UDK: 37:377(71)(047.31)

10.58898/ijmt.v4i2.79-94

PERCEIVED EFFECTIVENESS AND ALIGNMENT OF TRAINING IN CANADIAN EDUCATIONAL INSTITUTIONS

Djordje Jovanovic[1*], Dejan Aladzic[2]

[1]Public College, British Columbia, Canada, djordjejovan@gmail.com

[2]Private College, British Columbia, Canada, dejan.aladzic@gmail.com

Abstract: This research examines how staff members at educational institutions in Canada perceive the effectiveness of training, whether the types of training they deem necessary align with those they participate in, and whether their perspectives change based on years of experience, educational attainment, and institution type. A cross-sectional online survey was conducted involving 50 employees. The effectiveness was evaluated using a five-point scale, with two items in multiple-choice format (types of training deemed necessary and attended). Within-subject comparisons, one-way analysis of variance for comparing multiple groups, and analysis of differences between two independent groups assuming unequal variances were utilized. The findings indicate a significant gap: participants identified more "necessary" training categories than they attended; no differences were observed based on experience; differences in education favour those with higher education levels (with higher and more consistent ratings); private institutions exhibit a more positive distribution of grades compared to public ones, although the difference in mean scores is not statistically significant in this sample. It is concluded that aligning training offerings with expressed needs and acknowledging the trainees' profiles is more crucial than the training experience itself. The study suggested conducting an annual needs assessment that aligns with strategic goals, designing training focused on practical application (through scenarios, practice, mentoring, and implementation strategies), engaging managers before and after training, customizing based on educational segments, and systematically evaluating outcomes with follow-up post-training. Furthermore, responses were anonymous, items were mandatory, and the scale ranged from 1 to 5; caution is advised when interpreting results from the small group with doctoral degrees.

Keywords: training effectiveness, professional development, training needs assessment, transfer of training, higher education.

Introduction

In an era of technological change and growing educational demands, professional training and employee development have become crucial components not only for the success of business and educational institutions, but also for employee satisfaction.

*Corresponding author: djordjejovan@gmail.com

In the past decade, educational institutions – particularly colleges across Canada – have made substantial investments in continuous training programs as strategic goals aimed at improving teaching quality, reducing employee turnover, and ensuring alignment with contemporary standards in pedagogy and administrative efficiency.

Given that institutions often organize various training programs, an important question frequently arises: Are these training programs and the trainings themselves aligned with the actual needs and expectations of employees? It should first be noted that previous studies on this topic have consistently suggested a growing disconnect between the types of training employees want – such as skills that enhance productivity or research competencies – and the trainings that are actually delivered, which often focus on generic or administrative competencies.

The aim of this study is to examine employees' perceptions regarding training and development opportunities in educational institutions across Canada. The study seeks to determine whether significant differences exist between the trainings provided and those employees consider necessary, taking into account the influence of demographic factors such as years of experience, level of education, and type of institution (public versus private), and how these factors shape employees' attitudes toward training.

To achieve the objectives of this study, the following four hypotheses were formulated:

H1: There is a statistically significant difference between the types of training employees believe they need and the types of training that are provided to them in their workplace.

H2: There is a statistically significant difference between respondents' years of work experience and their attitudes toward the process of employee development and training.

H3: There is a statistically significant difference between respondents' levels of formal education and their attitudes toward the process of employee development and training.

H4: There is a statistically significant difference between the type of institution in which respondents are employed (public or private) and their attitudes toward the process of employee development and training.

The findings of this study will contribute not only to a better understanding of training implementation programs in the Canadian educational sector, but also offer practical recommendations for improving institutional support for employee development.

Materials and methods

Research Design and Setting

This study employed a cross-sectional survey of employees working in educational institutions across Canada. The research examined not only perceptions of training and development but also differences in employee attitudes in relation to years of work experience, level of education, and type of educational institution (public/private).

Participants and Sample

As previously noted, data were collected from employees in educational institutions across Canada. Participation in the survey was voluntary and strictly anonymous, and no information that could directly identify individual respondents was collected. It is important to note that the total sample (N = 50) was

used for all analyses, and there were no cases of attrition due to missing data for the examined variables. Differences occurred only in the distribution of respondents across groups within each analysis. In H1, 50 paired measurements (needed/provided) were analysed; in H2, the "16+ years" category contained no observations and was therefore excluded from the ANOVA, while the remaining groups included a total of 50 respondents; in H3, the distribution was 8 (Bachelor's), 38 (Master's), and 4 (Doctorate); and in H4, 30 respondents were from private institutions and 20 from public institutions.

Instrument (Survey)

The survey included single-choice and multiple-choice questions, as well as Likert-scale items. The following variables were used in the analysis:

- **Years of Experience in Current Position:** with the options: 0–2 years, 3–5 years, 6–10 years, 11–15 years, 16+ years.
- **Level of Education:** High School Diploma, Diploma or Certificate, Associate Degree, Bachelor's Degree, Master's Degree, Doctorate.
- **Institution Type:** Private, Public.
- **How would you rate the effectiveness of the training programs you have attended?** Likert scale from 1–5 (Very Effective, Effective, Neutral, Ineffective, Very Ineffective).
- **What types of training do you feel are most needed?** (Select all that apply) multiple choice.*
- **What types of training programs have you attended in the past year?** (Select all that apply) multiple choice.*

(*Options include: Technical Skills, Soft Skills (e.g., communication, teamwork), Leadership Development, Research Skills, Teaching Skills, Productivity Skills, Other.)

Target Objective

The primary outcome in the analyses (Hypotheses H2–H4) is the *Effectiveness Score*, derived directly from the item *"How would you rate the effectiveness of the training programs you have attended?"* where respondents could select only one option on a Likert scale (1 = Very Ineffective, 2 = Ineffective, 3 = Neutral, 4 = Effective, 5 = Very Effective). Considering that the data represent a linearly distributed categorical scale with approximately equal intervals between response categories, the application of parametric tests is justified. Thus, one-way ANOVA (Fisher, 1925) was used for H2–H3, and a t-test (Student, 1908) for two independent groups for H1, while Welch's correction (Welch, 1947) was applied for H4 due to the assumption of unequal variances.

Instead of the Effectiveness Score in Hypothesis H1, the number of different types of training that respondents identified as needed and the number of training types they actually attended in the previous year were used. A two-tailed t-test ($\alpha = 0.05$) was then conducted to determine whether the average number of "needed" trainings was statistically higher than the average number of trainings "attended."

Standardization and Testing

The survey was designed and administered in a way that prevented missing responses; therefore, omission due to missing values was not applicable. Categories with zero frequency (e.g., in H2, "16+ years") were structurally excluded from the specific test. The testing was carried out by organizing the data in MS Excel, and through the available statistical tools, clear answers were obtained for the four

previously stated hypotheses. For H1, two counters were calculated for each respondent (the number of training types marked as "needed" and the number actually "attended"), while for H2–H4, groups were formed based on years of experience, level of education, and type of institution in relation to the effectiveness rating of the trainings attended.

Since the survey included 50 participants, a visual inspection of responses was performed, confirming that all answers were logical and acceptable, resulting in 50 usable responses for further processing. The analysis was conducted in *Data → Data Analysis* using a paired t-test for H1, one-way ANOVA for H2 and H3, and Welch's two-sample t-test for H4. Finally, the key metrics (M, Var/SD, n, t/F, df, p) were generated and the results interpreted at $\alpha = 0.05$. It is also important to emphasize that all between-group analyses assume independence of observations, while H1 is based on paired measurements within the same respondent.

Results

In this chapter, we present a concise overview of the results of the conducted analysis: (i) a sample overview, (ii) hypothesis testing, and (iii) the testing results for hypotheses H1–H4. All tests were two-tailed with $\alpha = 0.05$; where applicable, 95% confidence intervals and effect sizes are reported. On the 1–5 scale, higher values indicated greater perceived effectiveness, while in H1 a positive difference (Needed − Provided) indicated that respondents listed more needed than actually attended types of training.

Sample Overview

The analysis included the complete sample (N = 50); the group distribution depended on the hypothesis:

- **Hypothesis 1** — 50 pairs;
- **Hypothesis 2** — Years of experience: 0–2 (n = 17), 3–5 (n = 17), 6–10 (n = 12), 11–15 (n = 4), 16+ (n = 0);
- **Hypothesis 3** — Education: Bachelor (n = 8), Master (n = 38), Doctorate (n = 4);
- **Hypothesis 4** — Institution type: Private (n = 30), Public (n = 20).

Hypothesis Testing

Table 1 presents the results of the analysis for H1, where a t-test assessed employees' perceptions of "needed" versus "attended" trainings. The average number of training categories considered necessary by respondents was $M_{needed} = 3.62$ (SD ≈ 0.83), while the average number actually attended was $M_{provided} = 2.60$ (SD ≈ 1.98). The mean difference was $\Delta M = 1.02$ and was statistically significant: $t(49) = 3.396$, $p = 0.00136$. The estimated 95% confidence interval for ΔM was approximately [0.42; 1.62], and the Pearson correlation was $r \approx 0.03$ (Table 2).

Table 1. *T-Test: Paired Two-Sample for Means*

	Variable 1	***Variable 2***
Mean	3.62	2.6
Variance	0.689387755	3.918367347
Observations	50	50
Pearson Correlation	0.029800998	
Hypothesized Mean Difference	0	
df	49	
t Stat	3.396305351	
P(T<=t) one-tail	0.000681468	
t Critical one-tailed	1.676550893	
P(T<=t) two-tail	0.001362935	
t Critical two-tailed	2.00957524	

Table 2. *Supplementary metrics for the H1 paired t-test*

Metric	**Value**
Standard Deviation-SD-variable 1	0.83029378
Standard Deviation-SD-variable 2	1.97948664
Standard Deviation of the Difference scores-SD_D	2.12362801
Standard Error-SE	0.30032635
Lower limit	0.4164716
Upper limit	1.6235284

For H2, summarized data are shown in Table 3. A one-way ANOVA was used to examine perceived training effectiveness across groups based on years of experience (0–2, 3–5, 6–10, 11–15; no respondents were in the 16+ category). The scale coding was: “Very Effective” = 5, “Effective” = 4, “Neutral” = 3, “Ineffective” = 2, “Very Ineffective” = 1. Group means (scale 1–5) were: 0–2 M = 3.53 (SD ≈ 1.23), 3–5 M = 3.76 (SD ≈ 0.66), 6–10 M = 3.58 (SD ≈ 1.24), 11–15 M = 3.50 (SD = 1.00), with no data for the 16+ group. Variances ranged from ~0.44 to ~1.54, indicating moderate differences in within-group variability (Table 4).

Table 3. *Distribution of effectiveness ratings by years of experience*

Years of Experience in Current Position	**Very Ineffective**	**Ineffective**	**Neutral**	**Effective**	**Very Effective**	**Grand Total**
0-2 years	2	1	3	8	3	17
11-15 years			3		1	4
3-5 years			6	9	2	17
6-10 years		4		5	3	12
Grand Total	2	5	12	22	9	50

Table 4. *Descriptive statistics by years of experience*

Years of Experience in Current Position	Count	Sum	Average	Variance	Standard Deviation-SD
0-2 years	17	60	3.529412	1.514706	1.230734
3-5 years	17	64	3.764706	0.441176	0.664211
6-10 years	12	43	3.583333	1.537879	1.240112
11-15 years	4	14	3.5	1	1
16+ years	0	0	#DIV/0!	#DIV/0!	#DIV/0!

For **Between Groups**, the sum of squares (SS) was 0.5692, and the mean square (MS) was 0.1423. For **Within Groups**, SS = 51.2108 and MS = 1.1380. Based on this, F = 0.1250, p = 0.9727, with a critical value F_crit = 2.5787 at α = 0.05 (Table 5).

Table 5. *One-way ANOVA by years of experience*

Source of Variation	SS	df	MS	F	P-value	F crit
Between Groups	0.569216	4	0.142304	0.125045	0.972683	2.578739
Within Groups	51.21078	45	1.138017			
Total	51.78	49				

For the analysis of H3, with data presented in Tables 6–8, a one-way ANOVA was applied to compare perceived training effectiveness across groups based on education level (High School Diploma, Diploma or Certificate, Associate Degree, Bachelor's Degree, Master's Degree, Doctorate). The scale was coded as: "Very Effective" = 5, "Effective" = 4, "Neutral" = 3, "Ineffective" = 2, "Very Ineffective" = 1. The distribution of scores by group is shown in Table 6. It is also important to note that survey participants had education levels of Bachelor's Degree, Master's Degree, and Doctorate; the other groups were not represented.

Table 6. *Distribution of effectiveness ratings by level of education*

Level of Education	Very Ineffective	Ineffective	Neutral	Effective	Very Effective	Grand Total
Bachelor's Degree	*1*	*1*	*3*	*3*		*8*
Doctorate				*2*	*2*	*4*
Master's Degree	*1*	*4*	*9*	*17*	*7*	*38*
Grand Total	***2***	***5***	***12***	***22***	***9***	***50***

Table 7 presents individual values by education level, with means ranging from 3.00 (Bachelor's, n = 8) to 4.50 (Doctorate, n = 4), variances from 0.3333 to 1.1429, and standard deviations from 0.58 to 1.07 (Master's: M = 3.6579, Var = 0.9879, SD ≈ 0.99). For Between Groups, SS = 6.2274, df = 2, giving MS = 3.1137; for Within Groups, SS = 45.5526, df = 47, MS = 0.9692 (total SS = 51.78, df = 49). Based on this, F = 3.2126, slightly above F_crit = 3.1951, with p = 0.04923 (Table 8).

Table 7. *Descriptive statistics by level of education*

Level of Education	*Count*	*Sum*	*Average*	*Variance*	*Standard Deviation-SD*
Bachelor's Degree	*8*	*24*	*3*	*1.142857*	*1.069045*
Master's Degree	*38*	*139*	*3.657895*	*0.987909*	*0.993936*
Doctorate	*4*	*18*	*4.5*	*0.333333*	*0.57735*

Table 8. *One-way ANOVA by level of education*

Source of Variation	*SS*	*df*	*MS*	*F*	*P-value*	*F crit*
Between Groups	6.227368	2	3.113684	3.212617	0.049233	3.195056
Within Groups	45.55263	47	0.969205			
Total	51.78	49				

For the analysis of H4, with combined data presented in Tables 9–11, a t-test (assuming unequal variances) was applied to compare perceived training effectiveness between types of institutions (Private vs. Public). The scale coding was: "Very Effective" = 5, "Effective" = 4, "Neutral" = 3, "Ineffective" = 2, "Very Ineffective" = 1 (Table 9).

Table 9. *Distribution of effectiveness ratings by institution type*

Institution Type	Very Ineffective	Ineffective	Neutral	Effective	Very Effective	Grand Total
Public	1	2	7	9	1	20
Private	1	3	5	13	8	30
Grand Total	2	5	12	22	9	50

Table 10 shows the individual statistical parameters by institution type, with means of 3.80 and 3.35, and variances of 1.1310 and 0.871. Based on the t-test, t(44) = 1.5787, p = 0.1216 (two-tailed), with t_crit = 2.0154; the mean difference ΔM = 0.45 was not statistically significant. The standard deviation for the private institution was 1.06, while for the public institution it was 0.93. The estimated 95% confidence interval for ΔM was approximately [-0.12; 1.02] (Table 11).

Table 10. *T-Test: Paired Two-Sample for Means: Private vs. Public institution*

	Variable 1	*Variable 2*
Mean	3.8	3.35
Variance	1.131034483	0.8711
Observations	30	20
Hypothesized Mean Difference	0	
df	44	
t Stat	1.578667703	
P(T<=t) one-tail	0.060787742	
t Critical one-tailed	1.680229977	
P(T<=t) two-tail	0.121575484	
t Critical two-tailed	2.015367574	

Table 11. *Supplementary metrics for H4*

Metric	Value
Standard Deviation-SD-variable 1	1.06350105
Standard Deviation-SD-variable 2	0.933302004
Standard Deviation of the Difference scores-SD_D	2.015611334
Standard Error-SE	0.285050489
Lower limit	-0.124481512
Upper limit	1.024481512

Hypothesis Testing Results

The results for H1 indicate that the average difference reflects a clear pattern of a structural gap: the number of "attended" training types shows more than twice the variability compared to "needed" trainings (Var ≈ 3.92 vs. 0.69; SD ≈ 1.98 vs. 0.83), suggesting that attendance among respondents is uneven (some attend very few or none, while others attend more), whereas perceived needs are considerably more stable. Additionally, the correlation between paired observations is negligible (r ≈ 0.03), indicating that those who report more needs are not necessarily the ones who attend more—thus, provision/participation is not aligned with individual needs.

Considering the gap measure D = Needed − Provided, the mean difference is ΔM = 1.02 with SD_D ≈ 2.12 and 95% CI [0.42; 1.62]. This means that a "typical" respondent reported approximately one more category as needed than actually attended, while there is significant heterogeneity in the gap (some respondents have minimal or even negative gaps). The statistical significance itself (paired t(49) = 3.396; p = 0.00136) arises primarily from the consistently positive shift in the mean rather than a favorable covariance structure (since r ≈ 0). In other words, the result suggests a systematic but moderately large deficit in coverage of needed trainings at the sample level, with noticeable individual differences.

Conclusion (H1). The paired t-test showed that the difference between the average number of trainings respondents considered necessary and those actually attended is statistically significant: t(49) = 3.396, p = 0.00136 (two-tailed), ΔM = 1.02, 95% CI [0.42; 1.62]. In other words, H1 is confirmed, and it can be concluded that:

There is a statistically significant difference between the types of training that employees believe they need and the types of training that are provided to them in the workplace.

Regarding H2, four experience groups were examined, and the mean ratings were very close (3.50–3.76), although clear patterns of distribution are observable within them. In the 0–2 years group, **Effective** (8/17 = 47.1%) and **Very Effective** (3/17 = 17.6%) dominated, but all "very negative" ratings appeared in the sample (2/17 = 11.8%), along with one **Ineffective** and three **Neutral** ratings (5.9% and 17.6%, respectively), resulting in higher variability (SD ≈ 1.23). The 3–5 years group was the "most stable": no negative ratings occurred, with **Effective** (9/17 = 52.9%) and **Neutral** (6/17 = 35.3%) prevailing, yielding the highest mean (M = 3.76) and the lowest dispersion (SD ≈ 0.66).

In the 6–10 years group, a local decline is observed due to a higher share of **Ineffective** ratings (4/12 = 33.3%), alongside **Effective** (5/12 = 41.7%) and **Very Effective** (3/12 = 25.0%); this mixture produces greater variability (SD ≈ 1.24). The 11–15 years group is small (n = 4) and mostly neutral (Neutral 3/4), with one **Very Effective** rating, giving a mean of 3.50 and SD = 1.00.

However, these local differences in profiles do not translate into stable differences in means at the sample level. The one-way ANOVA was not significant, F(4,45) = 0.125, p = 0.9727, with most variability occurring **within groups** (SS_within = 51.21) rather than **between groups** (SS_between = 0.57). The estimated partial $\eta^2 \approx 0.011$ indicates a negligible effect.

Conclusion (H2). Considering the p-value = 0.9727 > 0.05 in this sample, H2 is **not confirmed**, and it can be concluded that:

There is no statistically significant difference between respondents' years of work experience and their attitudes toward the employee development and training process.

The one-way ANOVA for H3 results yielded a significant outcome: F(2,47) = 3.21, p = 0.049. The group means and dispersion were as follows: Bachelor's M = 3.00 (SD ≈ 1.07, n = 8), Master's M = 3.66 (SD ≈ 0.99, n = 38), and Doctorate M = 4.50 (SD ≈ 0.58, n = 4). A clear pattern underlies these means: the Doctorate group not only gives higher ratings but does so more consistently (lowest SD), indicating that most doctoral respondents consistently perceive trainings as effective. In contrast, the Bachelor's group shows a lower mean and greater variability, suggesting uneven experiences with trainings in that group.

The differences between means are also practically relevant: Doctorate − Bachelor's ≈ 1.50, Master's − Bachelor's ≈ 0.66, Doctorate − Master's ≈ 0.84, reflecting the expected order Doctorate > Master's > Bachelor's. The estimated partial effect size is $\eta^2_p \approx 0.12$ (moderate effect). However, results should be interpreted cautiously due to the small Doctorate group (n = 4) with lower variance; for final conclusions regarding group pairs, post hoc comparisons are recommended if variance inequality is suspected.

Conclusion (H3). Hypothesis H3 is confirmed: education level is associated with perceived effectiveness of training programs; respondents with higher education degrees provide, on average, higher and more consistent effectiveness ratings, with the strongest positive shift observed in the Doctorate group. Therefore, it can be concluded that:

There is a statistically significant difference between respondents' formal education level and their attitudes toward the employee development and training process

Looking more closely at H4, the distribution of ratings by institution type shows that very negative ratings are almost absent in both private and public institutions (5% in public vs. 3.3% in private), with most responses positioned at the top of the scale. In private institutions, high ratings dominate: **Effective + Very Effective = 21/30 (70%)**, with **Very Effective** at 26.7% (8/30). In contrast, public institutions show less representation at the top of the scale: **Effective + Very Effective = 10/20 (50%)**, with **Very Effective** only 5% (1/20), while **Neutral** is considerably more frequent — 35% (7/20) compared to 16.7% (5/30) in private institutions. Negative ratings are rare and similar in frequency across both sectors: 13.3% in private (Very Ineffective 3.3% + Ineffective 10%) versus 15% in public (5% + 10%).

This profile clearly explains the difference in means ($M_{private} = 3.80$ vs. $M_{public} = 3.35$; $\Delta M = 0.45$). The difference does not arise from negative ratings, which are rare in both sectors, but from the replacement of neutral ratings in public institutions with very positive ratings in private institutions.

Conclusion (H4). Although mean ratings are higher in private institutions ($M = 3.80$) than in public institutions ($M = 3.35$), the independent-samples t-test with unequal variances did not show a statistically significant difference: $t(44) = 1.58$, $p = 0.122 > 0.05$ (two-tailed), $\Delta M = 0.45$, 95% CI [−0.12; 1.02]. Therefore, H4 is **not confirmed**, and it can be concluded that

There is no statistically significant difference **between the type of institution where respondents work (public or private) and their attitudes toward the employee development and training process.**

Discussion

Brief Synthesis of Key Findings

This study reveals a persistent discrepancy between the training that employees perceive as necessary and the training they actually receive (H1). Notably, the duration of work experience did not significantly influence perceptions of training effectiveness (H2). In contrast, educational attainment emerged as a significant differentiating factor (H3); individuals with higher education levels consistently provided more favourable assessments of the training programs. Although a trend suggested that private institutions received higher ratings for effectiveness—indicated by a greater proportion of "Very effective" responses compared to "Neutral"—this observation did not achieve statistical significance in the current sample (H4).

The findings are interpreted through established theoretical frameworks, including human capital theory, learning motivation, and transfer climate, as articulated by Becker (1993), Baldwin and Ford

(1988), and Colquitt et al. (2000), as well as Blume et al. (2010). These frameworks offer valuable insights into how employees' evaluations of training effectiveness are influenced by their educational background, institutional affiliation, and intrinsic motivation to enhance their professional skills.

Interpretation of Research Questions

H1 — Needs and Supply Match

A significant gap in the "needs attended" metric indicates potential shortcomings in the needs assessment process (Training Needs Analysis, TNA) and/or limitations in training access (Aguinis & Kraiger, 2009; Salas, Tannenbaum, Kraiger, & Smith-Jentsch, 2012). This finding serves as a signal for management that the alignment of planning, prioritization, and resource allocation with employees' explicit needs may not be sufficiently robust. Furthermore, such a gap may contribute to diminished perceptions regarding the effectiveness of subsequent training programs (Arthur, Bennett, Edens, & Bell, 2003).

H2 — Seniority as a Non-Differentiator

The absence of discernible differences based on seniority suggests that the context and climate surrounding the transfer of knowledge are more influential than the "age/experience effect" on training perceptions (Baldwin & Ford, 1988; Burke & Hutchins, 2007). It appears that learners, regardless of their seniority, encounter similar design and delivery formats, as well as comparable constraints regarding the application of acquired knowledge, leading to minimal impact of seniority on perceptions (Ng & Feldman, 2010).

H3 — Education as a Differentiator

Respondents with higher levels of education generally provided higher and more consistent ratings. This observation is consistent with theoretical models that correlate cognitive and motivational factors with learning outcomes (Kraiger, Ford, & Salas, 1993; Colquitt et al., 2000). Higher educational attainment may facilitate a better alignment between the content and prior knowledge, promote stronger self-efficacy, and clarify expected benefits, which collectively enhance the perceived value of the training program (Blume et al., 2010). However, this finding warrants cautious interpretation due to the small sample size within the Doctorate group.

H4 — Institutional Context

While private institutions reported a higher prevalence of "Very effective" ratings and fewer "Neutral" ratings than their public counterparts, the differences observed were not statistically significant. This may be attributed, in part, to unequal group sizes and the presence of a “ceiling” effect, wherein both sectors exhibited few negative ratings, thus limiting variance and the analytical power of the results (Boyne, 2002; Perry & Wise, 1990). The disparities likely arise more from variations in climate, support, and conditions for transfer rather than solely from the type of institution (Salas et al., 2012).

Integration of Findings, Theoretical Frameworks, and Perception Mechanisms of Training Effectiveness

The empirical findings can be effectively understood through three thematic axes and their theoretical foundations.

First, there is the alignment of needs and supply. The central outcome hinges on the gap between what employees recognize as necessary and what training they actually receive. According to human capital theory, investments in training yield results only when they target relevant competencies and address genuine development gaps (Becker, 1993). Therefore, "gap-closure" should become an explicit goal of the training and development (T&D) system, which should involve steps such as diagnosis, prioritization, design and delivery, and transfer monitoring (Salas et al., 2012; Arthur et al., 2003).

Second, the characteristics of the trainees play a significant role. Research serves as a crucial differentiator in the perception of training effectiveness. This aligns with models that highlight how cognitive readiness, motivation to learn, and goal orientation enhance the perceived value and transfer of learning (Kraiger, Ford, & Salas, 1993; Colquitt, LePine, & Noe, 2000; VandeWalle, 1997). Higher education levels likely lead to a better "fit" between the training content and existing knowledge schemas, as well as increased self-efficacy, resulting in more positive and consistent outcomes (Blume, Ford, Baldwin, & Huang, 2010). In contrast, simply having numerous years of experience—without a supportive environment for applying new learning—rarely changes perceptions (Baldwin & Ford, 1988; Burke & Hutchins, 2007).

Third, the organizational context, whether public or private, impacts perceptions. Differences do not manifest merely through a higher proportion of negative ratings; instead, they arise in the distribution of ratings between "neutral" and "very effective." This suggests that the crux of the issue lies in the design, resources, and support for transfer (such as mentoring, feedback, and job relevance), rather than ownership of the training programs themselves (Boyne, 2002; Perry & Wise, 1990; Salas et al., 2012). Organizations with a stronger post-training implementation framework are more likely to see shifts in ratings from "neutral" to "very effective," even when the programmatic themes remain consistent.

Lastly, measurement and evaluation should extend beyond just the level of reaction. A taxonomy of outcomes—cognitive, skill-based, and affective—provides a clear map of what to assess and when (Kraiger et al., 1993). Additionally, a practical evaluation framework (such as the Kirkpatrick Model, which includes levels 1–3 routine and level 4 selective) instills discipline in tracking transfer and impact (Kirkpatrick & Kirkpatrick, 2006). By anchoring these thematic axes in proven mechanisms, the likelihood of achieving gap-closure increases—not just in perception, but also in actual behaviour and work outcomes.

Implications for Training Policy

The practical implications can be summarized in a few key steps that influence both the quality of the training program and its measurable outcomes. First, it is essential to implement an annual training needs assessment (TNA) that is linked to the institution's strategic goals. This will help ensure that "closing the gap" between the training needed and the training attended is recognized as a formal outcome, measured through official Key Performance Indicators (KPIs). Establishing this link ensures that development investments are directed toward areas with the most significant development gaps and the tremendous potential for impact.

Additionally, a transfer-oriented design should become the standard approach. This involves incorporating real-world scenarios and tasks, providing guided practice in the workplace, offering

mentoring support, and establishing clear post-training implementation plans (i.e., specifying who will do what and by when). These elements significantly enhance the likelihood that participants will apply what they learn effectively in their work environments (Salas et al., 2012). Moreover, personalization based on educational levels—adapting depth, pace, and examples to Bachelor's, Master's, and Doctorate profiles—is crucial, as different levels of initial knowledge and expectation patterns affect learning acquisition and effectiveness assessment.

A key factor in this process is the organizational climate and support. Involving managers in setting goals before training and in systematically following up after training (monitoring implementation, providing feedback, and addressing obstacles) has been shown to encourage knowledge transfer and reinforce the effects of the program (Burke & Hutchins, 2007). This process—needs assessment, design for transfer, and post-training support—creates a continuous loop between individual learning and organizational outcomes.

When it comes to measurability and evaluation, we recommend implementing an operational dashboard that tracks key indicators. Key metrics should include participation and completion rates, post-test results, indicators of knowledge transfer (such as structured follow-up at 30–90 days), an overall behavioural change rating, and relevant KPIs (e.g., output quality, cycle time). Evaluations should be conducted systematically according to the Kirkpatrick framework, with routine assessments at levels 1–3 and selective assessments at level 4, gradually incorporating more objective performance measures (Aguinis & Kraiger, 2009; Arthur et al., 2003). Only after accumulating data on the impact on business metrics should a cautious assessment of ROI be conducted, keeping in mind that complex development effects can rarely be boiled down to a single figure (Aguinis & Kraiger, 2009).

Study Limitations, Future Research Directions, and Final Message

This study has several methodological limitations that affect the scope of our conclusions. First, the design is cross-sectional and relies on self-reports, which limits our ability to establish causality and may introduce perceptual biases. Second, the groups in the study are unequal, with a small sample size (n=4) in the Doctorate group, which reduces statistical power and the stability of the estimates. Third, a ceiling effect was observed, where there were few negative ratings, reducing variability and the sensitivity of the tests to detect subtle differences. Lastly, while we treated the 1–5 scale as approximately interval—a common practice in survey research—the results would benefit from confirmation using nonparametric or robust approaches, as well as longitudinal follow-up on transfer (Baldwin & Ford, 1988; Blume et al., 2010).

To address these limitations, future research should include: (a) post hoc comparisons using methods for unequal variances (e.g., Games–Howell) for Hypothesis 3, to better estimate the differences among the Doctorate, Master's, and Bachelor's groups; (b) detailed documentation of program characteristics (such as format, duration, work experience, and mentoring), as design is known to influence outcomes significantly (Arthur et al., 2003); and (c) a mixed-method approach that incorporates qualitative insights into the "gap" identified in Hypothesis 1 and the barriers to transfer. Additionally, expanding the sample size, particularly in public institutions and within the Doctorate group, while

introducing longitudinal metrics, will enhance our ability to assess behavioural changes and performance, rather than just immediate reactions.

Ultimately, our practical focus should be on closing the gap identified in Hypothesis 1 through systematic needs assessments linked to strategic goals and transfer-oriented design. This involves incorporating job scenarios, on-the-job practice, mentoring support, and implementation plans. By personalizing education across different segments and systematically evaluating outcomes (using Kirkpatrick Levels 1–3 routinely and Level 4 selectively), educational institutions can improve both perceived and actual effectiveness. This emphasis aligns with evidence suggesting that thoughtful design and organizational support have a more significant impact on outcomes than formal attributes, such as tenure, while also acknowledging that the educational profile of the learner remains an important moderator to be considered upfront (Aguinis & Kraiger, 2009; Salas et al., 2012; Colquitt et al., 2000).

Conclusions

This research examined the connection between estimated training needs, actual programs participated in, and perceptions of effectiveness, while considering the diverse characteristics of participants and the institutional context within educational organizations in Canada. The key findings are as follows:

1. There is a consistent disparity between the training considered "necessary" and what is actually "attended."

2. Years of experience do not appear to be a distinguishing factor in perceptions of effectiveness.

3. Individuals with higher levels of formal education tend to report higher and more consistent effectiveness ratings.

4. A trend favouring private institutions exists, though no statistically significant difference is noted when compared to public institutions.

In conclusion, the outcomes suggest that the connection between needs and supply, as well as the characteristics of participants, significantly influence results more than the formal measurement of “length of experience.” The institutional context (public or private) affects the distribution style of ratings, but does not indicate a reliable difference within the sampled environments. Therefore, training policies should focus on needs assessments, be designed for practical application, and be tailored to various educational segments.

Practical implications include:

1. Establishing annual training needs assessments (TNAs) aligned with strategic goals and KPIs, with a clear objective to "bridge the gap" identified in finding 1.

2. Implementing transfer-oriented design consistently, incorporating real business tasks, on-the-job practice, mentoring support, and an execution plan, along with systematic follow-ups.

3. Customizing the depth, pace, and content examples based on educational level (Bachelor/Master/Doctorate) and conducting routine evaluations (Kirkpatrick levels 1–3, and selectively level 4) with a greater emphasis on objective performance metrics.

Theoretical implications suggest that the findings align with theories of human capital, learning motivation, and training transfer. The educational background of trainees serves as an important moderator in assessing effectiveness, while years of service do not show a different impact in this scenario.

Limitations of the study include a cross-sectional design, reliance on self-reported data, imbalanced groups (especially a small sample size in the Doctorate category), and "ceiling" effects. The sample context constrains the generalizability of the findings. Recommendations for future research involve broadening the sample (particularly in public institutions and among higher education levels), applying post-hoc comparisons under conditions of unequal variances (e.g., Games–Howell) for finding 3, incorporating longitudinal follow-ups on training transfer, utilizing more objective performance indicators, and documenting additional design characteristics of the training.

In summary, educational institutions can enhance both the perceived and actual effectiveness of their programs by systematically narrowing the gap between needs and provisions, designing training that supports practical application, and tailoring educational pathways to align with the profiles of the trainees.

Acknowledgments

We appreciate the time and thoughtful responses of all survey participants. We also extend our gratitude to the colleagues who helped with language editing, data preparation, and proofreading. Any remaining errors are solely our responsibility.

Conflict of interest

The authors declare that they have no conflict of interest.

References

Aguinis, H., & Kraiger, K. (2009). Benefits of training and development for individuals and teams, organizations, and society. Annual Review of Psychology, 60, 451–474. https://doi.org/10.1146/annurev.psych.60.110707.163505

Arthur, W., Jr., Bennett, W., Jr., 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. https://doi.org/10.1037/0021-9010.88.2.234

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

Becker, G. S. (1993). Human capital: A Theoretical and Empirical Analysis with Special Reference to Education (3rd ed.). University of Chicago Press.

Blume, B. D., Ford, J. K., Baldwin, T. T., & Huang, J. L. (2010). Transfer of training: A meta-analytic review. Journal of Management, 36(4), 1065–1105. https://doi.org/10.1177/0149206309352880

Boyne, G. A. (2002). Public and private management: What's the difference? Journal of Management Studies, 39(1), 97–122. https://doi.org/10.1111/1467-6486.00284

Burke, L. A., & Hutchins, H. M. (2007). Training transfer: An integrative literature review. Human Resource Development Review, 6(3), 263–296. https://doi.org/10.1177/1534484307303035

Colquitt, J. A., LePine, J. A., & Noe, R. A. (2000). Toward an integrative theory of training motivation: A meta-analytic path analysis of 20 years of research. Journal of Applied Psychology, 85(5), 678–707. https://doi.org/10.1037/0021-9010.85.5.678

Fisher, R. A. (1925). Statistical methods for research workers. Oliver & Boyd.

Games, P. A., & Howell, J. F. (1976). Pairwise multiple comparison procedures with unequal N's and/or variances. Journal of Educational Statistics, 1(2), 113–125.

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

Kraiger, K., Ford, J. K., & Salas, E. (1993). Application of cognitive, skill-based, and affective theories of learning outcomes to new methods of training evaluation. Journal of Applied Psychology, 78(2), 311–328. https://doi.org/10.1037/0021-9010.78.2.311

Ng, T. W. H., & Feldman, D. C. (2010). The relationships of age with job attitudes: A meta-analysis. Personnel Psychology, 63(3), 677–718. https://doi.org/10.1111/j.1744-6570.2010.01184.x

Perry, J. L., & Wise, L. R. (1990). The motivational bases of public service. Public Administration Review, 50(3), 367–373. https://doi.org/10.2307/976618

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. https://doi.org/10.1177/1529100612436661 (Original work published 2012)

Student. (1908). The probable error of a mean. Biometrika, 6(1), 1–25. https://doi.org/10.1093/biomet/6.1.1

VandeWalle, D. (1997). Development and validation of a work domain goal orientation instrument. Educational and Psychological Measurement, 57(6), 995–1015. https://doi.org/10.1177/0013164497057006009

Welch, B. L. (1947). The generalization of "Student's" problem when several different population variances are involved. Biometrika, 34(1–2), 28–35. https://doi.org/10.1093/biomet/34.1-2.28