leadership theories and models that they may then know how to use them, and be flexible enough to be able to alter their behaviour in such a way as to

do what the theory/model suggests. Schriesheim (2003, 186) argued that there may be individuals “that aren’t as diagnostically capable and/or as behaviourally flexible as others.” Similarly, Santos and Stuart (2003) demonstrated that some of their participants found it difficult to translate the newly gained cognitive insights into behavioural changes, due to an insufficient motivation, or the self-confidence needed to apply what they learned on their return to work. Our study has drawn similar conclusions and confirms that there are Middle Managers who do not or cannot change; and, for them, leadership training and development may not be functional (or perhaps may even be dysfunctional). This may be, in part, due to the several learning transfer factors, that include trainee characteristics (such as their readiness to learn, within-person variation, personal circumstances (Hannah and Avolio 2010), programme-related factors (such as ineffective facilitation, failure to integrate goal-setting tasks and relapse prevention strategies etc (Hutchins and Burke 2007; Richman-Hirsch 2001) found in our LDP, and other organisational variables (cited in Ford, Baldwin and Prasad 2018; Huang, Ford and Ryan 2017; Sørensen 2017; Beer, Finnström, and Schrader 2016; Edmondson and Woolley 2003).

What is more interesting to note, here, is that our specific focus on the temporal dimension in this illustrative case provided us with an additional explanation for why certain participants articulate experiencing very little or no outcomes at all. In the past, scholars (Bluedorn and Standifer 2006; Bluedorn and Jaussi 2007; Bluedorn and Jaussi 2008; Ancona and Chong 1996) provide us with a temporal construct called ‘entrainment’, which was useful in partly explaining the reason for these types of outcomes. Entrainment refers to “the adjustment of the pace of cycle of an activity to match or synchronise with that of another activity” (Ancona and Chong 1996, 253). A team member consistently arriving a few minutes before the appointment with her line manager is an example of entrainment. It has been argued

that entrainment can occur within and between levels of individuals and collectives (Bluedorn and Jaussi 2007). Our findings suggest that there were many occasions where managers adjusted their post-programme high-energy levels to match their department's slow and stagnant routines, thus contributing to nil outcomes or minimal change. Entrainment is expressed well in the following extract:

I was promised a 'Gala' dinner at the course. But when I came back to work, what was served was only burger and chips. The department is simply not ready for me to apply most of the things that were taught in the course. In some cases, my manager is not ready, in other cases, team members are not ready....In fact, after that high-energy programme, I literally have to slow down simply to wait for the right time.... that never arrives (T3:6).

Taken together, and because of an explicit inclusion of a temporal question, this study revealed that time does not have a uniform effect on the outcome-experience of participants, and that the temporal and contextual barriers that inhibit the transfer of learning (Blakeley and Higgs 2014) need to be recognised in evaluations. In this light, we argue that evaluators who do not consider temporality in LDP evaluations run the risk of either over- or under-estimating a programme's effect on participants, and, consequently, are making inaccurate judgements about the programme's value.

Discussion

Our intention in this paper is to argue for the use of time sensitive evaluation frameworks in appraising LDPs. We have highlighted the benefits of incorporating time in the design of evaluation studies and identified reasons for the neglect of time. To illustrate the potential benefits of adopting a temporal perspective, we used our own evaluation study of a healthcare leadership programme. This case example, besides strengthening our argument, is significant

for at least two reasons. First, our study shines a light on a set of programme outcomes that reveal their temporality more clearly. For example, our categories do not reveal the all-too-familiar Kirkpatrickian taxonomy of Reactions-Learning-Behaviour-Results. Instead, we discovered that participants experienced LDP outcomes as symbols, rejuvenation, discovery, change, engagement and personal transformation. Despite some similarities, such as ‘developing networks’ in other studies, (e.g. Simmonds and Tsui 2010), the symbolic nature of the LDP was one of the highlights of our discovery, as many evaluation frameworks, which are prescribed by many HRD scholars, do not hint at such an outcome. In an era of disengagement, specifically in healthcare organisations, identifying the symbolic nature of LDPs – that effectively begin to emerge even before the actual programme – becomes critical. Second, adopting a temporal orientation enabled us to conclude that outcomes emerge at various points in time, and that time does not have uniform effects on the participants’ experience of programme impact. By clearly articulating the variation in learners’ outcome experience and the varied influence of time on their experience, we were able to make more accurate and credible judgements about the merit and value of the programme, in our reports to the sponsors. Even those who evaluate short LDPs that do not have the luxury of 22 cohorts, as we had in our case, could still specify what outcomes are experienced by whom, and what differential effects time could have played in learners’ experience of the outcomes observed.

Our study has some limitations. We analysed outcome experiences of learners from 22 cohorts, with elapsed time of six weeks to two years since the end of the programme. This one-time, cross-sectional evaluation falls short of revealing the differential effects of these different lengths of time. The use of phenomenographic data analytical techniques helped us reveal the range of meanings within the sample group, as a collective, and not the range of meanings for each individual within the group, for example, by differentiating individuals as old-time

graduates and more-recent completers. This means, despite revealing a full range of experiential differences in the whole group, our phenomenographic “snapshot” is not able to show the possible effects of time *on each individual*, as that is not the focus of a phenomenographic analysis (Akerlind 2005). Moreover, the call to conduct this evaluation came to us only *after* the programme; this after-event opportunity offered two advantages: to maintain the goal-free stance, and to include the temporal question in the study. It is possible that factors, not related to the chosen LDP could have influenced their outcome experience, over time. We also note that the six outcomes identified in our case are on an individual or group level effects and none of them are related to Kirkpatrick’s Results level. Data collected longitudinally and at a different time could have revealed such impacts, if any. Recognising the intrinsic complexities of integrating temporal dimension in evaluation, we acknowledge that for many LDPs, judging the appropriate time period may be challenging. We concur with Day (2011), who acknowledged potential limitations of what a single research study can achieve in its lifetime, and yet similar studies are needed.

In a similar vein, there is a need to understand *the speed* with which these outcomes grow stronger or become weaker (Huy 2001), and their trajectories over time (Day and Sin 2011). Studies that explore when an outcome starts to emerge and ceases to exist, and whether or not an outcome transforms during its life-time, when participants progress in their career, remains relevant. Those contributions (for example Ancona et al. 2001; Mumford 2006) that comprehensively deal with many temporal aspects of leadership may be useful starting points for such studies. Particularly, those time-related categorisation concepts, as proposed by Bluedorn and Jaussi (2008), which are including entrainment, polychronicity, temporal depth (leadership learners’ perception of distances into the past and future), and learners’ preference for transferring the learning fast, may be a useful starting point for advancing conceptual

developments within the LDP evaluation. By providing new theoretical insights, such attempts would enable us to be clear about when and how often we should measure key variables and outcomes, and how to measure the ‘correct’ lag across the various types of programme outcomes, as implied in the classical work of Mitchell and James (2001).

The possible identification of new and emerging outcomes *over a period of time* also raises questions about two other issues, namely (1) the measurement of those outcomes, and (2) the timing of the evaluation itself. First, aggregating all these known and yet-to-be-known outcomes into meaningful constructs, and devising measurement strategies, together with appropriate time frames, will potentially be very challenging when it is not clear what we mean by programme impact, and what it is exactly that we are measuring in LDP evaluation. There becomes a need to seriously reconsider and define the summative evaluation’s unit(s) of analysis. It is because it is highly likely that programme outcomes might appear across levels, in-between levels, and in other unfamiliar places, and with the passage of time. This means, whilst acknowledging the powerful utility of a multi-level analysis, we also speculate that we may even need to have additional analytical approaches, where time (or duration of an outcome-experience) is seen as a new level of analysis - a possibility implied in some scholarly works (Pentland 1999; Zimbardo and Boyd 1999). Thus, by emphasising the temporal dimension of programme impact and influence that have been currently ignored, we have also highlighted new technical challenges of outcomes-measurement associated with programme impact. Second, depending on the timing of the evaluation, what an evaluator hopes to see, capture and report might also vary. This inherent time-related limitation of evaluation must also be acknowledged by evaluation-researchers, when communicating programme results and value (Russ-Eft 2007).

The above consideration of temporal dimensions of outcomes particularly challenge those

who advocate ROI frameworks (e.g. Cascio and Boudreau 2011; Archer 2013), specifically on two issues. The first issue concerns the retrospective nature of ROI studies. For example, in their meta-analytical study, Avolio et al. (2010) proposed a strategy for examining the ROI of leadership development programmes by focusing only on the direct impact of the intervention on the leader and follower. They hasten to acknowledge the extent to which leadership development interventions could have positive and cascading effects on indirect followers, and on the unit /organisational climate and on the culture (and those benefits, noted as ‘unknown unknowns’), and these were not quantified in their calculations. They make clear the extent to which leadership is connected to employee commitment, and, in turn, customer engagement, stating “that the return on investment may have underestimated the positive impact on customer retention, repeat purchases and so forth” (779). Thus, by measuring outcomes at a particular point *in* time, and not outcomes *over* time (Aragon-Sanchez et al. 2003), these scholars, despite their acknowledgement, were not able to include the potentially on-going unanticipated outcomes. The question here is how they might reliably quantify the value of these new emerging outcomes. It is possible to speculate that there may be other subtle outcomes (the unknown unknowns), which are yet to be identified and understood, and the effects of which we have yet to quantify meaningfully. Therefore, it is possible to argue that either most of the ROI advocates tend to underestimate/report the actual value of LDPs, or they provide only a partial picture of the overall impact of a programme, in their ROI calculations. This is not to argue that evaluators should not attempt to estimate the value of LDPs because such calculations are complex and unreliable. Quite the contrary, we believe that only by engaging in more innovative efforts to identify the effect of time on programme outcomes will we be able to make better estimates of actual programme value.

Secondly, organisations that intend to use ROI frameworks in the future should not underestimate the benefits of LDPs by ignoring the cascading effects of a programme. Similarly, they should not overestimate programme value by ignoring the *possible outcomes decay* in their calculation. If they are to gain an accurate understanding of how and whether investments in leadership development results in quantifiable benefits, they need to be mindful of these potential biases. It would also be useful for evaluators and researchers to record *when* outcomes are observed and measured. They should also acknowledge that there is a possibility of the emergence of other outcomes over time, and the influence of time on those outcomes, as argued by Shamir (2011).

Overall, in line with Shipp and Cole (2015), we alert that LDP evaluators should avoid temporal blind spot evaluations based on simplistic designs, such as a one-shot survey. Evaluators should not assume that a short evaluation effort will identify all of the possible programme effects. We encourage evaluation-practitioners to select the appropriate time periods in which to conduct evaluations, so that they are then able to capture a possible temporal influence on programme outcomes. If evaluators carry out a snap-shot observation of programme outcomes, it is highly likely that they will either miss many of the programme outcomes, or will overestimate the value of certain outcomes, as some outcomes tend to wear off with time – a limitation of our own study too. A guiding question could be: What time frames make sense given the objectives, nature, intensity and the theoretical basis of the programme being evaluated? In our case, for example, the data suggests a range of time periods that include a point even ‘before’ the programme’s actual start date (for symbolic outcome), certain points ‘during’ the programme and ‘in-between’ modules (for rejuvenation and discovery), a set of multiple points ‘after’ the programme (for change and engagement programmes, and some ‘specific, peak, disruptive moments, experienced suddenly, at random

points' of the LDP, when some learners experienced rapid, personal, positive transformation outcomes. The key point here is that the timing and duration of evaluation must be contextually relevant to the particular programme being evaluated. In essence, time periods selected by evaluators determine what they learn about the programme. Avolio et al. (2010) cautions that evaluators who engage in evaluating LDPs that are based on theories such as transformational and collective leadership specify the deep and lasting relationship changes in and among leaders and followers. The evaluators of such programmes should carefully consider the length of their evaluation studies, and they need to be mindful that the evaluation studies should be proportionate to the nature of what is being developed. For example, if the programmes are based on collective leadership theories, which aim to develop mutually influencing behaviours among managerial teams, then evaluators should plan their evaluations accordingly, as such behaviours will take time to develop. However, this is not another call for more longitudinal studies, although such studies are important, and scholars have long argued for more such studies elsewhere (Hunter et al. 2007; Avolio et al. 2009; Ford and Sinha 2008). Rather, what is suggested here is the need for evaluators to look for the possible effects of time on programme outcomes, with a view to understanding how a programme influences participants over time, focusing on the duration and stability of programme outcomes. Such a focus will allow practitioners to identify programmes that produce outcomes over time and the resulting evidence will, in turn, enable them to better demonstrate the value of LDPs to key stakeholders.

Implications for LDP evaluation theory, practice and policy

To advance LDP evaluation theory, it is important that LDP evaluation scholars innovate *time-sensitive evaluation frameworks* in the future. By intentionally incorporating time in conceptual evaluation frameworks, such innovations could potentially affect the way an

evaluator conceptualises and measures outcomes. As a result, the time at which something occurs, as well as the time at which an evaluator engages in evaluation, and the duration of an evaluation, and the possible effect of time on outcomes could all then become integral parts of the measurement in LDP evaluation. This temporal focus will enable us purposively to look for outcomes (such as mental states, identity shifts or collaborative networks) that may be seen as part of a sequence or a cycle, and that flow through multiple levels, and time periods, and may be multiple occasions and repeated rounds.

We need evaluation studies that are specifically designed to examine how time affects maintenance or the deterioration of programme outcomes. It is possible that certain outcomes (e.g. skill development, changed mind-set, or a newly acquired identity) reciprocally influence other outcomes (e.g. improved networks, inter-team collaboration), suggesting a possible dynamic and non-linear cyclical relationships among the outcomes that influence other outcomes with the passage of time. It is possible that time affects an outcome in a number of ways. For example, a manager who gained confidence may actively engage in discovering new networks, and show more engaged leadership patterns in his or her team. After some time, when the strength of this initial kick has worn off, it is possible that the manager could become disengaged, focusing only on certain contractual obligations to get things done in his/her department, without actually engaging in inspiring others. In a healthcare organisation, for example, it is possible that this outcome-decay may have consequences over time for other wards and departments whose workflow depends on resources and tasks situated in the recently disengaged leader's ward, as shown in a study by Joseph-Richard and Hazlett (2014). It is equally possible that there may be variations in the way that these changes affect, over time, those team members who work in remote teams. Prior research (Dvir et al. 2002) has established that the physical distance between leaders and team members affect the strength of

the relationship existing between them. This means that the individual changes experienced by a leadership-learner could have various cascading effects on team members that are more distant from them than on those who work closely with him/her in the same team or group.

The focus on temporal effects of programme outcomes have implications for evaluation commissioners and policy makers too. If they advise evaluators not to neglect time in evaluations by instructing them to mention, explicitly, the time-limitations of their studies, together with details of the timing of their study and its duration, it is possible to learn about what a particular study has actually achieved in a given situation. Such explicit acknowledgements are to be encouraged with a view to inform key stakeholders about the possible limitations of a study's findings in the context of the timing of their study or its duration, as encouraged by Shamir (2011, 312). Evaluation Commissioners should also ask evaluators questions about what happens to programme-influence on learners, as time passes by, and demand explanations for time lags in evaluation reports. This might mean that the evaluators would need to provide a detailed description of the context within which outcomes emerge, and to show the stability, strength, and duration of the identified outcomes. In cases where this is not possible, as Shamir (2011) suggests, policy makers could demand that LDP evaluators provide reliable explanations for their neglect of time issues. This is critical for commissioners and evaluation policy makers because a programme that does not create the type of outcomes *over time* would have to be questioned in terms of its value. Only the time-sensitive reporting of programme outcomes could help researchers to understand better the actual change that occurs in a given context, and that, subsequently, lead to more accurate interpretation of the findings that are of use to key stakeholders. Understanding whether LDP outcomes are decayed, enhanced or maintained can help HRD managers and leadership development evaluators in public, private and voluntary organisations to strategically invest in

these developmental activities to better prepare their workforce to help them thrive in turbulent times.

Conclusion

To sum up our argument, our review illuminates that leadership, when viewed from a process perspective (Fairhurst 2017; DeRue and Myers 2014), is a dynamic and longitudinal process which inherently involves time, and the relationships between leaders and followers also occur over time. It takes time to be seen by others as a leader and to influence others, as well as to develop leadership competencies and mind-sets through LDPs and other methods. It takes time to learn and to apply that learning at work. Therefore, any attempt to understand the outcomes of such a process must incorporate temporal variables in evaluations; time lags and effects of time on programme outcomes, if any, must specifically be addressed in LDP evaluation. In this paper, we highlight that time has been relatively overlooked in most evaluation studies, and questions, such as ‘*when* outcomes emerge’ and ‘what happens to the reported outcomes with a passage of time’, have not been fully examined in LDP evaluation literature. Considering the many theoretical, practical and policy-related implications, we argue that there is merit in incorporating temporal dimensions in LDP evaluation, and in adopting a multi-temporal view of outcomes, so that a programme’s actual value can be identified and more targeted investments can be made to develop leaders and leadership in contemporary organisations. We have taken a first step in articulating this case with a view to provide a stimulus for incorporating time in LDP evaluation studies. We believe that an explicit consideration of the temporal dimension of outcomes in future LDP evaluation studies would improve the quality of LDP outcomes research and advance LDP evaluation theory, policy and practice.

References

- Ancona, Deborah G., Paul S. Goodman, Barbara S. Lawrence, and Michael L. Tushman. 2001. "Time: A new research lens." *Academy of Management Review* 26, no. 4: 645-663.
- Ancona, Deborah G., Gerardo A. Okhuysen, and Leslie A. Perlow. 2001. "Taking time to integrate temporal research." *Academy of Management Review* 26, no. 4: 512-529.
- Ancona, Deborah, and Chee-Leong Chong. "Entrainment: Pace, cycle, and rhythm in organizational behavior." 1996. *Research in Organisational Behaviour*, VOL 18, 1996 18: 251-284.
- Åkerlind, Gerlese S. 2005. "Variation and commonality in phenomenographic research methods." *Higher Education Research & Development* 24, no. 4: 321-334.
- Aragón-Sánchez, Antonio, Isabel Barba-Aragón, and Raquel Sanz-Valle. 2003. "Effects of training on business results1." *The International Journal of Human Resource Management* 14, no. 6: 956-980.
- Archer, Stephen. 2013. "Leadership development—can there be a return on investment?" *Development and Learning in Organizations: An International Journal*, 27, no. 3: 18-21.
- Ardichvili, Alexandre, Kristina Natt och Dag, and Steven Manderscheid. 2016. "Leadership development: Current and emerging models and practices." *Advances in Developing Human Resources* 18, no. 3: 275-285.
- Avolio, Bruce J., James B. Avey, and David Quisenberry. 2010. "Estimating return on leadership development investment." *The Leadership Quarterly* 21, no. 4: 633-644.
- Avolio, Bruce J., Rebecca J. Reichard, Sean T. Hannah, Fred O. Walumbwa, and Adrian Chan. 2009. "A meta-analytic review of leadership impact research: Experimental and quasi-experimental studies." *The Leadership Quarterly* 20, no. 5: 764-784.
- Baldwin, Timothy T., and J. Kevin Ford. 1988. "Transfer of training: A review and directions for future research." *Personnel Psychology* 41, no. 1: 63-105.
- Bee, Frances, and Roland Bee. 2003. *Learning needs analysis and evaluation*. Chartered Institute of Personnel and Development, London.
- Beer, Michael, Magnus Finnström, and Derek Schrader. 2016. "Why leadership training fails—and what to do about it." *Harvard Business Review* 94, no. 10: 50-57.
- Berson, Josh. 2019. "Why Leadership Development Feels Broken: And How We're Fixing It". <https://joshbersin.com/2019/07/why-leadership-development-feels-broken-and-how-were-fixing-it/>

- Birdi, Kamal. 2010. "The Taxonomy of Training and Development Outcomes (TOTADO): A New Model of Training Evaluation", *Annual BPS Division of Occupational Psychology Conference*, Brighton, UK.
- Black, Alice M., and Garee W. Earnest. 2009. "Measuring the outcomes of leadership development programs." *Journal of Leadership & Organizational Studies* 16, no. 2: 184-196.
- Blakeley, Karen, and Malcolm Higgs. 2014. "Responsible leadership development—crucible experiences and power relationships in a global professional services firm." *Human Resource Development International* 17, no. 5: 560-576.
- Bluedorn, Allen C., and Rhetta L. Standifer. 2006. "Time and the temporal imagination." *Academy of Management Learning & Education* 5, no. 2: 196-206.
- Bluedorn, Allen C., and Kimberly S. Jaussi. 2007. "Organizationally relevant dimensions of time across levels of analysis." *Research in Multi-Level Issues*, no. 6: 187-223.
- Bluedorn, Allen C., and Kimberly S. Jaussi. 2008. "Leaders, followers, and time." *The Leadership Quarterly* 19, no. 6: 654-668.
- Boak, George, and Sarah Crabbe. 2019. "Evaluating the impact of coaching skills training on individual and corporate behaviour." *European Journal of Training and Development*. 40, no. 1:97-106
- Bolden, Richard. 2011. "Distributed leadership in organizations: A review of theory and research." *International journal of management reviews* 13, no. 3: 251-269.
- Burgoyne, John, and Turnbull-James, Kim. 2001. "Leadership development: Best practice guide for organisations", *Council for Excellence in Management and Leadership*, London, UK.
- Burgoyne, John, John Mackness, and Sadie Williams. 2009. "Baseline study of leadership development in higher education." *Leadership Foundation for Higher Education, London*.
- Burgoyne, John, Wendy Hirsh, and Sadie Williams. 2004. *The Development of Management and Leadership Capability and its Contribution to Performance: The evidence, the prospects and the research need*. London: Department for Education and Skills.
- Burke, Lisa A., and Alan M. Saks. 2009. "Accountability in training transfer: Adapting Schlenker's model of responsibility to a persistent but solvable problem." *Human Resource Development Review* 8, no. 3: 382-402.
- Carden, Lila L., and Jamie L. Callahan. 2007. "Creating leaders or loyalists? Conflicting identities in a leadership development programme." *Human Resource Development International* 10, no. 2:169-186.
- Cascio, Wayne, and John Boudreau. 2011. "Investing in people: Financial impact of human resource initiatives". 2nd ed. Pearson Education Inc, Upper Saddle River, NJ.

- Chartered Management Institute. 2014. "The Future Forecast: Priorities and Expectations for 2015", CMI, London.
- Clarke, Nicholas. 2012. "Evaluating Leadership Training and Development: A Levels-of-Analysis Perspective." *Human Resource Development Quarterly* 23, no. 4: 441-460.
- Collins, Doris B., and Elwood F. Holton III. 2004. "The effectiveness of managerial leadership development programs: A meta-analysis of studies from 1982 to 2001." *Human Resource Development Quarterly* 15, no. 2: 217-248.
- Cunliffe, Ann L., and Matthew Eriksen. 2011. "Relational leadership." *Human Relations* 64, no. 11: 1425-1449.
- Davidson, E. Jane. 2005. *Evaluation Methodology Basics: The nuts and bolts of sound evaluation*. Sage.
- Davies, Rick. 2018. "Representing theories of change: Technical challenges with evaluation consequences." *Journal of Development Effectiveness* 10, no. 4: 438-461.
- Day, David V. 2011. "Leadership development." *The SAGE Handbook of Leadership* 22: 37-50.
- Day, David V. 2013. "Training and developing leaders: Theory and research." *The Oxford Handbook of Leadership*, 76-93.
- Day, David V. 2014. "The future of leadership: Challenges and prospects." in *The Oxford Handbook of Leadership and Organizations*, edited by D.V. Day, 859-867, Oxford university press, NY.
- Day, David V., and Lisa Dragoni. 2015. "Leadership development: An outcome-oriented review based on time and levels of analyses." *Annual Review of Organisational Psychology and Organisational Behaviour* 2, no. 1: 133-156.
- Day, David V., John W. Fleenor, Leanne E. Atwater, Rachel E. Sturm, and Rob A. McKee. 2014. "Advances in leader and leadership development: A review of 25 years of research and theory." *The Leadership Quarterly* 25, no. 1: 63-82.
- Day, David V., and Michelle M. Harrison. 2007. "A multilevel, identity-based approach to leadership development." *Human Resource Management Review* 17, no. 4: 360-373.
- Day, David V., Michelle M. Harrison, and Stanley M. Halpin. 2008. "An integrative approach to leader development: Connecting adult development, identity, and expertise". Routledge.
- Day, David V., and Hock-Peng Sin. 2011. "Longitudinal tests of an integrative model of leader development: Charting and understanding developmental trajectories." *The Leadership Quarterly* 22, no. 3: 545-560.

- DeRue, D. Scott, and Christopher G. Myers. 2014. "Leadership development: A review and Agenda for future research." In *Oxford Handbook of Leadership and Organizations*, edited by D.V. Day. 832-54, Oxford university press, Oxford.
- Dexter, Barbara, and Christopher Prince. 2007. "Evaluating the impact of leadership development: a case study." *Journal of European Industrial Training*. 31, no. 8: 609-25.
- Dvir, Taly, Dov Eden, Bruce J. Avolio, and Boas Shamir. 2002. "Impact of transformational leadership on follower development and performance: A field experiment." *Academy of Management Journal* 45, no. 4: 735-744.
- Eagan, Jane. 2011. "Evaluating Learning and Talent Development: Factsheet" CIPD, London, UK.
- Edmondson, Amy C., and Anita Williams Woolley. 2003. "Understanding outcomes of organizational learning interventions." (2003). In *Blackwell Handbook of Organizational Learning and Knowledge Management*, edited by M. Easterby-Smith and M. Lyles. Malden, MA: Blackwell Publishing.
- Edwards, Gareth, Carole Elliott, Marian Iszatt-White, and Doris Schedlitzki. 2015. "Using creative techniques in leadership learning and development: an introduction." *Advances in Developing Human Resources* 17, no. 3: 279-288.
- Edwards, Gareth. and Gill, Roger. 2005. "Evaluating leadership across in organisations" in *Leadership under Microscope*, edited by R Gill, 23-34, The Leadership Trust Foundation, Herefordshire.
- Edwards, Gareth, and Sharon Turnbull. 2013. "Special issue on new paradigms in evaluating leadership development." *Advances in Developing Human Resources* 15, no. 1: 3-9.
- Feser, Claudio, Nicolai Nielsen, and Michael Rennie. 2017. "What's missing in leadership development." *McKinsey Quarterly*, no. 3: 20-24.
- Fairhurst, Gail. 2017. "Leadership process." *The SAGE Handbook of Process Organization Studies*. 497-512, London: Sage.
- Ford, J. Kevin, Timothy T. Baldwin, and Joshua Prasad. 2018. "Transfer of training: The known and the unknown." *Annual Review of Organizational Psychology and Organizational Behaviour*, no. 5: 201-225.
- Ford, Jackie, and Nancy Harding. 2007. "Move over management: We are all leaders now." *Management Learning*, 38, no. 5: 475-493.
- Ford, J. Kevin, and Ruchi Sinha. 2008. "Advances in training evaluation research." In *Oxford Handbook of Personnel Psychology*, 291-316, edited by Cartwright, Susan, and Cary L. Cooper, Oxford Handbooks, Oxford.

- Getha-Taylor, Heather, Jacob Fowles, Chris Silvia, and Cullen C. Merritt. 2015. "Considering the effects of time on leadership development: A local government training evaluation." *Public Personnel Management* 44, no. 3: 295-316.
- Gentry, William A., and Jennifer W. Martineau. 2010. "Hierarchical linear modelling as an example for measuring change over time in a leadership development evaluation context." *The Leadership Quarterly* 21, no. 4: 645-656.
- Giber, David, Samuel Lam, Marshall Goldsmith and Justin Bourke. 2009. "Best practices in Leadership Development Handbook" 2nd edn, Pfeiffer, San Francisco, CA.
- Gill, Roger. 2011. "Theory and practice of leadership" 2nd ed. Sage. Thousand Oaks, CA
- Grove, J., B. Kibel, and Taylor Haas. 2007. "EvaluLEAD: An open-systems perspective on evaluating leadership development" in *The Handbook of Leadership Development Evaluation*, 71-110, edited by K. Hannum, J.W. Martineau and C. Reinelt, Jossey-Bass, San Francisco, CA.
- Gutierrez, M., and T. Tasse. 2007. "Leading with theory: Using a theory of change approach for leadership development evaluations." *The Handbook of Leadership Development Evaluation*, 48-70, edited by K. Hannum, J.W. Martineau and C. Reinelt, Jossey-Bass, San Francisco, CA.
- Hallinger, Philip, and Suparak Suriyankietkaew. 2018. "Science mapping of the knowledge base on sustainable leadership 1990–2018," *Sustainability* 10, no. 12: 4846.
- Han, Heeyoung, and David Boulay. 2013. "Reflections and future prospects for evaluation in human resource development." *New Horizons in Adult Education and Human Resource Development* 25, no. 2: 6-18.
- Han, Feifei, and Robert Ellis. 2019. "Using phenomenography to tackle key challenges in science education." *Frontiers in psychology*, no. 10: 1414.
- Hannah, Sean T., and Bruce J. Avolio. 2010. "Ready or not: How do we accelerate the developmental readiness of leaders?" *Journal of Organizational Behaviour* 31, no. 8: 1181-1187.
- Hezlett, Sarah A. 2016. "Enhancing experience-driven leadership development." *Advances in Developing Human Resources* 18, no. 3: 369-389.
- Hirsh, Wendy, Penny Tamkin, Valerie Garrow, and John Burgoyne. 2011. "Evaluating Management and Leadership Development: New Ideas and Practical Approaches." *Key Research Findings. Brighton: Institute for Employment Studies.*
- Howard, Ann, and Richard S. Wellins. 2008. "Global leadership forecast 2008| 2009: Overcoming the shortfalls in developing leaders." Executive Summary, Pittsburgh, PA.

- Huang, Jason L., J. Kevin Ford, and Ann Marie Ryan. 2017. "Ignored no more: Within-person variability enables better understanding of training transfer." *Personnel Psychology* 70, no. 3: 557-596.
- Hunter, Samuel T., Katrina E. Bedell-Avers, and Michael D. Mumford. 2007. "The typical leadership study: Assumptions, implications, and potential remedies." *The Leadership Quarterly* 18, no. 5: 435-446.
- Nguyen Huy, Q. U. Y. 2001. "Time, temporal capability, and planned change." *Academy of Management Review* 26, no. 4: 601-623.
- Joseph-Richard, Paul and Hazlett, Shirley-Ann. 2014. "Variation in Middle-Managers' Conceptions of Leadership Development: A Phenomenographic Study". Paper presented at the 2014 *Annual Meeting of the Academy of Management*, Philadelphia. Academy of Management, Philadelphia, PA.
- Kaufman, Roger, and John M. Keller. 1994. "Levels of evaluation: Beyond Kirkpatrick." *Human Resource Development Quarterly* 5, no. 4: 371-80.
- Kearns, Paul, and Tony Miller. 1996. "Measuring the Impact of Training and Development on the Bottom Line: An Evaluation Toolkit to Make Training Pay". Technical Communications.
- King, Elizabeth, and Paul Nesbit. 2015. "Collusion with denial: Leadership development and its evaluation." *Journal of Management Development*, 34, no. 2: 134-152
- Kirkpatrick, Donald. 1996. "Great ideas revisited" *Training & Development*, 50, no.1: 54-60.
- Kraiger, Kurt. 2014. "Looking back and looking forward: Trends in training and development research." *Human Resource Development Quarterly* 25, no. 4: 401-408.
- Kruse, Kevin. 2020. Top 5 Leadership Development Trends For 2020.
<https://www.forbes.com/sites/kevinkruse/2020/01/15/top-5-leadership-development-trends-for-2020/#3fb54e101f3a>
- Lacerenza, Christina N., Denise L. Reyes, Shannon L. Marlow, Dana L. Joseph, and Eduardo Salas. 2017. "Leadership training design, delivery, and implementation: A meta-analysis." *Journal of Applied Psychology* 102, no. 12: 1686-1718.
- Leskiw, Sheri-Lynne, and Parbudyal Singh. 2007. "Leadership development: Learning from best practices." *Leadership & Organization Development Journal*, 28, no. 5: 444-64.
- Marton, Ference. 1986. "Phenomenography—a research approach to investigating different understandings of reality." *Journal of thought*, 21, no. 3: 28-49.
- Marton, Ference, and Shirley A. Booth. 1997. *Learning and awareness*. Psychology Press.

- Marton, Ference, Maj Asplund Carlsson, and László Halász. 1992. "Differences in understanding and the use of reflective variation in reading." *British Journal of Educational Psychology* 62, no. 1: 1-16.
- McAlearney, Ann Scheck. 2010. "Executive leadership development in US health systems." *Journal of Healthcare Management* 55, no. 3: 206-222.
- McCauley, Cynthia D., D. Scott DeRue, Paul R. Yost, and Sylvester Taylor. 2013. Experience-driven leader development: Models, tools, best practices, and advice for on-the-job development. John Wiley & Sons.
- McBain, Richard, Abby Ghobadian, Jackie Switzer, Petra Wilton, Patrick Woodman, and Gemma Pearson. 2012. "The business benefits of management and leadership development." *CMI and Penna*.
- McCall Jr, Morgan W. 2004. "Leadership development through experience." *Academy of Management Perspectives* 18, no. 3: 127-130.
- Mitchell, Terence R., and Lawrence R. James. 2001. "Building better theory: Time and the specification of when things happen." *Academy of Management Review* 26, no. 4: 530-547.
- Modisane, Keba T. 2018. "Finding the niche to reposition leadership in Africa's developing economies for the global highway: review of literature on leadership development programmes—methods and techniques." *Human Resource Development International* 21, no. 1: 12-23.
- Mumford, Michael D. 2006. *Pathways to outstanding leadership: A comparative analysis of charismatic, ideological, and pragmatic leaders*. Lawrence Erlbaum Associates Publishers, Erlbaum Press, Mahwah, NJ.
- Mumford, Michael D., Samuel T. Hunter, Dawn L. Eubanks, Katrina E. Bedell, and Stephen T. Murphy. 2007. "Developing leaders for creative efforts: A domain-based approach to leadership development." *Human Resource Management Review* 17, no. 4: 402-417.
- Packard, Thomas, and Loring Jones. 2015. "An outcomes evaluation of a leadership development initiative." *Journal of Management Development*. 34, no. 2: 153-68
- Passmore, Jonathan and Maira Joao Velez. 2014. Training Evaluation. In *the Wiley Blackwell Handbook of the Psychology of Training, Development, and Performance Improvement*, edited by Kraiger, Kurt, Jonathan Passmore, Nuno Rebelo dos Santos, and Sigmar Malvezzi, John Wiley & Sons.
- Pentland, Brian T. 1999. "Building process theory with narrative: From description to explanation." *Academy of Management Review* 24, no. 4: 711-724.
- Peters, Larry, John Baum, and Gregory Stephens. 2011. "Creating ROI in leadership development." *Organizational Dynamics*, 2, no. 40: 104-109.

- Phillips, Jack J., and Patti Phillips. 2007. "Measuring return on investment in leadership development." In *The Handbook of Leadership Development Evaluation*, 137-166, edited by K.M. Hannum, J.W. Martineau and C. Reinelt, John Wiley & Sons Ltd, San Francisco, CA.
- Prosser, Michael. 1994. "Some experiences of using phenomenographic research methodology in the context of research in teaching and learning." *Phenomenographic research: Variations in method*: 31-43. Melbourne: RMIT: EQARD.
- Reyes, Denise L., Julie Dinh, Christina N. Lacerenza, Shannon L. Marlow, Dana L. Joseph, and Eduardo Salas. 2019. "The state of higher education leadership development program evaluation: A meta-analysis, critical review, and recommendations." *The Leadership Quarterly* 30, no. 5: 101-311.
- Richman-Hirsch, Wendy L. 2001. "Post-training interventions to enhance transfer: The moderating effects of work environments." *Human Resource Development Quarterly* 12, no. 2: 105-120.
- Russ-Eft, Darlene. 2007. "Communication in evaluation: A systems approach." *The Handbook of Leadership Development Evaluation*, 487-510, edited by K.M. Hannum, J.W. Martineau and C. Reinelt. John Wiley & Sons Ltd, San Francisco, CA.
- Russ-Eft, Darlene, F. Marcie J. Bober-Michel, Tiffany A. Koszalka, and Catherine M Sleezer. 2014. "Evaluator Competency Field Book". Charlotte, NC: Information Age.
- Russ-Eft, Darlene, and Hallie Preskill. 2009. "Evaluation in organizations: A systematic approach to enhancing learning, performance, and change". Basic Books.
- Sandberg, Jörgen. 2005. "How do we justify knowledge produced within interpretive approaches?" *Organizational Research Methods*, 8, no. 1: 41-68.
- Santos, Amalia, and Mark Stuart. 2003. "Employee perceptions and their influence on training effectiveness." *Human Resource Management Journal* 13, no. 1: 27-45.
- Saks, Alan M., and Lisa A. Burke. 2012. "An investigation into the relationship between training evaluation and the transfer of training." *International Journal of Training and Development* 16, no. 2: 118-127.
- Schriesheim, Chester A. 2003. "Why leadership research is generally irrelevant for leadership development." In *The future of leadership development*, 181-197, edited by Murphy, Susan Elaine, and Ronald E. Riggio, Psychology Press.
- Scott, Kimberly S. 2017. "An integrative framework for problem-based learning and action learning: Promoting evidence-based design and evaluation in leadership development." *Human Resource Development Review* 16, no. 1: 3-34.
- Scriven, Michael. 1991. "Evaluation Thesaurus" 4th edn, Sage: Thousand Oaks, CA.

- Scriven, Michael. 1991. "Pros and cons about goal-free evaluation." *Evaluation Practice* 12, no. 1: 55-62.
- Shamir, Boas. 2011. "Leadership takes time: Some implications of (not) taking time seriously in leadership research." *The Leadership Quarterly* 22, no. 2: 307-315.
- Shipp, Abbie J., and Michael S. Cole. 2015. "Time in individual-level organizational studies: What is it, how is it used, and why isn't it exploited more often?" *Annual Review of Organisational Psychology and Organisational Behaviour*, 2, no. 1: 237-260.
- Simmonds, David, and Ophelia Tsui. 2010. "Effective design of a global leadership programme." *Human Resource Development International*, 13, no. 5: 519-540.
- Sinar, Evan, Wellins, Richard. S., Ray, Rebecca, Abel, Amy Lui, Neal, Stephanie N. 2014. "Ready-Now Leaders: Meeting Tomorrow's Business Challenges-Global leadership Forecast 2014/2015". Development Dimensions International. The Conference Board Inc.
- Sinar, Evan. et al. 2018. "Global leadership forecast 2018" 25 Research insights to fuel your people strategy". Development Dimensions International. The Conference Board Inc.
- Singer, Judith D., John B. Willett, and John B. Willett. 2003. "Applied longitudinal data analysis: Modelling change and event occurrence". Oxford university press, Oxford.
- Smith, Carmelina Lawton. 2015. "How coaching helps leadership resilience: The leadership perspective." *International Coaching Psychology Review* 10, no. 1: 6-19.
- Sonnentag, Sabine. 2012. "Time in organizational research: Catching up on a long-neglected topic in order to improve theory." *Organizational Psychology Review* 2, no. 4: 361-368.
- Sørensen, Peter. 2017. "What research on learning transfer can teach about improving the impact of leadership-development initiatives." *Consulting Psychology Journal: Practice and Research* 69, no. 1: 47.
- Stead, Valerie. 2014. "The gendered power relations of action learning: A critical analysis of women's reflections on a leadership development programme." *Human Resource Development International* 17, no. 4: 416-437.
- Stein, Danielle, and Craig Valters. 2012. "Understanding theory of change in international development." (JSRP Paper 1). Justice and Security Research Programme, International Development Department, London School of Economics and Political Science, London, UK.
- Stenfors-Hayes, Terese, Hakan Hult, and Madeleine A. Dahlgren. 2013. "A phenomenographic approach to research in medical education." *Medical Education* 47, no. 3: 261-270.
- Storey, John. 2004. "Changing theories of leadership and leadership development." In *Leadership in organizations: Current Issues and Key Trends*, edited by Storey, John, 11-38, Psychology Press.

- Tight, Malcolm. 2016. "Phenomenography: The development and application of an innovative research design in higher education research." *International Journal of Social Research Methodology* 19, no. 3: 319-338.
- Tomlinson, Michael, Dermot O'Reilly, and Mike Wallace. 2013. "Developing leaders as symbolic violence: Reproducing public service leadership through the (misrecognized) development of leaders' capitals." *Management Learning*, 44, no. 1: 81-97.
- Trigwell, Keith. 2000. "A phenomenographic interview on phenomenography" In *Phenomenography*, edited by Bowden, J. and Walsh, E. 62–82. Melbourne, RMIT University Press.
- Trust. 2017. "Leadership Development Strategy". Belfast. UK.
- Turnbull, Sharon. 2005. "Evaluation as anthropology: Cultural approaches to evaluating leadership development." in *Leadership under Microscope*, edited by Roger Gill, 101-115, The Leadership Trust Foundation, Herefordshire, UK.
- Tyler, Sheila. 2004. "Making leadership and management development." *Leadership in organizations: Current Issues and Key Trends* edited by Storey, John, 11-38, Psychology Press, London.
- Watkins, Karen E., Ingunn Hybertsen Lysø, and Kathleen deMarrais. 2011. "Evaluating executive leadership programs: A theory of change approach." *Advances in Developing Human Resources* 13, no. 2: 208-239.
- Wentz, Michelle and Toni Hodges. 2005. "Measuring ROI for a leadership mastery program in ROI at work: Best practice case studies from the real world" in *ROI at work: Best-practice case studies from the real world*, edited by Phillips, Jack J., and Patricia Pulliam Phillips. ASTD.
- Weiss, Carol H. 1997. "How can theory-based evaluation make greater headway?" *Evaluation Review* 21, no. 4: 501-524.
- West, Michael, Kirsten Armit, Lola Loewenthal, Regina Eckert, Thomas West, and Alan Lee. 2015. "Leadership and leadership development in health care: the evidence base." The King's Fund, London.
- Wilson, Kerry, and Sheila Corrall. 2008. "Developing public library managers as leaders." *Library Management*, 29, no. 6/7: 473-88.
- Yost, Paul R., and Mary Mannion Plunkett. 2009. "Real time leadership development" Vol. 7. John Wiley & Sons.
- Zimbardo, Philip G., and John N. Boyd. 1999. "Putting time in perspective: A valid, reliable individual-differences metric." *Journal of Personality and Social Psychology* 77, no. 6: 1271-1288.

Table S1. Participants' profile: A purposive sample of 43 managers to ensure maximum variation.

23 Midlevel Nurse Managers							20-General (midlevel non-nurse) healthcare managers						
Living Leadership: Leading for Success Program for Nurse Managers							Living Leadership Program for General Managers						
S. No.	Age	Gender	Managerial Experience	Team Size	Cohort	Department	S. No.	Age	Gender	Managerial Experience	Team Size	Cohort	Department
1	55	M	23	34	1	Neurology	24	52	F	19	20	1	Physiotherapy Services Management
2	36	F	6	7	1	Anaesthetics – Critical Care Team	25	42	F	22	13	1	Human Resources Development
3	37	F	6	36	2	Mental Health	26	50	M	10	22	2	Nursing Learning and Development
4	37	F	15	21	2	Older People Program of Care	27	27	F	3	6	2	Cardiology (Outpatients) Management
5	40	F	13	33	3	Nephrology	28	41	M	6	17	3	Clinical Psychology Services Management
6	37	F	13	27	3	Cancer Services	29	52	M	4	19	3	HRM (Workplace Planning)
7	47	F	26	23	4	Older People Services –Care Home M	30	64	M	35	45	4	Medical Physics Management
8	50	F	31	26	4	General Medicine	31	33	F	22	36	4	Patient, Client Services Management
9	41	F	19	14	5	Anaesthetics – Pain Management	32	54	F	35	26	5	Patient, Client Services Management
10	53	F	15	52	5	Brain Injury Services	33	37	F	16	21	5	Anaesthetics Services Management
11	47	F	22	36	6	Twilight coordinator – Community Nursing	34	37	F	17	13	6	Medical Workforce Planning (HRM)
12	46	F	15	28	6	Orthopaedics - Surgery	35	65	F	41	8	6	HRM (Governance)
13	59	F	30	23	7	Women's Health – Obs&Gyn (General)	36	48	F	28	46	7	Community Health Services
14	58	F	30	32	7	Orthopedics - Surgery	37	48	F	19	34	7	Health Promotion (HIV prevention) Services
15	48	F	24	18	8	Women's Health – Obs&Gyn (Special Ward)	38	54	F	22	43	8	Governance & Audit Services
16	49	F	10	43	8	Theatres	39	46	F	26	12	8	Equality & Diversity Management
17	47	F	30	20	9	Women's Health – Obs&Gyn (Outpatients)	40	30	F	9	17	9	Health Promotion (Diet & Lifestyle Choices)
18	47	M	26	35	9	Orthopaedics -Outpatients	41	50	F	7	19	9	Health Promotion (Smoking Cessation)
19	45	M	15	27	10	Mental health Services – Shannon Clinic	42	54	M	10	22	10	Governance & Audit Services
20	47	M	26	26	10	Outpatients-Specialist Services	43	52	F	19	20	10	Information Communication Technology
21	53	F	20	28	11	Women's Health – Obs&Gyn (Outpatients)							
22	44	M	17	25	11	Mental Health - Adult							
23	36	M	6	18	12	Mental Health - Children							

Programs: Two LDPs; Variation in 43 Participants: Age range: 27 – 65 years; Gender: Female – 74% and Male – 26%; Managerial experience: 3-35 years; Team Size: 6 – 52 members; Cohorts: 22 different groups; Departments: 37