JOURNAL OF TECHNICAL EDUCATION AND TRAINING
ISSN: 2229-8932 e-ISSN: 2600-7932

JTET

Vol. 17 No. 4 (2025) 233-250
https://publisher.uthm.edu.my/ojs/index.php/jtet

Assessing TVET Leadership Development Programmes: Implications to TVET Leaders

Nor Hazana Abdullah[1], Chin Zhi Hong[1]*, Sam Faheeda Abdul Samad[2], Rabiah Ahmad[3]

[1] *Faculty of Technology Management and Business, Universiti Tun Hussein Onn Malaysia, Parit Raja, Batu Pahat, Johor, 86400, MALAYSIA*
[2] *Center of Leadership and Competency, Universiti Tun Hussein Onn Malaysia, Parit Raja, Batu Pahat, Johor, 86400, MALAYSIA*
[3] *Faculty of Engineering Technology, Universiti Tun Hussein Onn Malaysia, Pagoh Campus, MALAYSIA*

*Corresponding Author: zhihong6002@gmail.com
DOI: https://doi.org/10.30880/jtet.2025.17.04.018

Article Info

Received: 15th June 2025
Accepted: 4th October 2025
Available online: 30th December 2025

Keywords

TVET, leadership development, Kirkpatrick model, programme evaluation, competency-based training, higher education

Abstract

Leadership is a pivotal element in advancing the objectives of Technical and Vocational Education and Training (TVET), especially in the context of higher education institutions tasked with driving workforce readiness and institutional transformation. Although leadership development is widely implemented, evaluation practices remain superficial: globally, over 70% of organisations assess training only at Level 1 (reaction), fewer than 20% measure learning at Level 2, and fewer than 10–15% conduct any form of Level 3 behavioural evaluation. Similarly, in the TVET higher education context, most programmes rely solely on post-training satisfaction surveys with minimal follow-up beyond immediate reactions. This study assesses the effectiveness of leadership development programmes at Universiti Tun Hussein Onn Malaysia (UTHM), guided by the Kirkpatrick Training Evaluation Model. The evaluation focused on Levels 1 (Reaction), 2 (Learning), and 3 (Behavior), encompassing 12 leadership programmes under three clusters: ULead, Hi-Lead, and Hi-Per. A descriptive quantitative approach was used, involving post-training surveys, pre- and post-tests, and supervisor feedback three months post-programme. Findings reveal high participant satisfaction (mean scores between 4.30–4.90), statistically significant learning gains across all tested modules ($p < 0.05$), and varying degrees of behavioral changes (7.14% to 53.33%). This study is the first to provide a systematic, multi-level (Levels 1–3) empirical evaluation of leadership development programmes within a Malaysian TVET university, combining post-training satisfaction data, pre- and post-learning assessments, and supervisor-rated behavioural transfer and serve as a basis for leadership program improvement and sustainability.

1. Introduction

Technical and Vocational Education and Training (TVET) has increasingly been recognized as a critical component of national development strategies, especially in countries striving for industrial competitiveness and inclusive

growth. According to UNESCO (2022), TVET is essential for addressing youth unemployment, fostering innovation, and promoting sustainable economic development. In Malaysia, the importance of TVET is underscored by national blueprints such as the Twelfth Malaysia Plan and TVET Empowerment Cabinet Committee Report, which position TVET as a key enabler for producing skilled talent in alignment with Industry 4.0 and the digital economy (Economic Planning Unit, 2021).

Within higher education, TVET universities such as those under the Malaysian Technical University Network (MTUN) play a dual role: delivering advanced technical education and cultivating leadership capable of driving institutional transformation. Universiti Tun Hussein Onn Malaysia (UTHM), as a core MTUN institution, has responded to this mandate through initiatives like the UTHM Leadership Development Framework, aimed at equipping TVET leaders with competencies in strategic thinking, and change leadership.

However, despite growing investment in leadership development programmes, a significant gap remains in evaluating their effectiveness. Over 70% of organisations assess training only at Level 1 (reaction), fewer than 20% measure learning at Level 2, and fewer than 10–15% conduct any form of Level 3 behavioural evaluation (Salas et al., 2012; Tannenbaum et al., 2012; Bates, 2004; Yusof & Yacob, 2020). In the higher education and TVET context, studies similarly report that most programmes rely solely on post-training satisfaction surveys with minimal follow-up beyond immediate reactions (Colella et al., 2019; Paull et al., 2023). This limitation reduces the utility of such programmes in informing succession planning, performance improvement, and institutional capacity-building. To address this issue, the present study applies Kirkpatrick's Training Evaluation Model (Kirkpatrick & Kirkpatrick, 2006) to answer the following research objectives (ROs);

RO_1: To assess trainees' satisfaction with UTHM leadership development programmes using Kirkpatrick's Level 1 (Reaction).

RO_2: To determine the extent of learning gains achieved through the programmes using Kirkpatrick's Level 2 (Learning).

RO_3:To evaluate the behavioural application of leadership competencies in the workplace using Kirkpatrick's Level 3 (Behavior).

By situating the evaluation within the institutional context of UTHM's Leadership Development Framework, this study contributes to the limited but growing literature on leadership capacity-building in TVET. It also responds to calls for more evidence-based approaches to training evaluation and leadership development in higher education (Noe, Clarke, & Klein, 2014; Ahmad & Arifin, 2021), with particular relevance for institutions seeking to build agile, future-ready TVET leaders.

2. Literature Review

The following subsection introduces UTHM Competency-Based Talent Development Framework and UTHM Leadership Talent Development Model as the contextual canvas of this study. It follows with the discussion of Kirkpatrik's Model and previous studies on leadership development programmes effectiveness.

2.1 UTHM Competency-Based Talent Development Framework

The UTHM Competency-Based Talent Development Framework is a structured and strategic approach designed to align the university's human resource development with its long-term aspirations under GTU2030. This framework integrates the principles of Competency-Based Talent Management (CBTM) to ensure that talent development is both systematic and aligned with UTHM's strategic objectives.

J. of Technical Education and Training Vol. 17 No. 4 (2025) p. 233-250

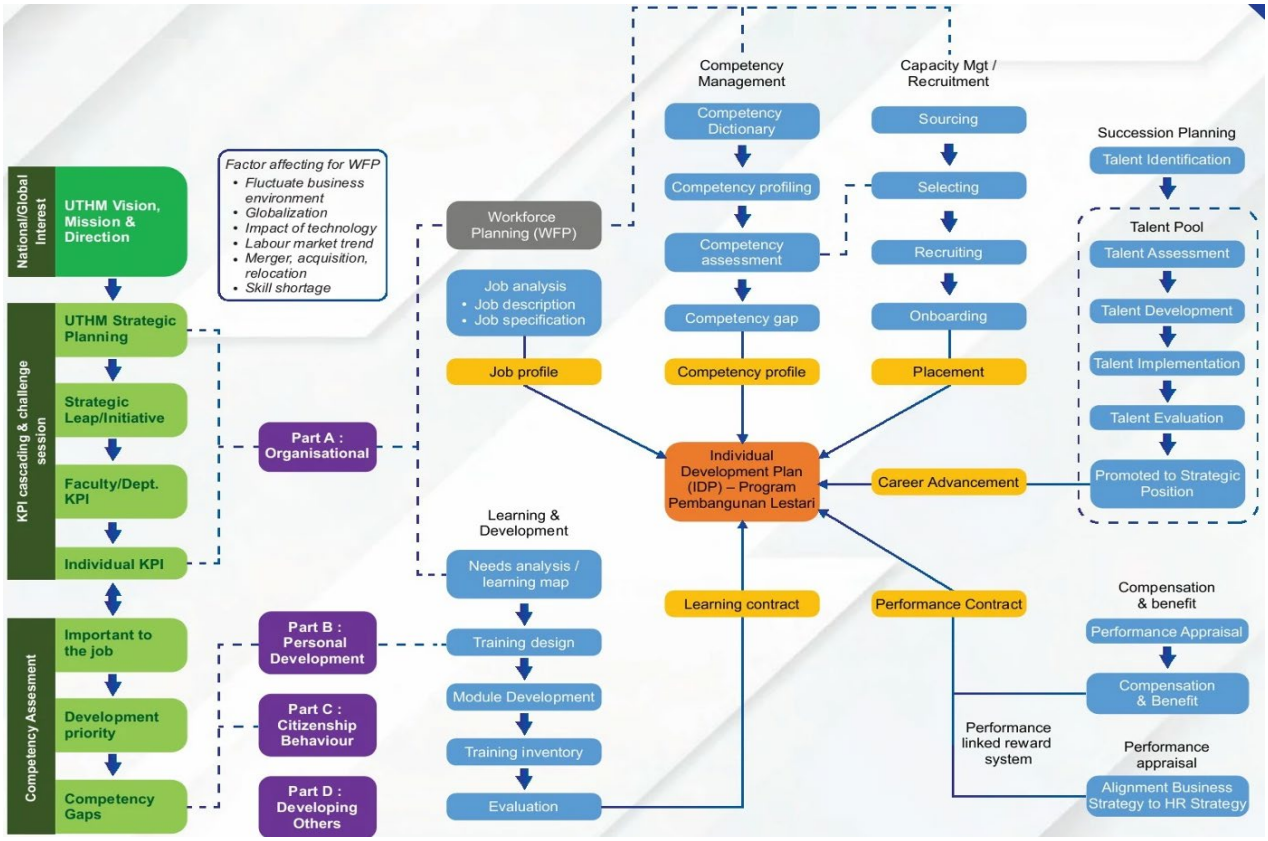

Fig. 1 *UTHM competency-based talent development framework*

A strategic mapping process is implemented to synchronize the university's vision, mission, and strategic plan with the Key Performance Indicators (KPIs) at the faculty, department, and individual levels. This alignment enables clear expectations and responsibilities for each staff member, fostering a culture of accountability and performance. Competency assessment and profiling are conducted based on a well-defined Competency Dictionary, encompassing four main domains:

- Core Competencies – Fundamental skills required across all job roles,
- Generic Competencies – Broader skills that enhance versatility and adaptability,
- Functional Competencies – Specific skills related to job functions, and
- Leadership Competencies – Capabilities essential for effective leadership and decision-making.

Identified competency gaps are used as the foundation for targeted Learning and Development (L&D) initiatives, which are structured according to the 70:20:10 learning model. This model emphasizes:

- 70% experiential learning through job tasks and practical experience,
- 20% social learning through coaching, mentoring, and peer interactions, and
- 10% formal learning via courses and structured training programs.

To address diverse development needs, training programs are segmented into three categories:

- Short-term: Mandatory courses aimed at immediate skill enhancement,
- Mid-term: Internal and external training sessions to broaden knowledge and capabilities, and
- Long-term: Leadership talent development programs to build future leaders.

Each staff member is provided with an Individual Development Plan (IDP), which outlines their learning trajectory and tracks progress towards closing competency gaps. These development initiatives feed into the university's Leadership Talent Pool and contribute to Succession Planning, ensuring that UTHM maintains a pipeline of capable leaders ready to step into key roles as needed. This entire framework is supported by the myLestari System, which integrates learning progress, performance evaluations, and competency assessments.

2.2 UTHM Leadership Talent Development Model

Fig 2 illustrate the UTHM Leadership Talent Development Model which is an essential part of UTHM Competency-Based Talent Development Framework. The leadership development programs (LDPs) are developed corresponding to the hierarchy of programmes for senior leaders of UTHM.

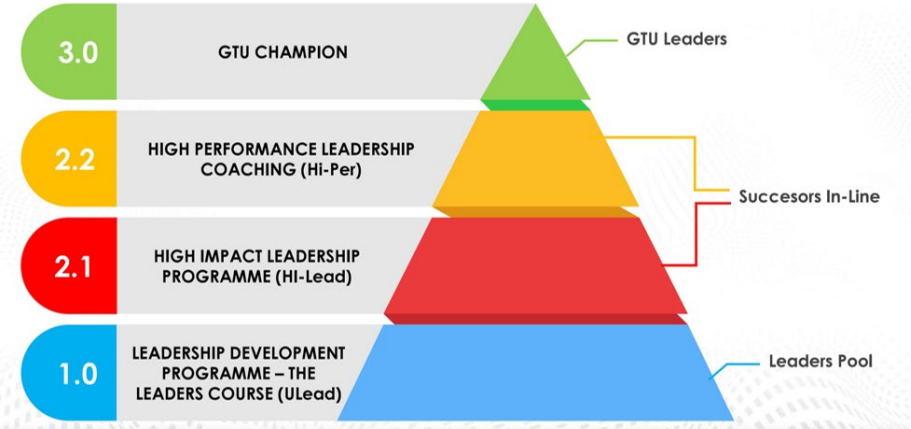

Fig. 2 *UTHM leadership talent development model*

Table 1 showed the mapping of the LDPs with 15 core competencies which are clustered by five domains: Personal Effectiveness, Cognition, Leading, Influence & Impact, and Achievement & Action. Each of these competency domains supports TVET leaders in navigating the increasingly complex demands of higher education institutions with a vocational and technical orientation. For example, in the Personal Effectiveness cluster, competencies such as Integrity, Values & Ethics and Organizational Commitment promote principled decision-making and institutional alignment—both vital in maintaining stakeholder trust and upholding public accountability in vocational education institutions (Dempster, Lovett, & Flückiger, 2011; OECD, 2020). Such values-based leadership is essential as TVET leaders often navigate competing demands from government, industry, and community.

Competencies like Visioning, Creative Thinking, and Business Acumen, in Cognitive domain, allow leaders to anticipate future trends, foster innovation, and strategically manage resources (Day, Fleenor, Atwater, Sturm, & McKee, 2014; UNESCO-UNEVOC, 2022). These are increasingly relevant as TVET institutions are expected to align programmes with emerging industry needs under the IR 4.0 and digital economy agendas (Economic Planning Unit, 2021).

The Leading domainr—comprising competencies such as Teamwork & Team Leadership and Leveraging Diversity—enhances a leader's ability to create inclusive, high-performing teams (Goleman, Boyatzis, & McKee, 2013). In the context of TVET, these competencies support inter-disciplinary collaboration, which is vital for managing practical-based programmes, coordinating with technical staff, and integrating work-based learning.

Similarly, Change Leadership, Networking, Tact & Courtesy, and Stakeholder Engagement are instrumental in facilitating partnerships with industry, accreditation bodies, and community groups (Bolden, 2016; Lizzio, Dempster, & Neumann, 2011). These skills allow TVET leaders to navigate complex ecosystems and lead transformation initiatives effectively, especially when managing curriculum reforms or institutional restructuring.

The Achievement & Action cluster, which includes Achievement Orientation, Initiative, and Stakeholder Focus, supports a culture of performance, continuous improvement, and user-centric leadership (Fernandez, Cho, & Perry, 2010; Gigliotti, 2020). For TVET leaders, these competencies directly affect outcomes such as graduate employability, programme accreditation, and industry responsiveness—key metrics for institutional success.

By systematically developing and evaluating these 15 leadership competencies, TVET leaders become more agile, adaptive, and strategically aligned with national education and workforce development priorities. Embedding such competencies within leadership pipelines also supports long-term institutional sustainability and succession planning, ensuring that leadership continuity is both intentional and competency-driven (Rowe, 2020; Hamzah & Ismail, 2020).

J. of Technical Education and Training Vol. 17 No. 4 (2025) p. 233-250

Table 1 *Mapping of LDPs and leadership competencies*

	List of Leadership Talent Development Programs	Cluster 1: Personal Effectiveness			Cluster 2: Cognition			Cluster 3: Leading			Cluster 4: Influence & Impact			Cluster 5: Achievement & Action		
		Integrity ,Values & Ethics	Organisational Commitment	Self Confidence	Visioning	Conceptual / Creative Thinking	Business Acumen	Teamwork & Team Leadership	Leveraging Diversity	Change Leadership / Adaptability	Impact & Influence	Networking / Relationship Building	Tact & Courtesy	Achievement Orientation	Initiative / Proactive Behavior	Stakeholder Focus
UTHM Leadership Development Programme – Leaders Course (ULead)																
1	Leading with Integrity and Ethics	/														
2	The Rise and Fall of Strategic Planning				/								/	/	/	
3	Stakeholder Engagement: Changes in Practice and Power									/		/			/	/
4	Finance for Non-Finance Managers						/									
5	HR from the Outside In					/										
UTHM High Impact Leadership Program (Hi-Lead)																
1	Change Leadership for GTU Leaders	/	/	/				/		/		/				
2	Holistic Leadership for Senior Administrator			/				/			/			/		
3	Masterclass for Efficient Leader			/				/			/			/		
4	Strategic and Successful Negotiation			/				/			/			/		
UTHM High Performance Leadership Coaching (Hi-Per)																
1	Self-Impact Leadership			/						/						
2	Team Leadership							/				/				/
3	Organizational Leadership		/			/	/		/						/	

2.3 Theoretical Underpinning

The theoretical underpinning of this study is primarily guided by Kirkpatrick's Evaluation Model, which provides a structured approach to assessing training effectiveness through trainees' reactions (Level 1), learning gains (Level 2), and behavioural application in the workplace (Level 3). While Kirkpatrick offers a valuable framework for determining what to evaluate, understanding why learning does or does not translate into behaviour requires deeper theoretical support. Therefore, this study incorporates principles from Baldwin and Ford's (1988) Training Transfer Model and Holton's (1996) HRD Evaluation and Research Model, both of which emphasise the influence of trainee characteristics, training design, motivation to transfer, and workplace environment on behavioural outcomes. Integrating these perspectives strengthens the analytical lens of the current study: high satisfaction and learning outcomes do not necessarily guarantee behavioural change unless conducive transfer conditions are present. Thus, the combined use of Kirkpatrick's evaluation framework with established transfer theories provides a robust theoretical foundation for examining the effectiveness of leadership development programmes across Levels 1 to 3.

2.4 Kirkpatrik's Model

The Kirkpatrick Model was developed by Donald Kirkpatrick and first introduced in 1959 through a series of articles (Kirkpatrick, 1959). It was later refined in 1976 and 1994 for practical use (Kirkpatrick, 1976; Kirkpatrick, 1994). The model consists of four levels for evaluating training: reaction, learning, behavior, and results criteria (Kirkpatrick, 1959, 1976, 1994). It is based on the understanding that evaluating training effectiveness requires multiple levels of analysis to fully capture its impact. A single outcome measure may overlook key factors influencing learning, behavior change, and organizational results (Kirkpatrick, 1959). Although new models have

been proposed (Kirkpatrick Partners, 2016), the four level model of training evaluation and criteria continues to be the most popular and often cited (Bates, 2004).

In the literature, Kirkpatrick's model has served as the foundational framework for many training evaluation models (Bates, 2004; Falletta, 1998). Essentially, the model provides a structured approach for evaluating training programs (Jain et al., 2021; Reio et al., 2017; Saad & Mat, 2013; Ruiz & Snoeck, 2018). Furthermore, it is simple, practical, and efficient to administer when evaluating training (Bates, 2004). The model has gained status as a well-accepted and influential approach in the field of training evaluation (Newstrom, 1995). For example, it has been applied to training programs across sectors, including corporate, governmental, and educational institutions (Kirkpatrick & Kirkpatrick, 2006). The model is a valuable resource for training professionals, particularly in applied contexts (Kaufman, Keller & Watkins, 1995). As Bates (2004) highlights, it enables organizations to conduct meaningful evaluations with limited resources by emphasizing outcome-based measures instead of controlling external factors (Bates, 2004).

The first level of evaluation includes measuring how trainees respond to the training on an emotional and subjective level (Kirkpatrick, 1959, 1976, 1994). In a more recent definition, this level measures the extent to which trainees react with favorable responses to the learning event. (Kirkpatrick Partners, 2016). According to Kirkpatrick (1994), favorable reactions from trainees are an important indicator of the training's potential success, as positive perceptions can influence their engagement and willingness to learn. The common method for evaluating this level involves having trainees complete a post-course survey to express their opinions (Smidt et al., 2009). Furthermore, feedback is typically gathered through surveys or questionnaires at the end of the training session (Zhao et al., 2023). A positive reaction to the training program is essential for trainees to be motivated to learn and apply the skills and knowledge acquired during the training.

Level 2 refers to the extent to which trainees acquire the intended knowledge, skills, attitudes, confidence, and commitment from a training program (Kirkpatrick, 1959, 1976, 1994). This level focuses on measuring the changes in trainees' learning outcomes, which can be achieved through assessments conducted before, during, and after the training. Essentially, learning addresses three primary components: the intended knowledge, skills, attitudes, or competencies that trainees acquire as a result of attending the training program (Kirkpatrick Partners, 2016). The assessment methods for Level 2 often include pre-and post-tests, simulations, or skill demonstrations to objectively determine the level of learning achieved (Kirkpatrick and Kirkpatrick, 2006). For instance, self-evaluation such as written assessments as indicators to measure what trainees have learned (Ruiz et al., 2018; Smidt et al., 2009).

According to Kirkpatrick (1998), Level 3: Behavior is described to the extent to which trainees apply what they learned during the training program to their job or real-world situations (Kirkpatrick, 1959, 1976, 1994). The degree of transfer of learning that has resulted in meaningful changes in the trainees' behavior and performance is significant because positive results cannot be expected unless a positive behavioral change has occurred (Kirkpatrick & Kirkpatrick, 2006). Furthermore, Kirkpatrick (1998) emphasizes the importance of evaluating behavior-level outcomes over time, as it may take several weeks or months for trainees to fully implement the new skills that they learned. The typical method of evaluation involves observing on-the-job behavior and conducting interviews or questionnaires with trainees, supervisors, or peers (Paull et al. 2016; Smidt et al., 2007).

For the purposes of this study, assessment was done on Levels 1 to 3 only, as Level 4—Return on Investment (ROI)—requires longitudinal data and organizational performance tracking over time, which cannot be achieved in a cross-sectional study design. As emphasized by Kirkpatrick and Kirkpatrick (2006) and echoed by Phillips and Phillips (2016), Level 4 assessments necessitate complex outcome attribution, control for external variables, and extended observation periods, making it unsuitable for short-term or one-time evaluations.

2.5 Effectiveness of LDPs Using Kirkpatrick's Levels 1–3

A growing body of empirical research supports the effectiveness of leadership development programs (LDPs) when evaluated through the first three levels of Kirkpatrick's model—reaction, learning, and behaviour. Thus, the following sub sections would focus on previous studies using the model for evaluation LDPs effectiveness.

2.4.1 Level 1: Reaction

Numerous studies report consistently high participant satisfaction following leadership and professional development programmes in higher education and related settings. For instance, Colella et al. (2019) recorded a mean satisfaction score of 9.3/10 among emergency medicine clinicians, while Hall et al. (2024), Hopkins et al. (2017), and Throgmorton et al. (2016) all reported average ratings exceeding 4.3/5. Similar findings are seen in educational leadership programmes, where trainees routinely express that the content is engaging, empowering, and relevant (Franzen, 2020; He & Leeman, 2021).

However, the consistently high satisfaction scores across studies must be interpreted with caution. These ratings derive from heterogeneous instruments and scales, and are potentially influenced by timing (immediate

post-training), cultural response tendencies, and researcher–participant relationships (Hall et al., 2024). More importantly, as the training evaluation literature warns, positive reactions do not necessarily translate into deeper learning or workplace behaviour change (Baldwin & Ford, 1988; Holton, 1996).

The variability observed in some contexts—such as Ketut et al.'s (2022) finding that fewer than half of trainees were satisfied—suggests that trainee satisfaction is sensitive to contextual factors such as relevance of content, facilitation quality, organisational support, and expectations. Taken together, these studies highlight that although satisfaction as in Level 1 (Reaction) is frequently high, it must be complemented by assessments of learning and behaviour to gain a valid picture of programme effectiveness.

2.4.2 Level 2: Learning

Numerous studies evaluating leadership development programmes (LDPs) report substantial improvements in trainees' knowledge, skills, and leadership capabilities. For example, Colella et al. (2019) documented significant gains across 24 competencies, while structured assessments in Masood and Usmani's (2015) programme demonstrated a 77.85% increase in knowledge. Similar positive learning outcomes have been observed across diverse contexts such as school leadership (Alsalamah & Callinan, 2021), primary care settings (Maulina et al., 2023), and women's leadership programmes (He & Leeman, 2021). These findings converge to suggest that LDPs are generally effective in enhancing cognitive and affective learning outcomes.

Beyond knowledge acquisition, several studies report increased confidence and behavioural intentions related to leadership practice (Franzen, 2020; Hopkins et al., 2017; Purwasetiawatik et al., 2024). While these self-perceptions are meaningful indicators of learning, they may not fully reflect actual workplace behaviour change. Much of the evidence relies on self-reported data, which can be subject to overestimation, social desirability bias, and inflated confidence immediately after training (Colella et al., 2019; Maulina et al., 2023).

Additionally, scholars highlight that immediate post-training assessments may not accurately capture long-term retention (Thomas et al., 2018; Ab Rahman et al., 2022). Short-term gains often diminish without reinforcement, coaching, or organisational support, aligning with the broader training literature that emphasizes the critical role of contextual factors in sustaining learning (Baldwin & Ford, 1988; Holton, 1996). Thus, although the evidence shows clear improvements in trainees' knowledge, skills, and confidence, the heavy reliance on self-reported and immediate post-training assessments raises questions about how well these gains are sustained over time

2.4.3 Level 3: Behaviour

Level 3 of Kirkpatrick's model evaluates whether training trainees apply newly acquired knowledge and skills in their work settings. Evidence across leadership development programs shows mixed outcomes. Several studies report positive behavioural transfer, such as improved delegation, communication, and feedback practices (Colella et al., 2019; Faisal-E-Alam et al., 2025), enhanced teamwork (Hopkins et al., 2017), and stronger interpersonal skills (Alsalamah & Callinan, 2021). However, others identify only partial or inconsistent behavioural change. Medical teacher training programmes demonstrated limited transfer without ongoing reinforcement (Masood & Usmani, 2015), and organisational barriers such as lack of departmental alignment and interpersonal tensions reduced the adoption of new behaviours (Lantu et al., 2020). In student leadership contexts, some behaviours persisted while others did not generalise beyond academic settings (Paull et al., 2016).

These discrepancies align with established transfer theories, which emphasise the importance of environmental support, opportunity to perform, and follow-up mechanisms (Baldwin & Ford, 1988; Holton, 1996). Furthermore, Level 3 evaluation is methodologically challenging: many studies rely on self-report or behavioural intentions rather than direct observation or multi-rater assessments, leading to potential overestimation of behavioural change. Overall, the literature suggests that although leadership development programmes frequently produce strong Level 1 and Level 2 outcomes, Level 3 transfer remains inconsistent and highly context-dependent, highlighting the need for more rigorous and long-term evaluation designs.

In summary, the critical analyses of previous studies across Levels 1 to 3 strongly justify the need for the present investigation. Existing literature consistently demonstrates high participant satisfaction and positive learning outcomes, yet behavioural change at the workplace remains variable, context-dependent, and often poorly measured. Many studies rely on self-reported data, immediate post-training evaluations, and limited follow-up, restricting the validity of their findings. Furthermore, comprehensive multi-level evaluations—particularly within higher-education and TVET leadership contexts—are scarce. These gaps in methodological rigour, long-term assessment, and transfer verification highlight the importance and relevance of the current study's systematic evaluation of leadership development programmes using Kirkpatrick's first three levels.

3. Methodology

3.1 Research Design

This study employed a quantitative descriptive evaluation design to assess the effectiveness of TVET leadership development programmes implemented at Universiti Tun Hussein Onn Malaysia (UTHM) under the UTHM Leadership Development Framework. The evaluation was structured using Kirkpatrick's Four-Level Training Evaluation Model (Kirkpatrick & Kirkpatrick, 2006), focusing on Level 1 (Reaction), Level 2 (Learning), and Level 3 (Behavior). These evaluation levels are inherently non-experimental and are designed to assess trainees' perceptions, knowledge gained, and behavioural change within the context of organisational training. Consequently, the use of a control group is not required nor typically feasible, as trainees are existing trainees enrolled based on organisational nomination and leadership pipeline criteria. Random assignment or the creation of a non-training control group would not be ethical or practical in this setting. Moreover, most studies adopting Level 1–3 metrics similarly do not incorporate control groups, as the purpose is programme evaluation rather than causal inference. Therefore, the absence of a control group aligns with accepted methodological practice in training evaluation research. In summary, evaluation data were derived from 12 leadership training programmes conducted throughout 2024, encompassing three levels of leadership development programmes:

- UTHM Leadership Development Program (ULead) – 5 programmes
- High Impact Leadership Program (Hi-Lead) – 5 programmes
- High Performance Leadership Coaching Programme (Hi-Per) – 2 programmes

A total of 370 trainees attended these programmes, with participation ranging from 20 to 45 individuals per course. Details of the LDPs are shown in Table 2.

Table 2 *Leadership development programmes for 2024*

List of UTHM Leadership Development Programs	Modules / Courses	Date	Target Participants	Number of Participants
Leadership Development Program (ULead)	Module 1: Leading with Integrity and Ethics	August 28	All position holders	42
	Module 2: The Rise and Fall of Strategic Planning	Sept 2-3	Administrator Grade 48 and above	42
	Module 3: Stakeholder Engagement – Changes in Practice and Power	Sept 18-19		45
	Module 4: Finance for Non-Finance Managers	August 20		44
	Module 5: HR from the Outside In	August 19		38
High Impact Leadership Program (Hi-Lead)	Academic Leadership Research & Innovation Leadership Student Affairs Leadership	Nov 24-26	Associate Dean / Associate Director / Vice Principal Department Chair / Deputy Principal	42
	Strategic and Successful Negotiation	June 11-13	Administrators at Grade 48 and above	36
	Masterclass for Efficient Leader	June 4-6		30

J. of Technical Education and Training Vol. 17 No. 4 (2025) p. 233-250

List of UTHM Leadership Development Programs	Modules / Courses	Date	Target Participants	Number of Participants
UTHM High Performance Leadership Coaching (Hi-Per)	HiPer Class:Impactful Leadership	Feb 24-27	Deputy Vice Chancellor/ Assistant Vice Chancellor/Deans / Directors	20
	Leadership Forum	June 5		
	Strategic Presentation	June 9		
	Upskilling Your Persona	Jan 30		31

3.2 Sampling Procedure

This study involved trainees who were existing trainees enrolled in the organisation's leadership development programmes. As membership in these programmes is determined entirely by organisational nomination procedures, and HR selection criteria, random or probabilistic sampling was neither practical nor appropriate. Therefore, consistent with prior training evaluation research that used non-probability or convenience sampling when working with intact cohorts (e.g. Lacerenza, Reyes, Marlow & Salas, 2017), this study adopted an intact-group sampling design. A total of 370 trainees participated in this study. All trainees were middle to senior academic and administrative staff involved in TVET leadership functions at UTHM.

3.3 Data Collection Instruments

Data collection instruments varied in each level. For Level 1 (Reaction), this study employed a standardized post-training survey that followed the reaction-level elements recommended by Kirkpatrick and Kirkpatrick (2007), which emphasise participant perceptions of course content, trainer delivery, relevance, facilities, and programme organisation. To operationalise these components, the study adopted the 20-item reaction instrument developed by Pandeani et al. (2020), which measures trainees' perceptions across content relevance, instructor effectiveness, learning environment, and programme support. The original instrument demonstrated excellent internal consistency, with a Cronbach's alpha of 0.948, indicating high reliability. Drawing from this validated scale, item scores were aggregated to produce mean satisfaction ratings representing overall participant reactions to the leadership development programmes. This study did not undertake rigorous validity and reliability testing which is consistent with other evaluation and education studies that employ study-specific reaction or opinion questionnaires for descriptive purposes without formal reliability and validity testing (e.g., Sadeghi, 2012;; O'Keeffe et al., 2017; Sanagavarapu et al., 2024)

In Level 2 (Learning), knowledge acquisition was assessed through pre- and post-tests developed by programme facilitators for the ULead and HiLead modules. The test items were mapped directly to the intended learning outcomes of each module to ensure content relevance and alignment with the curriculum. Consistent with training evaluation practices in literature, pre–post assessments are commonly used to measure short-term cognitive gains and have demonstrated validity in capturing learning improvement (e.g., Masood & Usmani, 2015; Paull et al., 2016). Prior to administration, the items were reviewed for content validity by two programme coordinators to ensure clarity and instructional alignment. Learning gains were quantified by comparing pre- and post-test scores, and the mean differences were analysed using paired-sample t-tests to determine whether the improvements were statistically significant at $p < .05$, indicating meaningful increases in knowledge following programme participation.

For Level 3 (Behaviour), behavioural transfer was assessed three months after programme completion using structured evaluation forms completed by trainees' immediate supervisors. Consistent with Kirkpatrick's (2007) emphasis on workplace application and transfer of learning, supervisors rated the extent to which trainees demonstrated the knowledge, skills, and leadership behaviours targeted in the programme within their actual job roles. Supervisor assessment is widely recommended in training transfer research as it reduces self-report bias and provides a more objective indication of behavioural change (Baldwin & Ford, 1988; Holton, 1996). Responses were compiled as descriptive percentages to reflect the proportion of trainees who exhibited measurable application of training outcomes in the workplace after the three-month period.

4. Findings

Findings of this study is presented based on the Kirkpatrick's levels 1 to 3 in the following subsections;

4.1 Reaction Level

Level 1 evaluation focused on measuring trainees' immediate reactions to the training programmes, capturing their satisfaction with course content, delivery, relevance, and overall training experience. The evaluation was conducted using a standardized post-training survey instrument (*Borang Penilaian Program*) utilizing a 5-point Likert scale, where 1 indicated "Strongly Disagree" and 5 indicated "Strongly Agree". A return rate of post-training survey for all programmes were 100%. The consistently high Level 1 scores (>4.30) also suggest strong alignment between programme content and the competency domains outlined in the UTHM Leadership Development Framework.

Across all twelve leadership programmes conducted in 2024, the Level 1 results demonstrated consistently high satisfaction levels (exceed 4). The mean scores ranged from 4.30 to 4.93, indicating that trainees generally found the programmes to be relevant, engaging, and well-structured. All five modules of the ULead programme recorded strong participant satisfaction as shown in Table 3. Overall, the combined mean satisfaction score across the five ULead programmes was 4.70, indicating a high level of participant satisfaction. ULead focuses on foundational management and administrative competencies designed to strengthen the capabilities of middle-level leaders. Feedback from participants showed that the majority expressed strong appreciation for the programme content, delivery, and its relevance to their day-to-day responsibilities. However, module on Strategic Planning and Lateral Thinking recorded the lowest mean of 4.3 due to poor assessment on the trainer (M=4.2) and content (M=4.1). It was found that the trainer kept referring to UTHM as UTM in his delivery, upsetting the trainees. Trainer's lack of preparation was evident which has serious implication to trainees' satisfaction. It is worth noting that all five modules of the ULead programme were conducted at the same venue, which further highlights the pivotal influence of trainer capability on participant satisfaction. A consistent pattern in the feedback indicates that when trainers demonstrate strong facilitation skills, subject-matter expertise, and positive interpersonal behaviours, participants are more likely to overlook other environmental or logistical limitations. This finding aligns with prior research showing that trainer effectiveness is one of the strongest predictors of positive reaction outcomes and perceived learning, often outweighing factors such as venue, facilities, or administrative arrangements (Alvarez, Salas, & Garofano, 2004; Iqbal & Khan, 2022; Tracey, Tannenbaum, & Kavanagh, 1995). These studies collectively emphasise that high-quality trainers can significantly enhance learner engagement and satisfaction, even when other aspects of the training context are less than ideal.

HiLead similarly recorded a high overall mean satisfaction score of 4.73, reflecting participants' strong positive reactions to the programme. The inclusion of sharing sessions with the respective Deputy Vice Chancellors was particularly well received, as these dialogues provided clearer insights into the strategic direction of each cluster and enabled trainees to relate programme content to their functional leadership roles. Low satisfaction (M=4.2) for "Strategic and Successful Negotiation" program was mainly due to training room which could not accommodate the 36 trainees comfortably especially during groups discussion and games.

HiPer recorded the highest satisfaction level among all programmes, with an overall mean score of 4.93**,** indicating exceptionally strong participant approval. This elevated rating can be attributed to the programme's advanced design, which integrates multiple high-impact learning components. The coaching sessions offered personalised guidance and facilitated deeper self-reflection on leadership behaviours, while the Leadership Forum provided exposure to strategic perspectives and best practices shared by senior leaders and invited experts. In addition, th**e** Strategic Presentation component allowed participants to apply their learning to organisational challenges, receive feedback from top management, and demonstrate their strategic thinking capabilities. Collectively, these elements created a highly engaging and meaningful learning experience that contributed to the programme's outstanding satisfaction ratings.

J. of Technical Education and Training Vol. 17 No. 4 (2025) p. 233-250

Table 3 *Trainees' satisfaction level on leadership development programmes for 2024*

UTHM Leadership Development Framework	Programmes	No Of Trainees	Level 1 Assessment				
			Facilities	Training Content	Training Organizer	Trainer	Overall
LEADERSHIP DEVELOPMENT PROGRAM (ULead)	"HR From the Outside In"	42	4.8	4.8	4.8	4.9	4.83
	"Finance for Non-Finance Managers"	42	4.7	4.6	4.7	4.66	4.67
	"Leading With Integrity and Ethics"	45	4.8	4.9	4.9	4.85	4.86
	"Managing Success with Strategic Planning and Lateral Thinking"	44	4.4	4.1	4.5	4.2	4.3
	"Stakeholder Engagement : Changes in Practice and Power"	38	4.9	4.8	4.8	4.9	4.85
HIGH IMPACT LEADERSHIP PROGRAM (Hi-Lead)	Academic Leadership Research and Innovation Leadership Student Services Leadership	42	4.8	4.8	4.8	4.88	4.82
	"Strategic and Successful Negotiation"	36	4.2	4.7	4.7	4.9	4.63
	Master Class for Efficient Leaders	30	4.7	4.8	4.8	4.9	4.8
UTHM HIGH PERFORMANCE LEADERSHIP COACHING (Hi-Per)	Impactful Self Leadership	20					NA
	Leadership Forum						NA
	Project Presentation						NA
	"Upskilling Your Pesona"	31	4	4.6	4.6	4.7	4.48

4.2 Learning Level

Level 2 of Kirkpatrick's model assesses the degree to which trainees acquire the intended knowledge, skills, and attitudes during the training. In this study, learning was evaluated using a pre-test and post-test method designed for each leadership module, administered before and immediately after the training sessions. The analysis focused on identifying statistically significant improvements in trainees' learning outcomes.

All training sessions except 'Upskilling Your Persona' that administered pre- and post-tests showed a statistically significant increase in scores ($p < 0.05$) as shown in Table 4, suggesting that the training content was effective in enhancing trainees' understanding and mastery of the subject matter. 'Upskilling Your Persona' is a grooming class where trainees were taught about social etiquettes and dress codes in formal setting. This result may be attributed to a ceiling effect, as many trainees already demonstrated a high baseline understanding of workplace grooming standards. Moreover, the nature of the content—emphasizing personal image, etiquette, and affective awareness—may not align with cognitive assessment instruments, a common limitation noted in similar studies (Smidt et al., 2009; Colella et al., 2019).

Table 4 *Trainees' learning on leadership development programmes for 2024*

UTHM Leadership Development Framework	Programmes	Total Trainees	Responses	Learning Gain	t-value	p-value
LEADERSHIP DEVELOPMENT PROGRAM (ULead)	"HR From the Outside In"	42	42	15%	3.76	<0.001
	"Finance for Non-Finance Managers"	42	33	72%	9.3	<0.001
	"Leading With Integrity and Ethics"	45	42	60%	7.52	<0.001
	"Managing Success with Strategic Planning and Lateral Thinking"	44	43	27%	4.13	<0.001
	"Stakeholder Engagement : Changes in Practice and Power"	38	32	75%	33.02	<0.001
HIGH IMPACT LEADERSHIP PROGRAM (Hi-Lead)	Master Class for Efficient Leaders Academic Leadership Research and Innovation Leadership Student Services Leadership	42	41	55%	14.7	<0.001
	"Strategic and Successful Negotiation"	36	36	58%	19.14	<0.001
	Master Class for Efficient Leaders	30	30	81%	31.95	<0.001
UTHM HIGH PERFORMANCE LEADERSHIP COACHING (Hi-Per)	Impactful Self Leadership Leadership Forum Project Presentation	20	Not applicable			
	"Upskilling Your Pesona"	31				

Across the five ULead modules, learning gains ranged substantially from 15% to 75%, with all improvements statistically significant at $p < .001$. The lowest increase was observed for "HR From the Outside In" (15%), while the highest gain emerged from "Stakeholder Engagement: Changes in Practice and Power" (75%). These variations suggest that participant baseline knowledge, content complexity, and instructional method differences may influence the magnitude of learning improvement. This pattern is consistent with findings from leadership development research, where modules closely aligned to participants' immediate job responsibilities or competency gaps generally produce higher learning gains (Colella et al., 2019). Similarly, Masood and Usmani (2015) reported uneven learning improvements across training topics, noting that content-heavy or conceptual modules often show lower short-term gains due to cognitive load and limited opportunities for practice. The present results reinforce this pattern, implying that ULead is more technical or applied modules, such as Finance and Stakeholder Engagement, may more effectively address clear knowledge gaps compared to concept-driven topics that require deeper reflection or longer reinforcement periods.

The HiLead programmes also demonstrated strong and statistically significant learning gains, ranging from 36% to 81%, indicating effective knowledge transfer across intermediate leadership modules. The module "Master Class for Efficient Leaders" produced the largest improvement (81%), followed by substantial gains in "Strategic and Successful Negotiation" (58%) and "Academic Leadership" (55%). These outcomes resemble previous studies in which mid-level leadership programmes yielded high learning improvements because participants were able to immediately connect the training content to operational challenges within their work settings (He & Leeman, 2021). Franzen (2020) similarly observed that leadership modules incorporating practical tools, case discussions, and skill-based exercises generated higher perceived and actual learning. The consistently high learning gains in HiLead reflect the programme's strong alignment between curriculum content and the functional leadership roles of its trainees, reinforcing the well-established principle that relevance, application, and experiential learning are central predictors of Level 2 outcomes (Tannenbaum et al., 2012).

J. of Technical Education and Training Vol. 17 No. 4 (2025) p. 233-250

Learning gain scores were not applicable for the HiPer modules, as these programmes emphasise advanced leadership capability building through coaching, leadership forums, and strategic presentations rather than content-based instruction. This design choice is consistent with the nature of senior leadership development, which typically focuses on behavioural refinement, strategic thinking, and reflective learning rather than knowledge acquisition measurable via pre–post tests. Prior research supports this distinction: senior leadership programmes often demonstrate more meaningful outcomes at Level 3 (Behaviour) rather than Level 2, as behavioural mastery and strategic insight evolve through practice, mentoring, and experiential immersion (Day & Dragoni, 2015). Thus, the absence of Level 2 testing for HiPer appropriately reflects its developmental purpose and aligns with leadership development best practices.

4.3 Behavior Level

Table 5 shows the changes of behavior after three months the trainees went to the LDPs. The response rate for Level 3 evaluation in this study was relatively low, consistent with findings in previous literature where post-training behavioral assessments often face data collection challenges. As Smidt et al. (2009) and Colella et al. (2019) noted, supervisor-based feedback mechanisms tend to yield poor returns due to workload constraints, unclear evaluation responsibilities, and limited follow-up infrastructure. This limitation underscores the need for institutionalizing behavioral follow-up within performance review systems or leveraging alternative data sources such as peer reviews and self-assessment for triangulation. For example, since the trainees of HiPer were all deputy vice chancellor, assistant deputy vice chancellor, deans and directors, their superiors were the Vice Chancellor himself. Thus, the behavioral change as reported by the superior are difficult to gauged. Nonetheless, for those who responded, the changes of behavior were evident and significant among trainees as shown in Table 5.

Table 5 *Trainees' learning on leadership development programmes for 2024*

UTHM Leadership Development Framework	Programmes	No of Trainees	% of Responses	Level 3: Participants had demonstrated changes in behavior
LEADERSHIP DEVELOPMENT PROGRAM (ULead)	"HR From the Outside In"	42	14.29%	100%
	"Finance for Non-Finance Managers"	42	7.14%	100%
	"Leading With Integrity and Ethics"	45	15.56%	100%
	"Managing Success with Strategic Planning and Lateral Thinking"	44	22.73%	100%
	"Stakeholder Engagement : Changes in Practice and Power"	38	13.16%	100%
HIGH IMPACT LEADERSHIP PROGRAM (Hi-Lead)	Master Class for Efficient Leaders	42	20.00%	100%
	Academic Leadership			
	Research and Innovation Leadership			
	Student Services Leadership			
	"Strategic and Successful Negotiation"	36	2.78%	100%
	Master Class for Efficient Leaders	30	20.00%	100%
UTHM HIGH PERFORMANCE LEADERSHIP COACHING (Hi-Per)	Impactful Self Leadership	20	10.00%	100%
	Leadership Forum		No applicable	No applicable
	Project Presentation		No applicable	No applicable
	"Upskilling Your Pesona"	31	53.33%	100%

5. Discussion

This study provides a systematic evaluation of UTHM's leadership development programmes within a TVET higher education context, using Kirkpatrick's Levels 1–3 to assess satisfaction, learning, and behavioural application. The findings demonstrate a generally positive trajectory in leadership competency development, though with notable variations that offer important insights for strengthening future TVET leadership initiatives.

At Level 1 (Reaction), the consistently high satisfaction scores across ULead, Hi-Lead, and Hi-Per programmes indicate that participants perceived the modules as relevant, engaging, and aligned with their leadership responsibilities. Programmes that addressed competencies such as ethical leadership, stakeholder management, and strategic thinking received particularly strong ratings, suggesting that these domains resonate with the leadership challenges faced in contemporary TVET institutions. This aligns with UNESCO (2022) and OECD (2020), which emphasise the increasing importance of values-based leadership and stakeholder responsiveness in navigating institutional transformation and industry expectations. However, the lower satisfaction in one ULead module, linked to trainer performance rather than curriculum relevance, reinforces evidence in the training literature that trainer quality is a dominant predictor of reaction outcomes (Alvarez et al., 2004; Tracey et al., 1995). This finding highlights the importance of trainer selection and preparation as core determinants of learner engagement within TVET leadership programmes.

At Level 2 (Learning), the significant pre–post improvements observed in all but one module reflect strong cognitive learning outcomes, particularly in technical or applied topics such as finance, negotiation, and stakeholder engagement. These results are consistent with prior studies showing that leadership programmes tend to generate measurable improvements in knowledge and confidence when learning objectives are clearly structured and aligned to participant needs (Colella et al., 2019; Masood & Usmani, 2015). The variation in learning gains across modules suggests that conceptual content—such as strategic planning—may require more experiential or problem-based methods to translate into knowledge acquisition, echoing similar patterns noted in He and Leeman (2021) and Franzen (2020). For TVET leadership development, this has direct implications: cognitive competencies in areas such as business acumen and creative thinking require ongoing reinforcement, coaching, and opportunities to apply learning in authentic institutional contexts.

Although the Level 3 evaluation received a lower response rate from supervisors—a common challenge in training transfer studies due to workload and availability constraints (Colella et al., 2019; Smidt et al., 2009)—the results are nonetheless meaningful. Importantly, all supervisors who responded reported observable behavioural improvements among participants, indicating that behavioural change did occur when follow-up assessment was possible. This aligns with Baldwin and Ford's (1988) assertion that transfer is most visible when learners receive adequate opportunity and support to apply their training. Thus, the low volume of supervisor feedback should not be interpreted as low behavioural change; rather, it reflects structural barriers to evaluation within academic environments. The consistent reports of improved communication, collaboration, leadership initiative, and application of acquired skills demonstrate that the programmes have begun to influence workplace behaviour among those assessed.

Overall, the integrated findings across Levels 1–3 point to several strategic implications for TVET leadership development. First, aligning programmes with the UTHM Competency-Based Talent Development Framework appears to support stronger learning outcomes, reinforcing the value of competency-based design. Second, the variation in behavioural transfer indicates a need to strengthen UTHM's post-training ecosystem—particularly supervisor involvement, follow-up mechanisms, and workplace opportunities to apply leadership competencies. Third, the strong reaction and learning scores provide a foundation upon which more advanced behavioural and results-based evaluation (Levels 3 and 4) can be developed, supporting evidence-based leadership pipelines essential for TVET transformation under GTU2030.

For TVET leaders, these findings highlight an important lesson: behavioural change can occur even with minimal structural reinforcement, but its visibility and sustainability depend on timely supervisor engagement. The fact that all assessed participants demonstrated behavioural improvements suggests that the programme design is effective in promoting leadership application. However, the limited number of supervisor evaluations indicates that TVET institutions must strengthen post-training mechanisms, such as formalised check-ins, structured coaching, and integration of leadership behaviours into performance dialogue. Embedding such systems will not only improve Level 3 evaluation quality but also enhance long-term leadership capability within the TVET ecosystem.

6. Conclusion

Taken together, the Level 1–3 findings demonstrate that UTHM's leadership programmes are achieving their intended developmental outcomes. High satisfaction (Level 1) and significant learning gains (Level 2) indicate strong programme relevance and instructional quality, while the positive behavioural reports from all responding supervisors (Level 3) confirm that trainees are beginning to translate knowledge into workplace practices. Although behavioural evaluation was limited by low response rates, the consistency of positive behavioural

observations suggests that meaningful leadership transfer is occurring. Strengthening organisational follow-up structures will allow TVET institutions to amplify these outcomes and support the development of a sustainable, competency-driven leadership pipeline.

Acknowledgement

Communication of this research is made possible through collaborative efforts and assistance from Registrar of Universiti Tun Hussein Onn Malaysia and UTHM Center of Leadership and Competency.

Conflict of Interest

Authors declare that there is no conflict of interests regarding the publication of the paper.

Author Contribution

The authors confirm contribution to the paper as follows: ***study conception and design:*** *Rabiah Ahmad and Nor Hazana Abdullah, Sam Faheeda Abdul Samad;* ***data collection:*** *Nor Hazana, Zhi Hong;* ***analysis and interpretation of results:*** *Nor Hazana and Zhi Hong, Rabiah Ahmad;* ***draft manuscript preparation:*** *Nor Hazana and Zhi Hong. All authors reviewed the results and approved the final version of the manuscript.*

References

Ab Rahman, A. A., Aziz, M. I., Uthamaputhran, S. A., Mohamad Anuar, N. I., Md. Isa@Yusuff, Y. Z., & Nik Muhammad, N. M. (2022). Changing entrepreneurial leadership knowledge competency in higher education: A way to move forward. *International Journal of Criminology and Sociology, 9*, 2566–2571. https://doi.org/10.6000/1929-4409.2020.09.315

Ahmad, A. R., & Arifin, S. R. (2021). Evaluating the effectiveness of leadership development programs in Malaysian public universities. *Asian Journal of University Education, 17*(1), 1–15. https://doi.org/10.24191/ajue.v17i1.8721

Alsalamah, A., & Callinan, C. (2021). Adaptation of Kirkpatrick's Four-Level Model of Training Criteria to Evaluate Training Programs for Head Teachers. *Education Sciences, 11*, 116. https://doi.org/10.3390/EDUCSCI11030116

Alvarez, K., Salas, E., & Garofano, C. M. (2004). An integrated model of training evaluation and effectiveness. *Human Resource Development Review, 3*(4), 385–416. https://doi.org/10.1177/1534484304270820

Baldwin, T. T., & Ford, J. K. (1988). Transfer of training: A review and directions for future research. *Personnel Psychology, 41*(1), 63–105. https://doi.org/10.1111/j.1744-6570.1988.tb00632.x

Bates, R. (2004). A critical analysis of evaluation practice: The Kirkpatrick model and the principle of beneficence. *Evaluation and Program Planning, 27*(3), 341–347.

Bolden, R. (2016). Leadership, management and organisational development. *European Journal of Training and Development, 40*(7), 610–628. https://doi.org/10.1108/EJTD-11-2015-0081

Colella, M. R., O'Connell, C. M., & Moore, A. T. (2019). Implementation and evaluation of an innovative leadership and teacher training program for non-physician emergency medicine practitioners in Uganda. *African Journal of Emergency Medicine, 9*(2), 25–30. https://doi.org/10.1016/j.afjem.2019.01.004

Day, D. V., Fleenor, J. W., Atwater, L. E., Sturm, R. E., & McKee, R. A. (2014). Advances in leader and leadership development: A review of 25 years of research and theory. *The Leadership Quarterly, 25*(1), 63–82. https://doi.org/10.1016/j.leaqua.2013.11.004

Day, D. V., & Dragoni, L. (2015). Leadership development: An outcome-oriented review. *Annual Review of Organizational Psychology and Organizational Behavior, 2*, 133–156. https://doi.org/10.1146/annurev-orgpsych-032414-111356

Dempster, N., Lovett, S., & Flückiger, B. (2011). Strategies for leadership development in teacher education. *School Leadership & Management, 31*(2), 105–120. https://doi.org/10.1080/13632434.2011.572986

Economic Planning Unit. (2021). *Twelfth Malaysia plan 2021–2025: A prosperous, inclusive, sustainable Malaysia*. Prime Minister's Department.

Falletta, S. V. (1998). Book Review: Evaluating Training Programs: The Four Levels. *American Journal of Evaluation*, *19*(2), 259-261. https://doi.org/10.1177/109821409801900215

Faisal-E-Alam, M., Begum, Z. A., & Islam, A. R. M. T. (2025). Unveiling training effectiveness through behavior and performance evaluations: A case from a developing country. *Evaluation and Program Planning, 110*, 102553. https://doi.org/10.1016/j.evalprogplan.2025.102553

Fernandez, S., Cho, Y. J., & Perry, J. L. (2010). Exploring the link between integrated leadership and public sector performance. *The Leadership Quarterly, 21*(2), 308–323. https://doi.org/10.1016/j.leaqua.2010.01.009

Franzen, M. (2020). Evaluating behavioural outcomes in academic leadership development programs. *Leadership in Health Services, 33*(2), 205–220.

Gigliotti, R. A. (2020). Leadership development in higher education: A literature review and implications for practice. *Journal of Leadership Education, 19*(1), 86–97. https://doi.org/10.12806/V19/I1/T2

Goleman, D., Boyatzis, R., & McKee, A. (2013). *Primal leadership: Unleashing the power of emotional intelligence*. Harvard Business Review Press.

Hall, K., Bardier, C., Greer, D., Clayton, A., & Poblete, R. (2024). Unleashing leadership potential in unprecedented times: Lessons from a virtual adaptive leadership programme. *Public Health in Practice, 8*, 100532. https://doi.org/10.1016/j.puhip.2024.100532

Hamzah, N., & Ismail, M. (2020). Succession planning and leadership development in Malaysian TVET institutions. *Journal of Technical Education and Training, 12*(1), 1–10.

He, Y., & Leeman, J. (2021). Empowering women leaders: Impacts of leadership development programs. *Journal of Higher Education Policy and Management, 43*(6), 624–640.

Heydari, M. R., Taghva, F., Amini, M., & Delavari, S. (2019). Using Kirkpatrick's model to measure the effect of a new teaching and learning methods workshop for health care staff. *BMC Research Notes, 12*, 388. https://doi.org/10.1186/s13104-019-4421-y

Holton, E. F. (1996). The flawed four-level evaluation model. *Human Resource Development Quarterly, 7*(1), 5–21.

Hopkins, M. M., O'Neil, D. A., & Stoller, J. K. (2017). Distinguishing competencies for leadership success in academia and healthcare. *Leadership in Health Services, 30*(2), 148–163.

Iqbal, M. Z., & Khan, R. A. (2022). Effect of trainer competency on training effectiveness: The mediating role of trainee motivation. *International Journal of Training and Development, 26*(1), 87–102. https://doi.org/10.1111/ijtd.12240

Jain, G., Sharma, N., & Shrivastava, A. (2021). Enhancing training effectiveness through blockchain-enabled training effectiveness measurement (BETEM). *Journal of Organizational Change Management, 34*(2), 439–461. https://doi.org/10.1108/JOCM-10-2020-0303

Kaufman, R., Keller, J., & Watkins, R. (1995). What Works and What Doesn't: Evaluation beyond Kirkpatrick. Performance and Instruction, 35, 8-12.https://doi.org/10.1002/pfi.4170350204

Ketut, R., Yudana, M. D., Natajaya, N., & Divayana, D. G. H. (2022). Leadership evaluation with the Kirkpatrick model. In I. Idris et al. (Eds.), *Exploring new horizons and challenges for social studies in a new normal* (pp. 164–176). Routledge. https://doi.org/10.1201/9781003290865-14

Kirkpatrick, D. L. (1959). Techniques for Evaluation Training Programs. Journal of the American Society of Training Directors, 13, 21-26.

Kirkpatrick, D. L. (1976). Evaluation of training. In R. L. Craig (Ed.), Training and development handbook: A guide to human resource development (2nd ed., pp. 301–319). McGraw-Hill.

Kirkpatrick, D. L. (1994). Evaluating training programs: The four levels (2nd ed.). American Society for Training and Development, Inc.

Kirkpatrick, D. L. (1998). *Evaluating training programs: The four levels*. Berrett-Koehler.

Kirkpatrick, D. L., & Kirkpatrick, J. D. (2006). Evaluating Training Programs: The Four Levels (3rd ed.). Berrett-Koehler.

Kirkpatrick Partners. (2016). *Kirkpatrick's four levels of training evaluation*. https://www.kirkpatrickpartners.com/Our-Philosophy/The-Kirkpatrick-Model

Lantu, D. C., Triana, R., Sitinjak, T., & Sari, D. (2020). Barriers to leadership skill transfer in organizations. *Journal of Management Development, 39*(7), 869–885.

Lacerenza, C. N., Reyes, D. L., Marlow, S. L., Joseph, D. L., & Salas, E. (2017). Leadership Training Design, Delivery, and Implementation: A Meta-Analysis. *Journal of Applied Psychology, 102,* 1686-1718. https://doi.org/10.1037/apl0000241

Lizzio, A., Dempster, N., & Neumann, R. (2011). Pathways to student leadership. *International Journal of Leadership in Education, 14*(1), 1–16.

Masood, A., & Usmani, T. A. (2015). Measuring the impact of teachers' training on learning outcomes. *Journal of Education and Practice, 6*(28), 76–81.

Maulina, H., et al. (2023). Leadership development in primary healthcare settings. *Journal of Health Management, 25*(3), 245–260.

Newstrom, J. W. (1995). Evaluating training programs: The four levels. *Journal of Continuing Education in the Health Professions, 15*(4), 249–250.

Noe, R. A., Clarke, A. D. M., & Klein, H. J. (2014). Learning in the twenty-first-century workplace. *Annual Review of Organizational Psychology and Organizational Behavior, 1*, 245–275.

O'Keeffe, M., et al. (2017). Student evaluations of teaching: Concerns about validity and reliability. *Higher Education Research & Development, 36*(5), 1017–1031.

OECD. (2020). *Strengthening the governance of skills systems: Lessons from six OECD countries*. OECD Publishing. https://doi.org/10.1787/3a4bb6ea-en

Pandeani, N. K. N., Suprastayasa, I. G. N. A., & Kartini, L. P. (2020). Employees' perception towards online training program: A case from a hotel. *Jurnal Bisnis Hospitaliti, 11*(1), 27–33. https://doi.org/10.22334/jbh.v11i1.186

Paull, M., Whitsed, C., & Girardi, A. (2016). Applying the Kirkpatrick model to evaluate an Interaction for Learning Framework intervention. *Issues in Educational Research, 26*(3), 490–507.

Phillips, P. P., & Phillips, J. J. (2016). *Handbook of training evaluation and measurement methods* (4th ed.). Routledge.

Purwasetiawatik, T. F., Ibrahim, M. M., & Rasyid, M. N. A. (2024). Leadership training program at Faculty of Psychology, Bosowa University. *Journal La Edusci, 5*(5), 293–308. http://dx.doi.org/10.37899/journallaedusci.v5i5.1760

Reio, T. G., Rocco, T. S., Smith, D. H., & Chang, E. (2017). A critique of Kirkpatrick's evaluation model. *New Horizons in Adult Education and Human Resource Development, 29*(2), 35–53. https://doi.org/10.1002/nha3.20178

Ruiz, J., & Snoeck, M. (2018). Adapting Kirkpatrick's evaluation model to technology-enhanced learning. In *MODELS '18: Companion Proceedings of the 21st ACM/IEEE International Conference on Model Driven Engineering Languages and Systems* (pp. 135–142). ACM. https://doi.org/10.1145/3270112.3270114

Rowe, W. G. (2001). Creating wealth in organizations: The role of strategic leadership. *Academy of Management Perspectives, 15*, 81-94. https://doi.org/10.5465/AME.2001.4251395

Saad, M., & Mat, T. Z. (2013). Evaluating training effectiveness using Kirkpatrick model: Evidence from Malaysian SMEs. *International Journal of Academic Research in Business and Social Sciences, 3*(8), 251–260.

Sadeghi, B. (2012). Professional development needs of EFL teachers. *Procedia – Social and Behavioral Sciences, 46*, 337–341.

Salas, E., Tannenbaum, S. I., Kraiger, K., & Smith-Jentsch, K. A. (2012). The science of training and development in organizations. *Psychological Science in the Public Interest, 13*(2), 74–101.

Sanagavarapu, P., et al. (2024). Student reaction surveys in online learning. *Journal of University Teaching & Learning Practice, 21*(1), 112–125.

Smidt, A., Balandin, S., Sigafoos, J., & Reed, V. (2009). The Kirkpatrick model: A useful tool for evaluating training outcomes. *Journal of Intellectual & Developmental Disability, 34*(3), 266–274.

Tannenbaum, S. I., & Yukl, G. (1992). Training effectiveness: A review and a model. *Annual Review of Psychology, 43*, 399–441.

Throgmorton, C.A., Mitchell, T., Morley, T., & Snyder, M. (2016). Evaluating a physician leadership development program - a mixed methods approach. *Journal of health organization and management, 30 3*, 390-407. https://doi.org/10.1108/JHOM-11-2014-0187

Tracey, J. B., Tannenbaum, S. I., & Kavanagh, M. J. (1995). Applying trained skills on the job: The importance of the work environment. *Journal of Applied Psychology, 80*(2), 239–252.

UNESCO. (2022). *Transforming technical and vocational education and training*. UNESCO.

UNESCO-UNEVOC. (2022). *Transforming TVET for successful transitions*. UNESCO.

Yusof, H., & Yacob, A. (2020). Evaluating TVET leadership programmes in Malaysia. *Journal of Technical Education and Training, 12*(2), 51–60.

Zhao, W., Liu, Z., Wang, T., Yin, X., Sun, Y., Zhang, X., & Yang, H. (2023). Assessment of a training project of English as a media of instruction (EMI) using Kirkpatrick model. *BMC Medical Education, 23*. https://doi.org/10.1186/s12909-023-04204-5