

Evaluating Leadership Development In Academia

Roberta Fenech, Higher Colleges of Technology, United Arab Emirates

ABSTRACT

Tertiary education institutions invest heavily in leadership development; taking into consideration the importance put on leadership development evaluating leadership development initiatives is crucial. The purpose of this research study is to evaluate a leadership development programme in a tertiary education institution, at a reaction, learning, behaviour and results level. The Kirkpatrick 4-level model is the theoretical framework of this study. This study contributes to existing literature as it adopts a quantitative approach to evaluating the effectiveness of a leadership development that does not merely settle to understand the participants' perception of the leadership development programme. The main conclusion of this research study is that leadership development evaluation, within and across all four-levels of the Kirkpatrick model, results in consistent and positive scores showing the effectiveness of leadership development programmes in academia at a reaction, learning, behaviour, and results level.

Keywords: Evaluation, Leadership Development, Kirkpatrick's Four-Level Model, Tertiary Education.

INTRODUCTION

Literature on the 21st century higher education institutions uses terminology such as the entrepreneurial university, the corporatization of higher education, and academic capitalism (Abdulla et al., 2022). Academic capitalism reflects the interrelations between markets, states, and higher education institutions. The market logic applied to higher education institutions is a complex phenomenon resulting in the intertwined actions of many actors at multiple levels (Sigahi & Saltorato, 2020). Academic leaders are key actors within this ever evolving and complex context and the development of such leaders is a cornerstone of success. The formal education of academics and non-academics in such institutions is a basis upon which to build leadership skills, however leaders also need to be formed to meet the specific needs of the tertiary education institution in which they operate (West et al., 2016). Tertiary education institutions invest heavily in leadership development; taking into consideration such costs and the importance put on leadership development, as a key player in the success of every institution, evaluating leadership development initiatives is crucial (King & Nesbit, 2015).

The purpose of this research study is to evaluate a leadership development programme in a tertiary education institution, at a reaction, learning, behaviour and results level. The research question is: What is the impact of a leadership development programme in higher education, using Kirkpatrick's theoretical framework of training evaluation? The Kirkpatrick 4-level model Kirkpatrick & Kayser-Kirkpatrick (2014) was selected, notwithstanding the varying opinions about this model that are presented in literature, as it is an established, widely used, and well-

recognized systematic approach to training and development evaluation (Paull et al., 2016). That has been well-adapted to higher education. In addition, it can provide a way to contextualize both short-term and long-term organizational outcomes (Yi et al., 2020).

Notwithstanding the strategic importance of leadership development, few organizations adequately evaluate the effectiveness of programs or their impact on performance (King & Nesbit, 2015). The same may be said of education institutions as evidence on the effectiveness of leadership development in education institutions is limited (Zeggelaar et al., 2020). This study contributes to existing literature as it adopts a quantitative approach to evaluating the effectiveness of a leadership development programme in a tertiary education institution using a tool that does not merely settle to understand the participants' perception of the leadership development programme. The latter, according to Martineau (2004), is what most evaluation techniques seek to measure and simply stop at the surface of participants' perceptions.

LITERATURE REVIEW

Leadership Development Evaluation

The evaluation of leadership development is not common. Reasons vary including the time and funding required for such practices as well as the gap in literature on the subject. Burn & Waring (2022) state that the quality of the few studies on the evaluation of leadership development is also of concern due to small sample sizes, lack of underpinning theory, survey instruments with inadequate reliability and validity, failure to measure important control variables, cross sectional designs, reliance on self-report and poor measurement of leadership. Hence there is a drive for leadership development evaluation to extend beyond the simple impact on the individuals participating in such development and to extend to organisational, industry and societal impact (Packard & Jones, 2015). The following is an account of recent studies carried out on the evaluation of leadership development programmes.

Cohrs et al. (2020) evaluated a two-day leadership development programme carried out in three companies in the manufacturing and accounting sectors. The methodology adopted was of a pre-test and post-test control group design and therefore included both a control and experimental group. Online surveys measuring transformational leadership and communication style were used to measure the impact of leadership development. Results of this study show an improvement in both transformational leadership and communication skills for leaders who attended the leadership development programme.

Zulfqar et al. (2021) evaluated a leadership development training programme designed for academics using Bloom's taxonomy as a framework. They focused on the remembering, understanding and application levels in evaluating their leadership programme. An experimental research design was also adopted however using interviews as a research tool. Indicator verbs were used to evaluate the outcomes of the leadership intervention. Findings showed that participants in the leadership development programme increased their awareness on leadership as well as adopted new leadership behaviour as a result of the development sessions. Positive behavioural changes were also noted for all the 6 dimensions of transformational leadership.

Burn & Waring (2022) in a study on the challenges of evaluating leadership development claim that one of the challenges of evaluating leadership development is the presence of many confounding variables when investigating the impact of leadership development on the leadership of participants. They explain the difficulty to isolate the effects of a leadership programme and in turn attribute these to the programme itself. Causation is definitely not a linear process. Burn & Waring (2022) conclude that we need to discover theory-based methods of evaluation that acknowledge and work with this complexity. The theory-based model that acknowledged the complexity of evaluation adopted in this study is the Kirkpatrick (1959) Model.

Kirkpatrick's Model

The Kirkpatrick (1959) was developed to provide managers with a tool to evaluate training outcomes and to this day remains part of all human resource management curricula for undergraduate students. There are four levels in this model, namely reaction, learning, behaviour and results. Participant's level of satisfaction and interest is measured at the reaction level, whilst skills and knowledge learnt is measured at the learning level. The third level is the behavioural level which assesses the application of what is learnt and finally the fourth level assesses the effect or results of such training on the organization (Paull et al., 2016).

Although widely used, and has served as an inspiration to many other tools for evaluation, this model has had its fair share of criticism Alsalamah & Callinan (2021) write about the criticisms around the assumptions made by this model. One assumption that has received criticism is that the levels are hierarchical with the fourth level (namely, results) having the most value. Another assumption that has been criticized is the causal link between levels with claims that such levels may be both simultaneous and/or distinct. A negative correlation between the levels has also been suggested. Others criticize the model for its emphasis on outcomes and not process. Despite all these criticisms it does remain a standard in the business and education field (Waddill, 2006; Alsalamah & Callinan, 2021).

Its wide spread use is also due to the praise it has received over the years. The Kirkpatrick (1959) add's value to training and development evaluation as it does not limit itself to attendance by also including the quality of the experience as the evaluation also measures application and results (Quinton et al., 2002). Cooley et al. (2015) state that the transfer or learning as well as the intention to transfer learning are important goals of all training and development programmes. These are measured by the Kirkpatrick four-level model indicating the value of such model.

The Kirkpatrick (1959) model has already been used to evaluate leadership development. Meta-analytical evidence, using the evaluation criteria of reaction, learning, transfer, and results, shows that leadership development is effective, resulting in improvements in leaders. These improvements have been narrowed down to improvements in skills and self-efficacy, whilst also resulting in positive effects on the team members, such as productivity, performance improvements, innovation, and improved health and safety at work.

In a recent research study on leadership development evaluation using the Kirkpatrick (1959) model. Lantu et al. (2021) also included the evaluations of superiors, co-workers, and

subordinates of the participants. They concluded that evaluating leadership programs is challenging as leadership itself is a complex construct. Therefore, in measuring leadership, understanding the context is essential. Overall, the results of all items in the participant self-rating questionnaire were higher than those by the participant's superiors, co-workers, and subordinates. The divergence of views between participants and others is not an unusual finding as similar outcomes were noted in other studies (Lantu et al., 2021). The methodology adopted in this study evaluation a leadership development program in academia is outlined in the section below.

METHODOLOGY

The leadership development programme being evaluated in this study is one designed using a contextualized social constructivist approach within a systemic theory and emotional intelligence framework. It was developed to train leaders in academia in the United Arab Emirates. This programme consists of six workshops, spread over a two-month period. The topics of these workshops are: self-awareness; leadership interdependence; purposeful feedback; emotionally intelligent and authentic leadership; leading high performing teams; effective communication; building a service culture.

An evaluation tool was designed based on the specific objectives of all six workshops and using the Kirkpatrick 4-level model (Kirkpatrick & Kayser-Kirkpatrick, 2014). Participants filled in the evaluation questionnaire prior to the first workshop and following the final workshop. Generic items were included to evaluate the reaction of participants to the overall leadership development experience. Items were also included evaluating the six workshops at a learning, behaviour and results level. The latter were designed to primarily match the objectives of each workshop (Table 1). This was done based on the recommendation by Lantu et al. (2021) to align the goals of the organisation and evaluative measures. Items were grouped together based on the workshop, for example items evaluating the first workshop at a learning, behaviour and results level were grouped together. A five-point Likert scale was used against 52 statements. Participants in this study read an introductory statement that included the purpose of the study, a statement of confidentiality and anonymity, as well as instructions on how to fill in the questionnaire. The questionnaire was administered on-line both in the first and sixth workshop.

Table 1 MEAN, STANDARD DEVIATION AND VARIANCE OF FINDINGS								
Workshop	Questionnaire Item	Level of Evaluation	Pre-leadership development			Post-leadership development		
			Mean	Standard Deviation	Variance	Mean	Standard Deviation	Variance
Workshop 1	1. I am knowledgeable on leadership	Learning	2.37	0.80	0.64	1.71	0.46	0.21

	interdependence							
	2. I am knowledgeable on the topic of systemic thinking		2.63	1.02	1.04	1.75	0.44	0.19
	3. I am knowledgeable on the topic of leadership styles		2.22	0.74	0.54	1.61	0.50	0.25
	4. I am aware of shadow systems		3.00	0.94	0.89	1.82	0.62	0.37
	5. I apply self-awareness skills at my place of work	Behaviour	2.18	0.78	0.60	1.61	0.50	0.25
	6. I apply systemic thinking in my leadership		2.43	0.89	0.78	1.71	0.60	0.36
	7. I practice self-management skills		2	0.73	0.53	1.86	0.52	0.28
	8. I practice self-awareness	Results	1.96	0.74	0.81	1.82	0.61	0.37
	9. I am confident in my ability to work in a team		1.43	0.50	0.25	1.57	0.50	0.25
	10. I work interdependently with other leaders		1.74	0.71	0.51	1.68	0.61	0.37
Workshop 2	11. I am knowledgeable on the nature of effective and purposeful feedback	Learning	1.82	0.68	0.47	1.75	0.52	0.27
	12. I am knowledgeable on the role of feedback in managing underperforming team		1.87	0.58	0.34	1.82	0.48	0.23

	members							
	13. I am knowledgeable on the obstacles in giving feedback online		1.93	0.62	0.38	1.75	0.52	0.27
	14. I apply effective feedback strategies and tools	Behaviour	2.15	0.67	0.44	1.89	0.42	0.18
	15. I apply effective feedback strategies with poor performers		2.22	0.73	0.53	1.89	0.52	0.26
	16. I am effective in giving feedback	Results	2.02	0.71	0.51	1.89	0.42	0.72
	17. I am effective in giving feedback to underperforming team members		2.13	0.86	0.74	1.79	0.42	0.17
Workshop 3	18. I am knowledgeable on emotional intelligence	Learning	2.07	0.74	0.55	1.82	0.55	0.30
	19. I understand the importance of authentic leadership		1.91	0.63	0.40	1.86	0.52	0.28
	20. I apply strategies to increase well-being of those around me at work	Behaviour	1.96	0.59	0.35	1.75	0.52	0.27
	21. I build trusting relationships at work		1.74	0.57	0.33	1.75	0.52	0.27
	22. I exert influence through my		2.07	0.74	0.55	1.81	0.40	0.16

	credibility							
	23. I consider myself to be an emotionally intelligent leader	Results	2.11	0.67	0.45	1.79	0.63	0.40
	24. I am an authentic leader		2.09	0.60	0.36	1.86	0.65	0.42
	25. I am a trusted leader		1.91	0.63	0.39	0.57	0.50	0.25
	26. I am a credible leader		1.91	0.60	0.36	1.75	0.50	0.34
	27. I am an influential leader		2.09	0.76	0.58	1.89	0.57	0.32
Workshop 4	28. I understand the impact of self-awareness on teams	Learning	2.04	0.76	0.58	1.68	0.55	0.30
	29. I understand the impact of other-awareness on teams		2.07	0.77	0.60	1.79	0.57	0.32
	30. I am knowledgeable about different team roles / styles that can be present in a team		2.07	0.57	0.33	1.79	0.50	0.25
	31. I practice team building techniques	Behaviour	2.24	0.75	0.82	1.82	0.55	0.30
	32. I use my knowledge of different team roles / styles that can be present in a team		2.09	0.66	0.44	1.86	0.45	0.20
	33. I leverage collaborative talent in teams		2.11	0.60	0.37	1.86	0.52	0.28
	34. I am able to build high	Results	2.04	0.67	0.44	1.75	0.44	0.19

	performing teams							
Workshop 5	35. I recognize the relevance of empathy in leadership communication	Learning	1.93	0.57	0.33	1.68	0.55	0.30
	36. I am knowledgeable on the topic of leadership communication skills		1.93	0.65	0.42	1.75	0.44	0.19
	37. I engage in effective leadership communication	Behaviour	2.02	0.58	0.34	1.79	0.42	0.17
	38. I am an emotionally intelligent communicator	Results	2.20	0.65	0.43	1.82	0.55	0.30
Workshop 6	39. I am knowledgeable on the way to promote a service culture	Learning	2.28	0.72	0.52	1.79	0.50	0.25
	40. I am knowledgeable of servant leadership		2.52	0.89	0.79	1.86	0.52	0.28
	41. I promote a service culture in my workplace	Behaviour	2.28	0.75	0.56	1.68	0.55	0.30
	42. I manage the performance of my subordinates in adopting a service culture		2.35	0.77	0.59	1.89	1.42	0.17
	43. I empower others through my leadership		2.02	0.68	0.47	1.79	0.42	0.17
	44. I am inclusive of others in my leadership		2.02	0.61	0.38	1.75	0.52	0.27
	45. I am	Results						

	successful in building a service culture		2.28	0.78	0.61	1.85	0.37	0.14
	46. My team are empowered		2.04	0.71	0.50	1.85	0.36	13
	47. I reward/ recognize whoever adopts a service culture		2.04	0.73	0.53	1.86	0.52	0.28
	48. Diversity and inclusion are characteristics of the team I lead		1.87	0.62	0.38	1.75	0.52	0.27
All workshops	49. Leadership development programmes are worth my time	Reaction	1.74	0.71	0.51	1.68	0.55	30
	50. Leadership development programmes are engaging		1.76	0.71	0.51	1.64	0.49	0.24
	51. Leadership development programmes are effective		1.71	0.64	0.41	1.70	0.44	0.19
	52. Leadership development programmes are relevant to my work		1.69	0.67	0.45	1.65	0.52	0.27

A pilot with 6 participants was conducted to assess the validity of the questionnaire. Feedback was collected from each participant. Final questionnaire items were assessed and modified based on the feedback analysis. Modifications were of a linguistic nature to increase clarity and specificity to items. 75 academic leaders across different divisions attended the leadership development workshops and participated in this research study. Data was collected over the five-month period when the workshops were being facilitated.

The different items in the questionnaire aimed at evaluating the workshops at the same level of evaluation, as established by the Kirkpatrick model (reaction, learning, behaviour and results), were tested for internal consistency using Cronbach's α . Cronbach's α provides information on how strongly the responses to a set of questions correlate. Cronbach's α for the items evaluating Reaction is 0.7 (acceptable level of internal consistency), Cronbach's α for the items evaluating learning is 0.9 (excellent internal consistency), Cronbach's α for items

evaluating behaviour is 0.9 (excellent internal consistency), and finally Cronbach's α for the items evaluating results is also 0.9 (excellent internal consistency). Cronbach's α for the questionnaire across all four levels also resulted in an excellent score of 0.9. Following these statistical results, establishing excellent internal consistency, the data gathered from both sets of questionnaires was further evaluated. The findings are shown in the following section on results.

RESULTS

The 75 participants in this research study were academic leaders holding a post-graduate degree related to their specialization. 30% were females. Participants selected a response for each of the 52 items on a Likert scale with 5 points ranging from Strongly Agree to Strongly Disagree. Strongly agree being at the lower end of the scale (value of 1) and strongly disagree being at the highest end of the scale (value of 5).

As shown in Table 1 and Table 2, findings from the pre-leadership development programme questionnaire show a preference for Agreement with statements across all 6 workshops and all 4 levels of evaluation (reaction, learning, behaviour, and results). The preference for this response option of Agreement on the Likert scale, across all 52 items, ranges from 73.91 % of responses (for item 30 in workshop 4 measuring learning) to 41.3% of responses (for item 2 in workshop 1 measuring learning). The mean for all items ranges from 1.43 to 2.63. The preference for the response option of Agreement was mainly given to items measuring Behaviour in Workshop 5 (titled Effective Communication), followed by Behaviour in Workshop 4 (titled Leading High Performing Teams), and Learning in Workshop 4 (titled Leading High Performing Teams). The data from the pre-leadership development programme questionnaire are clustered around the mean as the standard deviations are low ranging from 0.5 to 1. The variance is also low, indicating a small spread between results, ranging from 0.33 to 1.

Table 2 PRE- LEADERSHIP DEVELOPMENT RESPONSE DISTRIBUTION							
Workshop	Item	Level of Evaluation	Strongly Agree	Agree	Neither Agree nor Disagree	Disagree	Strongly Disagree
Workshop 1		Learning	6.52%	63.04%	17.39%	13.04%	0%
			10.87%	41.3%	23.91%	21.74%	2.17%
			11.11%	62.22%	20%	6.67%	0%
			6.52%	21.74%	39.13%	30.43%	2.17%
		Behaviour	11.11%	71.11%	6.67%	11.11%	0%
			8.7%	56.52%	17.39%	17.39%	0%
			19.57%	67.39%	6.52%	6.52%	0%
		Results	28.26%	56.52%	6.52%	8.7%	0%
			56.52%	43.48%	0%	0%	0%
	0		39.13%	50%	8.7%	2.17%	0%

Workshop 2	1	Learning	26.67%	68.89%	2.22%	0%	2.22%
	2		23.91%	65.22%	10.87%	0%	0%
	3		22.22%	62.22%	15.56%	0%	0%
	4	Behaviour	10.87%	67.39%	17.39%	4.35%	0%
	5		13.04%	56.52%	26.09%	4.35%	0%
	6	Results	19.57%	63.04%	13.04%	4.35%	0%
	7		21.74%	52.17%	17.39%	8.7%	0%
Workshop 3	8	Learning	21.74%	52.17%	23.91%	2.17%	0%
	9		24.44%	60%	15.56%	0%	0%
	0	Behaviour	19.57%	65.22%	15.22%	0%	0%
	1		32.61%	60.87%	6.52%	0%	0%
	2		21.74%	52.17%	23.91%	2.17%	0%
	3	Results	17.39%	54.35%	28.26%	0%	0%
	4		13.33%	64.44%	22.22%	0%	0%
	5		23.91%	60.87%	15.22%	0%	0%
	6		22.22%	64.44%	13.33%	0%	0%
	7		20%	55.56%	20%	4.44%	0%
Workshop 4	8	Learning	19.57%	63.04%	10.87%	6.52%	0%
	9		19.57%	60.87%	13.04%	6.52%	0%
	0		10.87%	73.91%	13.04%	2.17%	0%
	1	Behaviour	13.04%	60.87%	17.39%	6.52%	2.17%
	2		13.04%	69.57%	13.04%	4.35%	0%
	3		10.87%	69.57%	17.39%	2.17%	0%
	4	Results	19.57%	56.52%	23.91%	0%	0%
Workshop	5	Learning	19.57%	67.39%	13.04%	0%	0%

5							
	6		21.74%	65.22%	10.87%	2.17%	0%
	7	Behaviour	13.33%	73.33%	11.11%	2.22%	0%
	8	Results	10.87%	60.87%	26.09%	2.17%	0%
Workshop 6	9	Learning	8.7%	60.87%	23.91%	6.52%	0%
	0		8.7%	47.83%	26.09%	17.39%	0%
	1	Behaviour	8.7%	63.04%	19.57%	8.7%	0%
	2		8.7%	56.52%	26.09%	8.7%	0%
	3		19.57%	60.87%	17.39%	2.17%	0%
	4		17.39%	63.04%	19.57%	0%	0%
	5	Results	10.87%	58.7%	21.74%	8.7%	0%
	6		20%	57.78%	20%	2.22%	0%
	7		19.57%	60.87%	15.22%	4.35%	0%
	8		23.91%	67.39%	6.52%	2.17%	0%
All workshops	9	Reaction	41.3%	43.48%	15.22%	0%	0%
	0		37.78%	51.11%	8.89%	2.22%	0%
	1		34.78%	54.35%	10.87%	0%	0%
	2		42.22%	46.67%	11.11%	0%	0%

The results for the pre-leadership development programme questionnaire fit the profile of the participants. Participants are academics in leadership positions holding a post-graduate level of education for which reason it was not unexpected that the results would be skewed towards the Agree option on the 5-point Likert Scale. However results from the post-leadership development programme questionnaire still show an improvement as results are further skewed towards the Agree and Strongly Agree points on the Likert Scale denoting an increase in the level of agreement for all workshops and across all levels of evaluation. On average there is an increase of 6.5% between participants choosing Strongly Agree in the post-leadership development programme questionnaire and the same participants choosing Strongly Agree in the pre-leadership development programme questionnaire. There is also a decrease in participants opting for the Neither Agree or Disagree, Disagree, and Strongly Disagree points on the Likert

Scale in the post-leadership development programme questionnaire when compared to the pre-leadership development programme questionnaire (Table 3).

Table 3 POST- LEADERSHIP DEVELOPMENT RESPONSE DISTRIBUTION							
Workshop	Item	Level of Evaluation	Strongly Agree	Agree	Neither Agree nor Disagree	Disagree	Strongly Disagree
Workshop 1		Learning	28.57%	71.43%	0%	0%	0%
			25%	75%	0%	0%	0%
			39.29%	60.71%	0%	0%	0%
			28.57%	60.71%	10.71%	0%	0%
		Behaviour	39.29%	60.71%	0%	0%	0%
			35.71%	57.14%	7.14%	0%	0%
			21.43%	71.43%	7.14%	0%	0%
		Results	28.57%	60.71%	10.71%	0%	0%
			42.86%	57.14%	0%	0%	0%
	0		39.29%	53.57%	7.14%	0%	0%
Workshop 2	1	Learning	28.57%	67.86%	3.57%	0%	0%
	2		21.43%	75%	3.57%	0%	0%
	3		28.57%	67.86%	3.57%	0%	0%
	4	Behaviour	14.81%	81.48%	3.7%	0%	0%
	5		18.52%	74.07%	7.41%	0%	0%
	6	Results	14.29%	82.14%	3.57%	0%	0%
	7		21.43%	78.57%	0%	0%	0%
Workshop 3	8	Learning	25%	67.86%	7.14%	0%	0%
	9		21.43%	71.43%	7.14%	0%	0%
	0	Behaviour	28.57%	67.86%	3.57%	0%	0%
	1		28.57%	67.86%	3.57%	0%	0%
	2		18.52%	81.48%	0%	0%	0%
		Results	28.57%	67.86%	0%	3.57%	0%

	3						
	4		28.57%	57.14%	14.29%	0%	0%
	5		42.86%	57.14%	0%	0%	0%
	6		32.14%	60.71%	7.14%	0%	0%
	7		21.43%	67.86%	10.71%	0%	0%
Workshop 4	8	Learning	35.71%	60.71%	3.57%	0%	0%
	9		28.57%	64.29%	7.14%	0%	0%
	0		25%	71.43%	3.57%	0%	0%
	1	Behaviour	25%	67.86%	7.14%	0%	0%
	2		17.86%	78.57%	3.57%	0%	0%
	3		21.43%	71.43%	7.14%	0%	0%
	4	Results	25%	75%	0%	0%	0%
Workshop 5	5	Learning	35.71%	60.71%	3.57%	0%	0%
	6		25%	75%	0%	0%	0%
	7	Behaviour	21.43%	78.57%	0%	0%	0%
	8	Results	25%	67.86%	7.14%	0%	0%
Workshop 6	9	Learning	25%	71.43%	3.57%	0%	0%
	0		21.43%	71.43%	7.14%	0%	0%
	1	Behaviour	35.71%	60.71%	3.57%	0%	0%
	2		14.29%	82.14%	3.57%	0%	0%
	3		21.43%	78.57%	0%	0%	0%
	4		28.57%	67.86%	3.57%	0%	0%
	5	Results	15.38%	84.62%	0%	0%	0%
	6		14.81%	85.19%	0%	0%	0%
	7		21.43%	71.43%	7.14%	0%	0%

	8		28.57%	67.86%	3.57%	0%	0%
All workshops	9	Reaction	35.71%	60.71%	3.57%	0%	0%
	0		35.71%	64.29%	0%	0%	0%
	1		25%	75%	0%	0%	0%
	2		28.57%	67.86%	3.57%	0%	0%

As shown in Table 3, in the post-leadership development questionnaire the preference for the response option of Strongly Agree on the Likert scale, across all 52 items, ranges from 39% of responses (for item 5 in workshop 1 measuring Behaviour) to 15% of responses (for item 46 in workshop 6 measuring Results). The preference for the response option of Agree on the Likert scale, across all 52 items, ranges from 85% of responses (for item 46 in workshop 6 measuring Results) to 57% of responses (for item 6 in workshop 1 measuring Behaviour; item 24 and 24 in workshop 3 measuring results). The mean for all items ranges from 1.57 to 1.89. The data from the post-leadership development programme questionnaire are clustered around the mean as the standard deviations are very low ranging from 0.4 to 0.6. The variance is also very low, indicating a small spread between results, ranging from 0.2 to 0.4. The standard deviation and variance are lower for the findings of the post-leadership development programme questionnaire when compared to the already low ones of the pre-leadership development programme questionnaire. There is a homogeneity of scores across all levels of evaluation in both the pre- and post- leadership development programme questionnaire. These findings are discussed in the next section that also includes interpretations of findings.

DISCUSSION

The main finding of this research study is that whilst the pre-leadership development questionnaire of academics showed a skewed distribution of results towards the agreement of statements reflecting positive leadership development attitudes, knowledge, application, and outcomes, the post-leadership development questionnaire, with the same academics, showed an even stronger skewed distribution of results towards the strong agreement and agreement with these same statements at reaction, learning, behaviour, and results levels. Burn & Waring (2022) in a study on the challenges of evaluating leadership development claim that one of the challenges of interpreting the evaluation of leadership development is the presence of many confounding variables. This challenge will be taken into consideration in interpreting findings.

The pre-leadership development programme questionnaire results reflecting positive leadership development attitudes, knowledge, application, and results are best interpreted in the context of this study. This is in line with the recommendation by Lantu. All participants hold post-graduate degrees in their specialization together with professional recognition for teaching and supporting learning in higher education. Participants engage in yearly professional development as part of the conditions to maintain the latter professional recognition. These

factors may have influenced the results on the pre-leadership development programme questionnaire that are positive across all levels of evaluation. However, following the leadership development programme there was still an improvement in findings across all four levels of evaluation showing that, notwithstanding the initial scores, participants in the leadership development programme improved at the reaction, learning, behaviour, and results level. This result may also be interpreted within context as academics tend to be open to opportunities of training and development. Research results are in accordance with those by Lacerenza, Reyes, Marlow, Joseph, and Salas who also used the evaluation criteria of reaction, learning, transfer, and results and found positive effects on participants following leadership development.

Research results in this study show internal consistency of scores within all four levels of evaluation and across all four levels of reaction, learning, behaviour and results. Alsalamah & Callinan (2021) write that the link between all four levels is an assumption of the Kirkpatrick Model that has been criticized. It has been criticized on the basis that such levels may be distinct and may also be negatively correlated. Such criticism is not supported by this research study as the results across all levels show a good level of internal consistency and therefore a clear correlation between the results at all four levels. Recommendations that may be drawn from this study as well as limitations of the study are addressed in the section below which concludes this research paper (Zulfqar et al., 2021).

CONCLUSION

The main conclusion of this research study is that leadership development evaluation, within and across all four-levels of the Kirkpatrick model, results in consistent and positive scores showing the effectiveness of leadership development programmes in academia at a reaction, learning, behaviour, and results level. This may be concluded despite the limitations of this research study. The limitations of this study are of a methodological and logistical nature in that the span of time elapsing between the pre-leadership development questionnaires and the post-leadership development questionnaires may be considered too short to evaluate the leadership development programme at a behaviour and results level. The time span of five months was selected for this study as it matches the time span of the actual leadership development programme. The reasoning behind this is that retention of research participants may decrease if a more longitudinal approach is adopted and that research participants would drop out from the research study if contacted following the end of the leadership development programme. Another limitation of this research study is that only the perspective of the leadership development participants was sought. It would also be interesting to get the view of the participant's superiors, co-workers, and subordinates.

Notwithstanding the above limitations, recommendations may still be drawn from this research study. The main recommendation is to allocate resources to leadership development evaluation. The value of evaluation is such that it should not be considered as an appendix to leadership development programmes but as a very valuable exercise in itself that provides valuable information to all strata of management in an organisation. Incentives should be given to employees who cooperate with evaluation efforts.

A final recommendation is to evaluate leadership development programmes at high levels of evaluation that illustrate the impact on the organisation and the multiplier effect of such programmes. The multiplier effect of leadership development programmes in academia may be measured by considering their effect not only on the organisational aspect, however also on the educational and societal aspect.

REFERENCES

- Abdulla, A., Fenech, R., Kinsella, K., Hiasat, L., Chakravarti, S., White, T., & Rajan, P. B. (2022). Leadership development in academia in the UAE: creating a community of learning. *Journal of Higher Education Policy and Management*, 1-17.
- Alsalamah, A., & Callinan, C. (2021). The Kirkpatrick model for training evaluation: bibliometric analysis after 60 years (1959–2020). *Industrial and Commercial Training*, 54(1), 36-63.
- Burn, E., & Waring, J. (2022). The evaluation of health care leadership development programmes: a scoping review of reviews. *Leadership in Health Services*, (ahead-of-print).
- Cohrs, C., Bormann, K. C., Diebig, M., Millhoff, C., Pachocki, K., & Rowold, J. (2020). Transformational leadership and communication: Evaluation of a two-day leadership development program. *Leadership & Organization Development Journal*.
- Cooley, S. J., Cumming, J., Holland, M. J. G., & Burns, V. E. (2015). Developing the Model for Optimal Learning and Transfer (MOLT) following an evaluation of outdoor groupwork skills programmes. *European Journal of Training and Development*, 39(2), 104-121.
- King, E., & Nesbit, P. (2015). Collusion with denial: Leadership development and its evaluation. *Journal of Management Development*, 34(2), 134-152.
- Kirkpatrick, D. (1959). Techniques for evaluating training programs. *Journal of the American Society for Training and Development*, 13 (11), 3-9.
- Kirkpatrick, J. & Kayser-Kirkpatrick, W. (2014). *The Kirkpatrick four levels: A fresh look after 55 years*. Ocean Cityq: Kirkpatrick Partners.
- Lantu, D.C., Labdhagati, H., Razanaufal, M.W., & Sumarli, F.D. (2021). Was the training effective? Evaluation of managers' behavior after a leader development program in Indonesia's best corporate university. *International Journal of Training Research*, 19(1), 77-92.
- Martineau, J. (2004). Laying the groundwork: First steps in evaluating leadership development. *Leadership in Action: A Publication of the Center for Creative Leadership and Jossey-Bass*, 23(6), 3-8.
- Packard, T., & Jones, L. (2015). An outcomes evaluation of a leadership development initiative. *Journal of Management Development*, 34(2), 153-168.
- Paull, M., Whitsed, C., & Girardi, A. (2016). Applying the Kirkpatrick model: Evaluating an'interaction for learning framework'curriculum intervention. *Issues in Educational Research*, 26(3), 490-507.
- Quinton, M. L., Tidmarsh, G., Parry, B. J., & Cumming, J. (2022). A Kirkpatrick model process evaluation of reactions and learning from my strengths training for life. *International Journal of Environmental Research and Public Health*, 19(18), 11320.
- Sigahi, T. F. A. C., & Saltorato, P. (2020). Academic capitalism: distinguishing without disjoining through classification schemes. *Higher Education*, 80, 95-117.
- Waddill, D. D. (2006). Action e-learning: An exploratory case study of action learning applied online. *Human Resource Development International*, 9(2), 157-171.
- West, M., Smithgall, L., Rosler, G., & Winn, E. (2016). Evaluation of a nurse leadership development programme. *Nursing Management*, 22(10).
- Yi, Z. M., Zhou, L. Y., Yang, L., Yang, L., Liu, W., Zhao, R. S., & Zhai, S. D. (2020). Effect of the international pharmacy education programs: A pilot evaluation based on Kirkpatrick's model. *Medicine*, 99(27).

- Zeggelaar, A., Vermeulen, M., & Jochems, W. (2022). Evaluating effective professional development. *Professional Development in Education*, 48(5), 806-826.
- Zulfqar, A., Valcke, M., Quraishi, U., & Devos, G. (2021). Developing academic leaders: evaluation of a leadership development intervention in higher education. *SAGE Open*, 11(1), 2158244021991815

Received: 04-Mar-2023, Manuscript No. JMIDS-23-13334; **Editor assigned:** 06-Mar-2023, Pre QC No. JMIDS-23-13334 (PQ); **Reviewed:** 20-Mar-2023, QC No. JMIDS-23-13334; **Revised:** 22-Mar-2023, Manuscript No. JMIDS-23-13334(R); **Published:** 29-Mar-2023