Longitudinal Training Programme Evaluation of FITI

2025

Submitted by:

FINANCIAL INSTITUTIONS TRAINING INSTITUTE

Contents

1. Introduction 1
2. Literature Review 3
3. Methodology 3
   3.1. Data Source 4
   3.2. Variables and Measurement 4
      A. Programme Design and Content Quality 4
      B. Trainer Effectiveness 5
      C. Administrative Support and Facilities 5
      D. Overall Programme Rating and Duration 5
      E. Follow-up Interest and Recommendation 6
   3.3. Study Design 6
   3.4. Data Preparation 6
   3.5. Reliability Analysis 7
   3.6. Statistical Analysis 7
   3.7. Hypotheses 8
4. Results 10
   4.1. Descriptive Trends 10
   4.2. ANOVA Results 13
   4.3. Kruskal–Wallis Test Results 14
   4.4. Chi-Square Tests 15
   4.5. Discussion 16
5. Conclusion 25

List of Figures

Figure 1: Bar plot for follow-up interest 15
Figure 2: Bar plot Programme Recommendation 16
Figure 3: Program Content Quality Comparison (2023–2025) 17
Figure 4: Teaching Quality Trend (2023 - 2025) 19
Figure 5: Training Environment and Administrative Facilities 21
Figure 6: Participants' Perception of programme Duration 24

List of Tables

Table 1: Means across three years 10
Table 2: Standard Deviation 10
Table 3: Programme Design and Content Quality: Item-Level Results 11
Table 4: Trainer Effectiveness: Item-Level Results 12
Table 5 : Training Environment and Support: Item-Level Results 12
Table 6 : ANOVA Results 13
Table 7: Tukey HSD Post-hoc Comparisons (2023–2025) 13
Table 8: Kruskal–Wallis Test Results for Ordinal Outcomes across Years (2023–2025) 14
Table 9: Pairwise Wilcoxon Post-hoc Comparisons with Bonferroni Adjustment 14
Table 10: Chi-Square Tests for Binary Outcomes (2023–2025) 15
Table 11: Training Environment and Support: Item-Level Results 20
Table 12 : *Kruskal–Wallis Test Results for Ordinal Outcomes across Years (2023–2025)* 22

1. Introduction

Evaluating training programmes is essential for enhancing the effectiveness and strategic impact of organizational learning initiatives, as it provides evidence on how well training meets its objectives and contributes to organizational performance (Kirkpatrick & Kirkpatrick, 2006; Noe, 2020). Systematic evaluation enables organizations to identify strengths and weaknesses in training delivery, justify investments in human capital, and make informed decisions for future improvements (Salas et al., 2012). Without rigorous evaluation, training programmes risk perpetuating ineffective content, failing to meet learner needs, or misaligning with evolving industry demands.

The Financial Institutions Training Institute Limited (FITI) in Bhutan exemplifies a formal institutional response to the growing need for qualified professionals in the financial sector and beyond. FITI was established in 2011 by the Royal Monetary Authority of Bhutan in collaboration with the Financial Institutions Association of Bhutan as an autonomous apex institution with the mandate to train and develop human resource capacity, provide research and consultancy services, and act as a center of excellence for knowledge in the financial sector. Over time, the institute has expanded its scope to deliver relevant training programmes for diverse stakeholders, including corporate agencies and private sector professionals, thereby contributing to national human resource development in a knowledge-based society.

Prior research indicates that multiple factors influence the overall success of training programmes. Program design including clear learning objectives, relevant content, and alignment with professional competencies has been shown to significantly affect learning outcomes and transfer of training (Salas et al., 2012). Trainer effectiveness, reflecting a facilitator's expertise, instructional clarity, and ability to engage learners, is also critical in shaping participants' perceptions of training relevance and satisfaction (Burke & Hutchins, 2007). Additionally, administrative support, such as logistical arrangements, institutional backing, and availability of adequate facilities, plays a supportive role in facilitating a conducive learning environment that enhances participant engagement and training effectiveness (Arthur et al., 2003).

Although numerous studies assess the immediate outcomes of individual training programmes, there is a scarcity of research examining how trainee evaluations and program effectiveness evolve over time. Longitudinal analyses are necessary to identify trends and patterns in perceptions of program quality and to assess whether improvements or declines occur across successive years of implementation. Understanding these trends is particularly important for institutions like FITI, which continuously updates its training offerings and responds to shifting skill requirements in financial and related sectors. Therefore, the purpose of this study is to examine trends and patterns in participants' evaluations of FITI's training programmes across multiple years, focusing on dimensions such as program design, trainer effectiveness, and administrative support. This longitudinal perspective aims to provide actionable insights that support continuous improvement in training delivery and strategic human resource development planning.

2. Literature Review

Evaluating training programmes is critical for ensuring that learning initiatives achieve their intended outcomes and contributes to organizational performance (Noe, 2020; Salas et al., 2012). Well-designed programmes with clear objectives, relevant content, and alignment to professional competencies have been shown to improve knowledge acquisition, skills application, and overall participant satisfaction (Arthur, Bennett, Edens, & Bell, 2003). Without systematic evaluation, organizations risk delivering ineffective training that fails to meet learner needs or align with evolving industry standards.

Trainer effectiveness is another key determinant of training success. Research indicates that a facilitator's subject knowledge, instructional clarity, interaction with participants, and responsiveness to questions significantly influence both learning outcomes and participant engagement (Burke & Hutchins, 2007). High-quality facilitation can compensate for minor gaps in content or structure, whereas weak instruction can undermine otherwise well-designed programmes.

Administrative and logistical support also plays a vital role in shaping participants' perceptions. Adequate facilities, proper scheduling, and supportive institutional arrangements create a conducive learning environment, enabling participants to focus on the content and interact effectively with trainers and peers (Arthur et al., 2003). Conversely, poor organization or inadequate resources can distract learners and reduce training effectiveness, even if programme content and instruction are strong. Despite extensive research on training effectiveness, most studies focus on single programmes or short-term outcomes, leaving a gap in understanding how participant perceptions evolve over time (Salas et al., 2012). Longitudinal evaluation is essential for monitoring trends, identifying gradual improvements or declines, and supporting evidence-based interventions. For institutions like the Financial Institutions Training Institute Limited (FITI) in Bhutan, which continuously updates training offerings to meet sectorial skill needs, longitudinal feedback provides actionable insights for maintaining high-quality programmes, enhancing trainer performance, and improving participant experiences across successive years.

3. Methodology

3.1. Data Source

The study utilizes secondary data obtained from participant feedback surveys administered by the Financial Institution Training Institute (FITI) at the conclusion of each training programme conducted between 2023 and 2025. Feedback was collected immediately after each event to capture participants' immediate perceptions of programme effectiveness. Participation in the survey was voluntary, and responses were anonymized prior to analysis, in accordance with ethical standards for educational and training research (Creswell & Creswell, 2018).

The dataset contains records of training programmes delivered to professionals from various financial and public-sector institutions. Each record represents an individual participant's evaluation of a specific programme. Data cleaning involved the identification and correction of inconsistent coding; empty or incomplete responses were retained at this stage to allow analysis of the true effect of the programmes, though missing values were later handled appropriately during index construction.

3.2. Variables and Measurement

The feedback instrument primarily used ordinal **Likert-type scales**, where responses are ordered but not necessarily equally spaced. Likert-type scales are widely used in programme evaluation to measure attitudes, perceptions, and satisfaction (Boone & Boone, 2012). In this study, the variables are grouped as follows:

A. Programme Design and Content Quality

Programme design and content quality were assessed using four items reflecting participants' perceptions of the structure and substance of the training programme:

- Clarity of session objectives and aims
- Coverage of relevant topics
- Quality of learning activities and exercises
- Level and depth of programme content

These items were aggregated to form the **Programme Design index** for later analyses. Responses were recorded on an ordinal scale with four categories: *poor*, *fair*, *good*, and *very good*.

B. Trainer Effectiveness

Trainer effectiveness was evaluated using seven items covering the instructor's:

- Knowledge of the subject
- Teaching clarity
- Interaction and responsiveness
- Facilitation skills
- Pace of delivery
- Time management

These key performance indicators will later be combined into a **Trainer Effectiveness index**. Responses used the same four-point Likert-type scale ranging from *poor* to *very good*.

C. Administrative Support and Facilities

Administrative aspects were measured using four items:

- Meeting space
- Classroom facilities
- Administrative support
- Catering services

An **Administration index** will be constructed from these items. Responses followed the same four-point Likert-type scale.

D. Overall Programme Rating and Duration

- **Overall programme rating**: Participants rated the programme overall on the four-point scale (*poor*, *fair*, *good*, *very good*).

- **Perceived programme duration**: Participants indicated whether the programme was *short*, *right length*, or *long*.

E. Follow-up Interest and Recommendation

Two additional outcome variables captured participants' behavioral intentions:

- **Follow-up interest**: Whether participants would be interested in attending a follow-up or advanced training on the same subject (*yes*, *no*).
- **Programme recommendation**: Whether participants would recommend the programme to others (*yes*, *no*).

These variables provide insight into participants' overall engagement and the potential for programme continuation (Singer & Willett, 2003; Field, 2018).

3.3. Study Design

This research adopts a **longitudinal observational design** to examine trends and variations in participant feedback across a three-year period (2023–2025). Longitudinal designs are commonly used to assess changes in perceptions, attitudes, and programme effectiveness over time (Singer & Willett, 2003). The same feedback instrument was applied consistently across all programmes and years, ensuring that observed differences reflect genuine variation in participant perceptions rather than measurement artifacts.

3.4. Data Preparation

Prior to analysis, data were screened for inconsistent coding and missing values. Responses to Likert-type items were treated as **ordinal variables**, with mean scores calculated for descriptive purposes and index construction (Norman, 2010). Four main index variables were created for later analysis: **Programme Design**, **Trainer Effectiveness**, **Administration**, and **Overall Programme Rating**. Additional outcome variables included **Perceived Programme Duration**, **Follow-up Interest**, and **Programme Recommendation**. Categorical variables were factorized appropriately for inferential testing.

3.5. Reliability Analysis

The internal consistency of multi-item constructs Programme Design, Trainer Effectiveness, and Administration was assessed using **Cronbach's alpha** (Cronbach, 1951; Tavakol & Dennick, 2011). An alpha coefficient of 0.70 or higher was considered acceptable, confirming that items included in each index reliably measured the intended construct.

3.6. Statistical Analysis

The analysis was conducted using **R statistical software**, applying the following techniques:

1. **Descriptive statistics** – means, standard deviations, and frequency distributions summarized participant feedback for each year (2023–2025).
2. **Visualization** – trends in Programme Design, Trainer Effectiveness, Administration, Overall Rating, and Duration were plotted to detect variations across years.
3. **Inferential analysis** –
    - **Ordinal outcomes** (Programme Design, Trainer Effectiveness, Administration, Overall Programme Rating) were analyzed using **non-parametric tests** (e.g., Kruskal–Wallis test), suitable for Likert-type data (Norman, 2010).
    - **Categorical outcomes** (Perceived Duration, Follow-up Interest, Programme Recommendation) were analyzed using **Chi-square tests** or **Fisher's Exact Test** where expected counts were low.
4. **Significance level** – Statistical significance was set at 5% ($p < .05$) (Field, 2018).

3.7. Hypotheses

The study aims to examine **variation in participant feedback across the years 2023–2025**. The following hypotheses are tested:

1. **Programme Content Quality**
    - *$H0_1$*: There is no significant difference in programme content quality across the three years.
    - *$H1_1$*: There is a significant difference in programme content quality across the three years.
    - *Rationale*: Evaluates whether participant ratings of clarity of objectives, topics covered, quality of activities, and depth of content vary over time.
2. **Teaching Quality**
    - *$H0_2$*: There is no significant difference in teaching quality across the three years.
    - *$H1_2$*: There is a significant difference in teaching quality across the three years.
    - *Rationale*: Assesses whether participant perceptions of trainer effectiveness (teaching clarity, interaction, responsiveness, facilitation, pacing, and time management) differ across years.
3. **Administrative Support and Facilities**
    - *$H0_3$*: There is no significant difference in administrative support and facilities across the three years.
    - *$H1_3$*: There is a significant difference in administrative support and facilities across the three years.
    - *Rationale*: Examines variation in perceptions of meeting space, classroom facilities, administrative support, and catering services.
4. **Overall Programme Rating**
    - *$H0_4$*: There is no significant difference in overall programme ratings across the three years.
    - *$H1_4$*: There is a significant difference in overall programme ratings across the three years.

5. **Perception of Programme Duration**
    - $H0_5$: There is no significant difference in participants' perception of programme duration across the three years.
    - $H1_5$: There is a significant difference in participants' perception of programme duration across the three years.
6. **Interest in Follow-up Training**
    - $H0_6$: Interest in attending follow-up or advanced training is independent of year.
    - $H1_6$: Interest in attending follow-up or advanced training is associated with year.
7. **Programme Recommendation**
    - $H0_7$: Willingness to recommend the programme to others is independent of year.
    - $H1_7$: Willingness to recommend the programme to others is associated with year.

These hypotheses provide a structured framework for the statistical analysis to the research objectives and enabling assessment of year-wise variation in all key quality dimensions and participant outcomes.

4. Results

4.1. Descriptive Trends

Table 1*: Means across three years*

Year	Programme Design	Trainer Effectiveness	Administration Facilities	Overall Rating
2023	3.56	3.66	3.53	3.57
2024	3.66	3.76	3.62	3.75
2025	3.54	3.63	3.56	3.64

Table 2*: Standard Deviation*

Year	Programme Design	Trainer Effectiveness	Administration Facilities	Overall Rating
2023	0.45	0.4	0.51	0.61
2024	0.42	0.31	0.47	0.46
2025	0.46	0.36	0.46	0.53

Table 1 and 2 summarizes the year-wise mean scores and standard deviations for the major feedback dimensions from 2023 to 2025. Across all dimensions, mean scores remained above 3.5, indicating generally positive participant evaluations throughout the study period.

For programme design, the mean score increased from 3.56 in 2023 to 3.66 in 2024, before slightly declining to 3.54 in 2025. Trainer effectiveness followed a similar pattern, with the highest mean observed in 2024 (M = 3.76), compared to 2023 (M = 3.66) and 2025 (M = 3.63).

Mean scores for administrative facilities also peaked in 2024 (M = 3.62), with slightly lower but comparable values in 2023 (M = 3.53) and 2025 (M = 3.56). Overall programme ratings were highest in 2024 (M = 3.75), followed by 2025 (M = 3.64) and 2023 (M = 3.57).

Standard deviations across all years ranged from 0.31 to 0.61, indicating moderate variability in participant responses, with comparatively lower dispersion observed in 2024. Overall, the descriptive results suggest that 2024 recorded the most favorable and consistent participant evaluations across all key dimensions, providing a foundation for subsequent reliability and inferential analyses.

***Table 3**: Programme Design and Content Quality: Item-Level Results*

Variable	Uncorrected Item–Total Correlation	Corrected Item–Total Correlation	Mean	SD
Clarity of session objectives and aims	0.754	0.582	3.711	0.502
Topics covered in the session	0.751	0.552	3.623	0.556
Quality of learning activities and exercises	0.791	0.600	3.566	0.591
Level and depth of programme content	0.799	0.590	3.452	0.637

Item-level analysis was conducted to assess the contribution of individual items to the programme design and content quality scale. As shown in Table 3, all items demonstrated adequate to strong corrected item–total correlations, indicating that each item was meaningfully associated with the overall construct (Cronbach, 1951).

Corrected item–total correlations ranged from 0.55 to 0.60, with Quality of learning activities and exercises showing the strongest association with the overall scale (r = 0.60). Level and depth of programme content also exhibited a strong corrected item–total correlation (r = 0.59), supporting its relevance to the construct.

Mean scores across items ranged from 3.45 to 3.71, reflecting generally positive participant perceptions. Clarity of session objectives and aims received the highest mean rating (M = 3.71, SD = 0.50), while Level and depth of programme content recorded the lowest mean score (M = 3.45, SD = 0.64). Overall, these findings confirm the internal consistency and coherence of the programme design and content quality scale.

***Table 4**: Trainer Effectiveness: Item-Level Results*

Variable	Uncorrected Item–Total Correlation	Corrected Item–Total Correlation	Mean	SD
Knowledge and experience in subject matter	0.617	0.488	3.772	0.442
Clarity and effectiveness of teaching	0.732	0.621	3.698	0.486
Interaction with participants	0.698	0.577	3.728	0.485
Responsiveness to questions comments issues	0.720	0.601	3.725	0.503
Facilitation of activities and exercises	0.733	0.600	3.605	0.567
Pace of programme delivery	0.755	0.621	3.593	0.588
Organization and time management of programme	0.649	0.498	3.694	0.543

Item-level analysis indicated that all items contributed meaningfully to the Trainer Effectiveness scale. Corrected item–total correlations ranged from 0.49 to 0.62, with the strongest associations observed for Clarity and effectiveness of teaching and Pace of programme delivery (r = 0.62). All items met acceptable thresholds, supporting the internal consistency of the scale.

Mean scores ranged from 3.59 to 3.77, reflecting generally positive participant perceptions. Knowledge and experience in subject matter received the highest rating (M = 3.77, SD = 0.44), while Pace of programme delivery showed the lowest mean score (M = 3.59, SD = 0.59). Overall, the findings confirm the coherence and reliability of the Trainer Effectiveness scale.

***Table 5** : Training Environment and Support: Item-Level Results*

Variable	Uncorrected Item–Total Correlation	Corrected Item–Total Correlation	Mean	SD
Meeting space	0.769	0.617	3.707	0.521
Facilities in classroom	0.810	0.663	3.611	0.562
Administrative support	0.825	0.688	3.657	0.558
Catering services, Refreshment and Lunch	0.771	0.487	3.300	0.801

Item-level analysis showed that all items contributed adequately to the Training Environment and Support scale. Corrected item–total correlations ranged from 0.49 to 0.69, with administrative support (r = 0.69) and Facilities in the classroom (r = 0.66) demonstrating the strongest associations with the overall construct. Catering services, refreshments, and lunch displayed the lowest corrected item–total correlation (r = 0.49) but remained within acceptable limits.

Mean scores ranged from 3.30 to 3.71, indicating generally positive perceptions of the training environment. Meeting space received the highest mean rating (M = 3.71, SD = 0.52), while Catering services, refreshments, and lunch recorded the lowest mean score (M = 3.30, SD = 0.80). Overall, the results support the internal consistency and adequacy of the Training Environment and Support scale.

4.2. ANOVA Results

***Table 6** : ANOVA Results*

Variable	Source	df	Sum Sq	Mean Sq	F	p
Program Content Quality	Year	2	2.94	1.472	7.586	0.0005**
	Residuals	1025	198.82	0.194		
Trainer Effectiveness	Year	2	3.12	1.561	12.08	6.52e-06**
	Residuals	1025	132.41	0.129		
Administrative Facilities	Year	2	1.87	0.937	4.034	0.018*
	Residuals	1025	238.19	0.232		

Note: *p < 0.05, **p < 0.01.

***Table 7**: Tukey HSD Post-hoc Comparisons (2023–2025)*

Variable	Comparison	Mean Diff	95% CI Lower	95% CI Upper	p adj
Program Quality	2024–2023	0.104	0.029	0.180	0.0034**
	2025–2023	-0.017	-0.098	0.064	0.875
	2025–2024	-0.121	-0.204	-0.038	0.0018**
Trainer Effectiveness	2024–2023	0.099	0.038	0.161	0.00048**
	2025–2023	-0.032	-0.099	0.034	0.485
	2025–2024	-0.132	-0.199	-0.064	0.000017**
Administrative Facilities	2024–2023	0.100	0.017	0.182	0.013*
	2025–2023	0.042	-0.047	0.131	0.511
	2025–2024	-0.058	-0.148	0.033	0.296

The results of the one-way ANOVA and Tukey HSD post-hoc tests indicate that participant perceptions of programme content quality, teaching effectiveness, and administrative support varied significantly across the years 2023 to 2025. Specifically, the p-values for all three dimensions were below the conventional 0.05 threshold (Programme Content Quality: p = 0.0005; Teaching Effectiveness: p < 0.001; Administrative Facilities: p = 0.018), leading to the rejection of the null hypotheses $H0_1$, $H0_2$, and $H0_3$. This confirms that participants' evaluations were not consistent over time, with 2024 generally showing superior ratings. These findings

highlight the importance of longitudinal monitoring of training programmes to detect improvements or declines in instructional quality, administrative support, and overall programme effectiveness, supporting data-driven decisions for continuous programme enhancement.

4.3. Kruskal–Wallis Test Results

***Table 8**: Kruskal–Wallis Test Results for Ordinal Outcomes across Years (2023–2025)*

Variable	Test Statistic (χ^2)	df	p-value	Decision
Overall Programme Rating	18.135	2	0.000115***	Reject H_{04}
Programme Duration	38.224	2	< 0.001***	Reject H_{05}

Note: ***p < 0.001.

***Table 9**: Pairwise Wilcoxon Post-hoc Comparisons with Bonferroni Adjustment*

Overall Programme Rating		
Comparison	**Adjusted p-value**	**Interpretation**
2024 – 2023	0.00007***	Significant difference
2025 – 2023	0.448	Not significant
2025 – 2024	0.035*	Significant difference
Programme Duration		
Comparison	**Adjusted p-value**	**Interpretation**
2024 – 2023	0.0038**	Significant difference
2025 – 2023	< 0.001***	Significant difference
2025 – 2024	0.0014**	Significant difference

Note: *p < 0.05, **p < 0.01, ***p < 0.001.

Based on the Kruskal–Wallis test results, the null hypotheses H_{04} (overall programme rating) and H_{05} (programme duration perception) were rejected, as the associated p-values were well below the 0.05 significance level ($p < 0.001$). This provides strong statistical evidence that participants' overall evaluations of the programme and their perceptions of programme duration differed significantly across the years 2023–2025.

Post-hoc pairwise Wilcoxon tests with Bonferroni adjustment further clarified the nature of these differences. For overall programme ratings, a significant improvement was observed in 2024 compared to 2023, while ratings in 2025 did not differ significantly from 2023 but was significantly lower than those in 2024. For programme duration, significant differences were

detected across all year-to-year comparisons, indicating a consistent shift in how participants perceived the appropriateness of programme length over time.

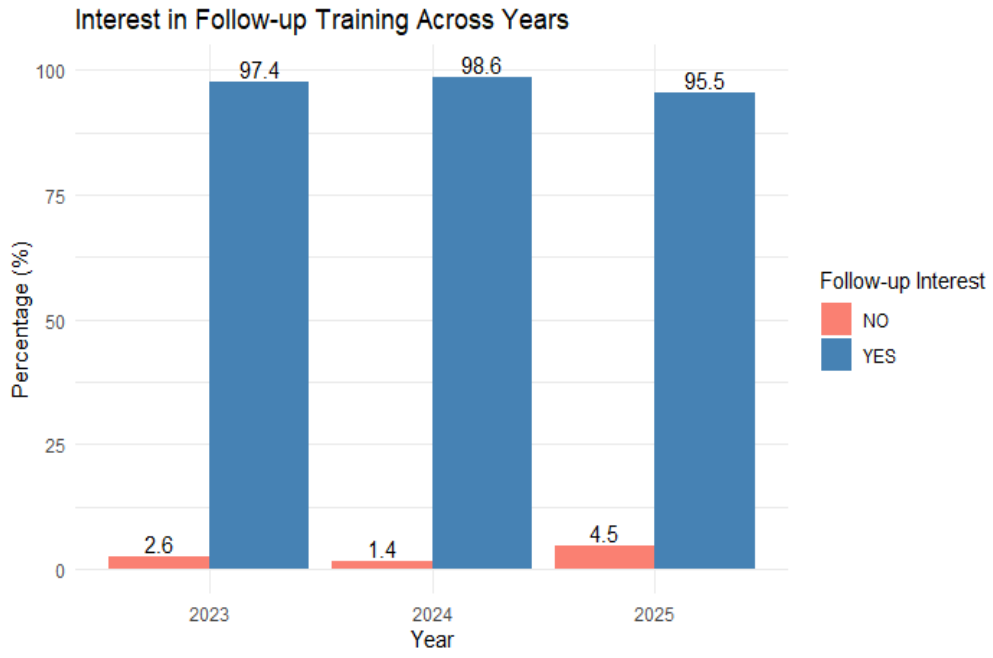
4.4. Chi-Square Tests

***Table 10**: Chi-Square Tests for Binary Outcomes (2023–2025)*

Variable	df	X^2	p-value	Decision
Interest in Follow-up Training	2	5.633	0.0598	Fail to reject H_{06}
Programme Recommendation	2	0.78952	0.6738	Fail to reject H_{07}

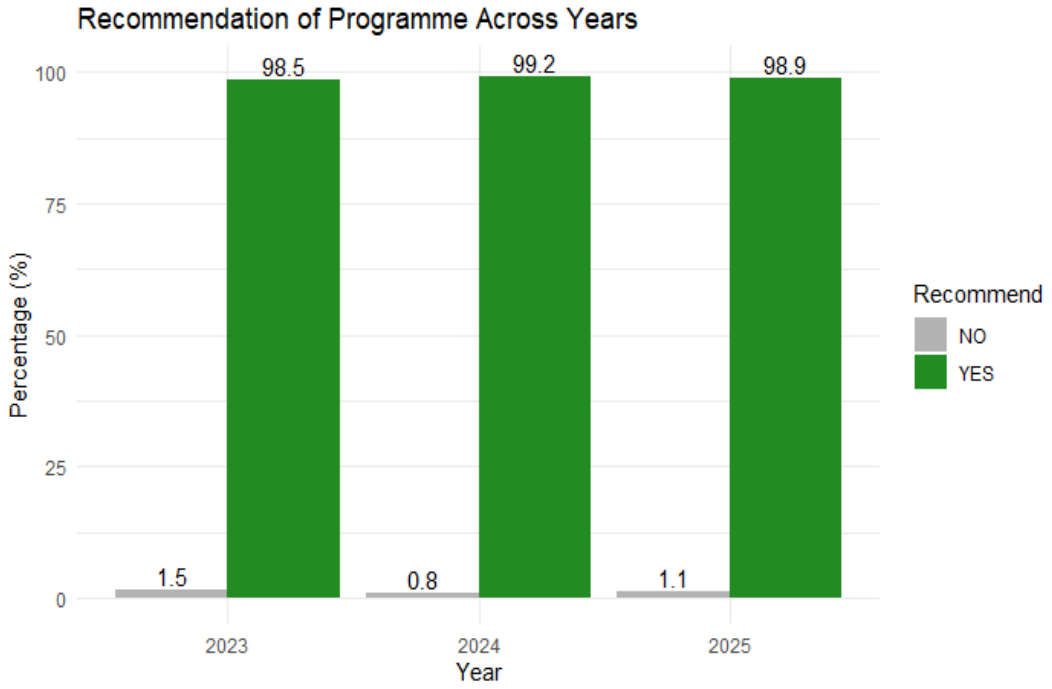
Follow-up Training Interest: The test was not statistically significant (p = 0.0598). Therefore, H_{06} is not rejected, indicating that participants' interest in attending follow-up or advanced training did not significantly change across the years 2023–2025. Moreover, referring figure 1 we can observe that the percentage for follow-up interest is comparatively above for those who have submitted that they are not interested.

***Figure 1**: Bar plot for follow-up interest*



Programme Recommendation: The test was not statistically significant (p = 0.6738). Therefore, H_{07} is not rejected, showing that participants' willingness to recommend the programme remained stable over the three years.

***Figure 2**: Bar plot Programme Recommendation*



Overall Interpretation: Despite minor fluctuations, participants' behavioral intentions regarding follow-up participation and programme recommendation were consistent, reflecting stability in engagement and endorsement even as programme content or delivery may have evolved.

4.5. Discussion

This study examined longitudinal trends in participant feedback on professional training programmes conducted by the Financial Institution Training Institute (FITI) from 2023 to 2025. By integrating descriptive statistics, reliability analysis, and inferential tests, the findings provide a comprehensive understanding of how programme quality, trainer effectiveness, administrative facilities, and overall perceptions evolved over time. We developed a nuanced understanding of how training quality and participant perceptions evolved over time. Our findings contribute to the limited body of longitudinal training evaluation research and provide practical insights for institutional quality assurance and continuous improvement.

Trends in Participant Perceptions across Years

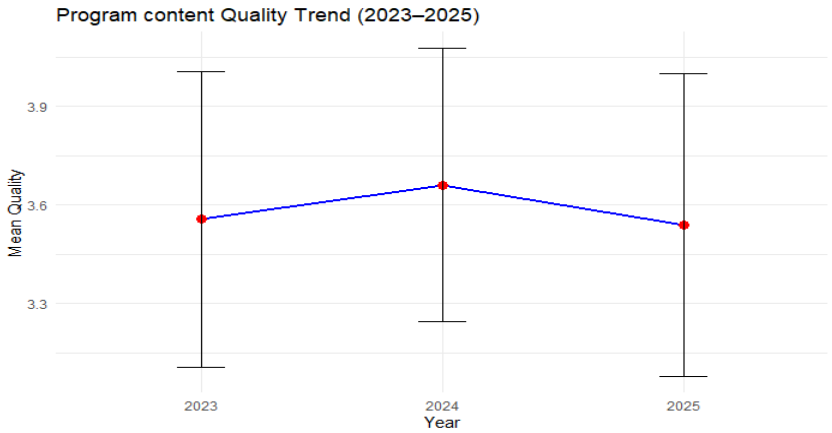
Descriptive results indicate that participant evaluations remained consistently positive throughout the three-year period, with mean scores across all dimensions exceeding 3.5. This suggests that FITI's training programmes met participant expectations at an overall satisfactory level across years. Notably, 2024 emerged as the most favorable year, recording the highest mean scores and the lowest variability across programme design, trainer effectiveness, administrative facilities, and overall programme ratings. The comparatively lower standard deviations in 2024 further indicate more consistent participant experiences during this period.

The slight decline observed in 2025 should be interpreted cautiously. Given that scores remained high and differences were relatively small in magnitude, this decline does not indicate deterioration in programme quality. Rather, it may reflect contextual factors such as increased participant expectations, broader programme diversification, or external constraints affecting delivery. From our perspective, this underscores the value of longitudinal evaluation in distinguishing meaningful changes from normal fluctuations in participant perceptions.

Programme Design and Content Quality

The findings related to programme design and content quality indicate that this dimension was both reliably measured and meaningfully differentiated across years. Item-level analysis demonstrated strong internal coherence, with all corrected item–total correlations exceeding accepted thresholds, confirming the robustness of the scale (Cronbach, 1951). This supports the use of a composite measure to evaluate participants' perceptions of programme content quality over time.

***Figure 3**: Program Content Quality Comparison (2023–2025)*



Participants consistently rated the clarity of session objectives most favorably, underscoring FITI's strength in clearly communicating training goals and expectations. Clear articulation of objectives is widely recognized as a critical component of effective instructional design, as it shapes learner engagement and guides content delivery. In contrast, the relatively lower mean score observed for the level and depth of programme content suggests potential scope for further enhancement. This finding may reflect the diverse professional backgrounds and experience levels of participants, indicating a need for greater differentiation or modularization of content depth to better accommodate varied learning needs.

As shown in Table 6, the ANOVA results revealed statistically significant differences in programme content quality across the three years ($F = 7.586$, $p = 0.0005$), leading to the rejection of the null hypothesis ($H0_1$) and acceptance of the alternative hypothesis ($H1_1$). Post-hoc comparisons indicated that 2024 significantly outperformed both 2023 and 2025, a pattern that is also visually evident in Figure 3. The convergence of inferential and descriptive evidence strengthens the conclusion that programme content quality peaked in 2024.

This pattern suggests that curriculum refinements or instructional design adjustments implemented in 2024 had a tangible positive impact on participant perceptions. Such improvements may have included better alignment between objectives and activities, enhanced learning materials, or more effective sequencing of content. The subsequent decline in 2025, while statistically significant, does not necessarily imply deterioration in quality but may reflect challenges in sustaining the same level of innovation or adapting content to evolving participant expectations.

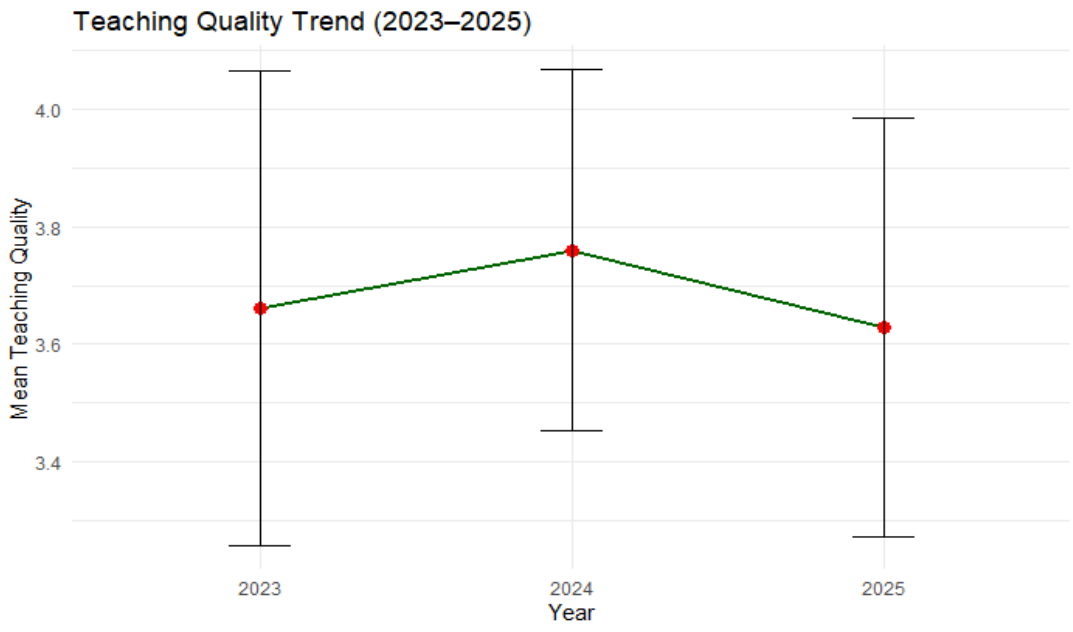
Overall, these findings highlight the importance of continuous curriculum review and instructional design enhancement in professional training contexts. While FITI demonstrates strong performance in defining and communicating programme objectives, the results point to opportunities for enriching content depth and complexity. From a quality assurance perspective, the observed year-to-year variation reinforces the value of longitudinal feedback analysis in identifying both periods of excellence and areas requiring renewed strategic focus.

Trainer Effectiveness

Trainer effectiveness emerged as one of the strongest and most consistently rated dimensions of FITI's training programmes across the three-year period. Participants reported particularly high levels of satisfaction with trainers' subject knowledge, teaching clarity, interaction, and responsiveness, underscoring the central role of instructional delivery in shaping positive training experiences. Item-level correlations further confirmed the internal reliability of the teaching quality scale, supporting its suitability for longitudinal comparison (Cronbach, 1951).

As presented in Table 6, the ANOVA results indicate a statistically significant difference in teaching quality across the three years ($F = 12.08$, $p < 0.001$), leading to the rejection of the null hypothesis ($H0_2$) and acceptance of the alternative hypothesis ($H1_2$). Tukey post-hoc comparisons revealed that 2024 significantly outperformed both 2023 and 2025, while differences between 2023 and 2025 were not statistically significant. This pattern is also clearly illustrated in Figure 4, which shows a pronounced peak in teaching quality during 2024.

***Figure 4**: Teaching Quality Trend (2023 - 2025)*



These findings suggest that investments in trainer development, facilitation skills, or instructional strategies implemented in 2024 may have contributed to enhanced learning experiences during that year. Such investments may include targeted trainer capacity-building

initiatives, improved instructional planning, or greater alignment between trainer expertise and programme content. Although a slight decline in teaching quality was observed in 2025 relative to 2024, this change remains modest in practical terms. Importantly, the average mean teaching quality score remained consistently above 3.6 across all three years, indicating sustained high performance. The observed variation may therefore reflect contextual influences (such as increased participant expectations, greater diversity in programme offerings, or logistical and delivery constraints) rather than substantive shortcomings in trainer effectiveness.

Overall, the results demonstrate that FITI has maintained a strong and reliable instructional foundation over time. The consistently high ratings for teaching quality reinforce the institution's emphasis on trainer competence and effective facilitation, while the peak observed in 2024 highlights the potential benefits of continued investment in trainer development as a strategic priority.

Training Environment and Administrative Facilities

Participants generally reported positive perceptions of the training environment and administrative facilities, highlighting the importance of logistical efficiency and institutional support in fostering a conducive learning atmosphere. As shown in Table 11, meeting space and administrative support received the highest mean ratings, indicating that FITI has been largely successful in providing appropriate physical infrastructure and responsive administrative services to support programme delivery.

***Table 11**: Training Environment and Support: Item-Level Results*

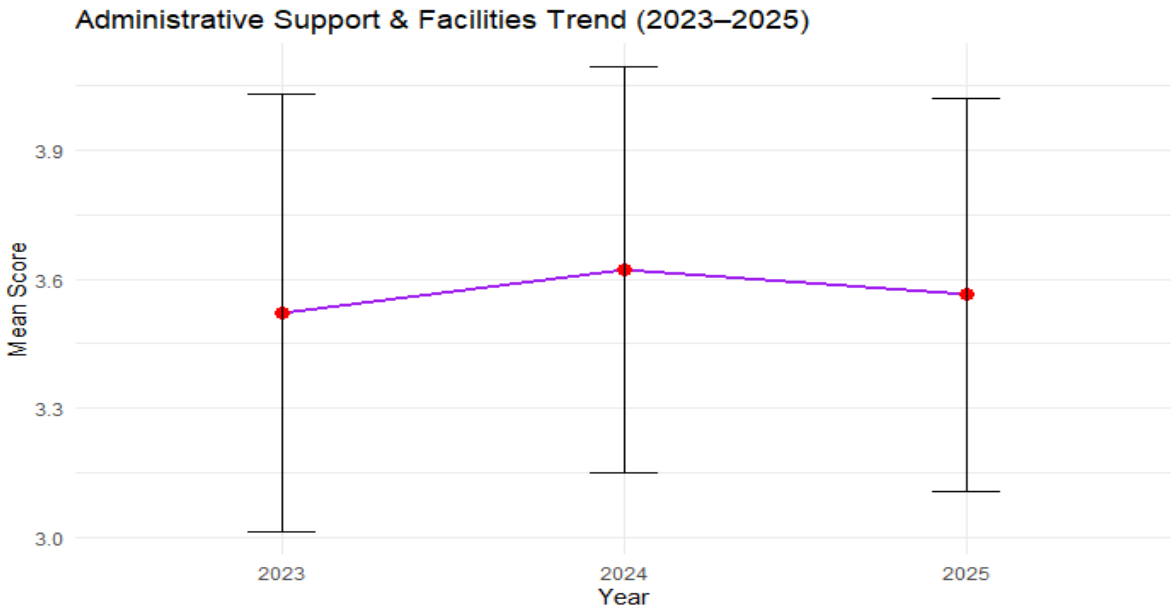
Variable	Mean	SD
Meeting space	3.707	0.521
Facilities in classroom	3.611	0.562
Administrative support	3.657	0.558
Catering services, Refreshment and Lunch	3.300	0.801

Meeting space emerged as the strongest contributor within this dimension, suggesting that venue suitability, seating arrangements, and overall learning ambience played a meaningful role in shaping participant satisfaction. Similarly, the consistently favorable ratings for administrative support reflect the effectiveness of FITI's coordination, communication, and on-site assistance, all of which are critical to smooth programme execution, particularly in professional training contexts.

In contrast, catering services, refreshments, and lunch received comparatively lower mean scores and exhibited greater variability. While such services may appear peripheral to core learning outcomes, the findings suggest that inconsistencies in catering quality can influence participants' overall programme perceptions, especially in multi-day or intensive training settings. Fatigue, comfort, and informal interaction opportunities during breaks can indirectly affect participant engagement and satisfaction, making this an area where targeted improvements could yield marginal yet meaningful gains.

Inferential analysis, as reported in Table 6, revealed a statistically significant difference in perceptions of administrative support and facilities across the three years, leading to the rejection of the null hypothesis ($H0_3$) and acceptance of the alternative hypothesis ($H1_3$). Post-hoc comparisons indicated that the improvement from 2023 to 2024 was statistically significant, while differences involving 2025 were not. This pattern suggests that enhancements implemented in 2024—possibly related to venue selection, logistical planning, or administrative processes—were positively received but were not further amplified in the subsequent year.

***Figure 5**: Training Environment and Administrative Facilities*



Overall, the relatively small year-to-year differences indicate that FITI's logistical and administrative arrangements remained stable and adequate throughout the study period. The findings reinforce the view that while instructional quality and content are central to training

effectiveness, the supporting environment plays a complementary role in shaping participant experiences. Continued attention to consistency in facilities and ancillary services, particularly catering, may therefore strengthen overall programme satisfaction and reinforce FITI's commitment to high-quality training delivery.

Overall Programme Rating and Programme Duration

The findings from the Kruskal–Wallis analyses indicate that participant perceptions of both overall programme quality and programme duration evolved significantly over the three-year period from 2023 to 2025. The significant improvement in overall programme ratings between 2023 and 2024 suggests that enhancements introduced during this period; such as refinements in programme content, instructional delivery, or facilitation practices, were positively received by participants. This improvement is consistent with broader trends observed across other evaluative dimensions, reinforcing the interpretation that 2024 represented a particularly strong year in FITI's training delivery.

Table 12 **:** *Kruskal–Wallis Test Results for Ordinal Outcomes across Years (2023–2025)*

Variable	Test Statistic (χ^2)	df	p-value
Overall Programme Rating	18.135	2	0.000115***
Programme Duration	38.224	2	< 0.001***

The slight decline in overall ratings observed in 2025, while statistically detectable, does not necessarily indicate a substantive deterioration in programme quality. Rather, it may reflect heightened participant expectations following earlier improvements or contextual variations such as changes in participant profiles, programme intensity, or delivery formats. Such fluctuations are common in longitudinal training evaluations and underscore the importance of continuous monitoring rather than reliance on single-year assessments.

With respect to programme duration, the results reveal statistically significant differences across all year-to-year comparisons; however, the magnitude of these differences was relatively modest. This suggests that participants' perceptions of programme length shifted incrementally rather than dramatically, indicating a general consensus that programme duration was appropriate throughout the study period. Minor variations in perception are likely attributable to small adjustments in scheduling, session pacing, or content density, rather than fundamental issues related to time allocation.

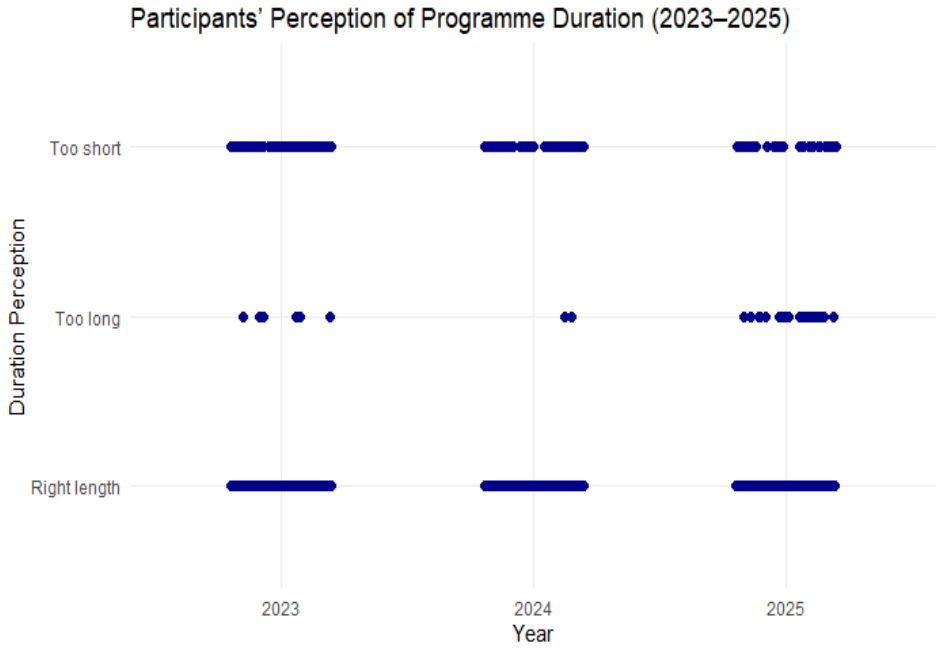
Taken together, these findings highlight the effectiveness of FITI's programme planning and scheduling practices while also emphasizing the value of ongoing, data-driven refinement. The combination of statistically significant yet practically modest changes suggests that FITI has maintained a stable training framework, within which gradual improvements have contributed to sustained participant satisfaction. From a quality assurance perspective, this reinforces the importance of longitudinal feedback analysis as a tool for identifying both periods of notable improvement and areas requiring renewed attention.

Participant Behavioral Intentions

The Chi-square analyses indicate that participants' behavioral intentions, measured through interest in follow-up training and willingness to recommend FITI programmes, remained stable across the three-year period (2023–2025). For both outcomes, the null hypotheses ($H0_6$ and $H0_7$) were not rejected, suggesting that these behavioral intentions were independent of year**.** This finding is particularly noteworthy in the context of other evaluative dimensions that exhibited statistically significant year-to-year variation.

The absence of significant differences in interest in follow-up or advanced training implies a consistently strong level of participant engagement with FITI programmes. Although perceptions of specific programme attributes (such as overall ratings or programme duration) fluctuated modestly over time, participants' desire to continue learning through additional FITI offerings remained largely unchanged. This stability suggests that FITI has succeeded in maintaining the relevance and perceived usefulness of its training programmes, even as incremental adjustments were made to programme structure or delivery.

***Figure 6**: Participants' Perception of programme Duration*



Similarly, the consistently high willingness to recommend the programme to others reflects sustained participant endorsement and institutional credibility. Recommendation behavior is often regarded as a robust indicator of perceived value and trust, as it reflects participants' readiness to associate their professional reputation with the programme. The lack of significant variation across years indicates that FITI's programmes consistently met baseline expectations for quality and effectiveness, regardless of minor shifts in other evaluative measures.

Taken together, these findings highlight an important distinction between evaluative perceptions and behavioral intentions**.** While participants may be sensitive to changes in specific programme characteristics, their overall commitment to and confidence in FITI appears resilient. From an institutional perspective, this underscores FITI's strong brand value and the success of its long-term training strategy. Moreover, the stability of behavioral intentions reinforces the argument that longitudinal feedback analysis should consider both perceptual and behavioral indicators to obtain a more comprehensive understanding of training effectiveness.

5. Conclusion

This study demonstrates the value of systematic longitudinal feedback analysis in evaluating and strengthening professional training programmes at the Financial Institution Training Institute. Across the three-year period from 2023 to 2025, participant feedback consistently reflected positive perceptions of programme design, trainer effectiveness, administrative facilities, and overall training quality. While statistically significant differences were observed across years, particularly highlighting 2024 as a peak period. The margins between years were relatively small, indicating stability rather than volatility in programme performance.

The findings suggest that improvements observed in 2024 may be attributed to curriculum refinement, enhanced trainer performance, and effective delivery practices. The modest declines in 2025 do not undermine overall programme quality but instead highlight the importance of continuous monitoring and incremental improvement. Importantly, participants' willingness to recommend programmes and engage in follow-up training remained stable throughout the study period; reinforcing confidence in FITI's training offerings.

From a broader perspective, this study underscores the usefulness of longitudinal feedback as a quality assurance mechanism in professional training institutions. Rather than relying on single-year evaluations, tracking participant perceptions over time enables institutions to identify emerging trends, assess the impact of interventions, and make evidence-based decisions for programme enhancement.

For future research, analyzing feedback at the individual programme level rather than aggregating across all programmes may yield more nuanced insights. Programme-specific analyses could reveal differential strengths, targeted areas for improvement, and distinct patterns of participant satisfaction that are masked in aggregated data. Incorporating qualitative feedback or linking participant perceptions with learning or performance outcomes would further strengthen understanding of training effectiveness and support more tailored programme development.

Overall, the study provides robust empirical evidence that FITI's training programmes have maintained a high standard of quality over time, while also highlighting the potential for more granular and programme-specific evaluations to drive continuous improvement.

References

Arthur, W., Bennett, W., Edens, P. S., & Bell, S. T. (2003). Effectiveness of training in organizations: A meta-analysis of design and evaluation features. ***Journal of Applied Psychology, 88***(2), 234–245.

Boone, H. N., & Boone, D. A. (2012). Analyzing Likert data. ***Journal of Extension, 50***(2), 1–5.

Burke, L. A., & Hutchins, H. M. (2007). Training transfer: An integrative literature review. ***Human Resource Development Review, 6***(3), 263–296.

Creswell, J. W., & Creswell, J. D. (2018). ***Research design: Qualitative, quantitative, and mixed methods approaches*** (5th ed.). Sage Publications.

Cronbach, L. J. (1951). Coefficient alpha and the internal structure of tests. ***Psychometrika, 16***(3), 297–334.

Field, A. (2018). ***Discovering statistics using IBM SPSS statistics*** (5th ed.). Sage Publications.

Financial Institutions Training Institute Limited. (n.d.). ***About us***. https://fiti.bt/about-us/

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

Noe, R. A. (2020). ***Employee training and development*** (8th ed.). McGraw-Hill Education.

Norman, G. (2010). Likert scales, levels of measurement and the “laws” of statistics. ***Advances in Health Sciences Education, 15***(5), 625–632.

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

Singer, J. D., & Willett, J. B. (2003). ***Applied longitudinal data analysis: Modeling change and event occurrence***. Oxford University Press.

Tavakol, M., & Dennick, R. (2011). Making sense of Cronbach’s alpha. ***International Journal of Medical Education, 2***, 53–55.