Canadian
Forces
College

Collège
des
Forces
Canadiennes

Training Evaluation in the Canadian Armed Forces

Lieutenant-Colonel Nicola Way

JCSP 48	PCEMI 48
Service Paper	**Étude Militaire**

Canada

CANADIAN FORCES COLLEGE – COLLÈGE DES FORCES CANADIENNES

JCSP 48 – PCEMI 48
2021 – 2022

Service Paper – Étude militaire

Training Evaluation in the Canadian Armed Forces

Lieutenant-Colonel Nicola Way

TRAINING EVALUATION IN THE CANADIAN ARMED FORCES

AIM

1.   This service paper will review training evaluation paradigms within the Canadian Armed Forces (CAF) and recommend potential training policy revisions to Chief Military Personnel (CMP). This will be achieved by analyzing CAF training policy and applying academic theory to assess whether training interventions within the cognitive and affective domains are subjected to the most appropriate and effective levels of evaluation. The recommendations presented in this service paper seek to improve learning transfer post-training, thus increasing the probability of training events that result in long-term behavioral and organizational change.

INTRODUCTION

2.   Training and professional development are the cornerstones of military capability. 'Conventional' military training is role-related, knowledge and skills-based, and is within the cognitive or psychomotor domains of learning.[1] Yet increasingly training is also seen as a preferred solution to cultural, behavioral, or attitudinal performance needs within the affective domain of learning.[2] These training requirements often result from an organizational desire to *do something* in response to an issue or complaint and can be hastily designed and implemented. The Deschamps review highlighted that, before Operation HONOUR, CAF training in the affective domain had been ineffective for several reasons,[3] and that by "failing to inform members about appropriate conduct or to inculcate an ethical culture in the CAF"[4] it had significantly contributed to the organization's overall failings. Defence is not alone in treating similar subject matter as a "check-in-the-box" procedure. Often individuals are assessed only to the degree to which they complete the training,[5] organizations resting instead with the belief that if training is well designed it must be effective.[6] Rather, organizations must ensure that individuals also understand their responsibilities in these often sensitive areas and can change their behaviors accordingly, not just pass an assessment of acquired knowledge.

---

[1] The domains of learning (cognitive, psychomotor, and affective), which were first coined by Bloom in 1956 in his seminal work on learning taxonomy, are well known and frequently cited within training and education professions including the military. They are the basis upon which Knowledge, Skills, and Attitude (KSA) learning outcomes were developed.

[2] Examples, drawn from CAF and close allies (UK and US), include training interventions for topics such as unconscious bias, sexual conduct, bullying and harassment, gender, race, respect for others, and ethics, amongst others.

[3] Marie Deschamps C.C. Ad. E. *External Review into Sexual Misconduct and Sexual Harassment in the Canadian Armed Forces,* (Ottawa: External Review Authority, 27 March 2015): 81-85. https://www.canada.ca/en/department-national-defence/corporate/reports-publications/sexual-misbehaviour/external-review-2015.html. For several reasons, including, failure to take the topic seriously, too much theory, not enough time to deliver, poor delivery methods, and ineffective trainers.

[4] *Ibid.*, vi.

[5] Eduardo Salas, Laura M. Milham and Clint A. Bowers. "Training Evaluation in the Military: Misconceptions, Opportunities, and Challenges." *Military Psychology,* 15, no. 1 (2003): 4. https://www.tandfonline.com/doi/abs/10.1207/S15327876MP1501_01?journalCode=hmlp20

[6] Salas, "Training Evaluation in…, 3.

3.   The Deschamps review recommended that CAF training be reanalyzed and two new training interventions created: "Bystander Intervention" and "Respect in the CAF" (RitCAF). To ensure success the review outlined that the CAF use external training deliverers, real-life scenarios, regular touchpoints, and allocate more training time. However, the report fails to mention evaluation or validation of this training, using the term "monitoring" instead.[7] Within training, monitoring means tracking completion rates, observing delivery, and ensuring training content. Monitoring does not imply an evaluation of behavioral change. This is not unusual; academic research on numerous organizations has found that even the most critical training programs rarely have an adequate plan for evaluating effectiveness.[8]

4.   This service paper will use Kirkpatrick's model of training evaluation to explain the criticality of thorough training evaluation. It will define key terms, analyze evaluation theory versus CAF practice, discuss why Kirkpatrick's level-three evaluation is the most important for long-term success, and provide tangible recommendations.

DISCUSSION

Training overview

5.   Training occurs after identifying a real or perceived deficiency in capability. Within Defence training is the means to "practice, develop and assure, within constraints, the practical application of a common military doctrine to deliver a military capability."[9] It is achieved through a series of complementary activities described below[10] that are common to all types of individual and collective military training:

   a.   *Training*. The activity that provides the KSA required by individuals to perform specific duties, functions, and tasks.[11] Training tackles known problems.[12]

   b.   *Education*. The activity that provides a base of knowledge and intellectual skills upon which information can be correctly interpreted and sound judgment

---

   [7] Deschamps. *External Review into…,* 85. Recommendation no. 10 states: "Assign to the center for accountability for sexual assault and harassment, in coordination with other CAF subject matter experts, responsibility for the development of the training curriculum, and the primary responsibility for monitoring training on matters related to inappropriate sexual conduct."
   [8] James D. Kirkpatrick and Wendy Kayser Kirkpatrick, *Four Levels of Training Evaluation,* (Virginia: ATD Press. 2006), 49. https://books.google.ca/books=onepage&q=kirkpatrick&f=false
   [9] British Army, *Army Command Standing Order No. 3249: Conduct of Army Training,* (UK MOD: Crown Copyright, October 2021), 3.
   [10] There are no universally accepted definitions for these activities, and variations are apparent across even the 5-eyes community. These definitions are taken from CAF doctrine but expanded to include critical concepts from US and UK doctrine.
   [11] Department of National Defence, A-P9-050-000/PT-001, *CFITES Manual of Individual Training and Education, Volume 1: Interim Guidance – Introduction/Description* (Ottawa: DND Canada, 2003), 5.
   [12] British Army, *Army Command Standing Order…*, 3.

applied.[13] Education develops a broader range of outcomes and prepares for novel and uncertain situations.[14]

     c.   *Experience*. The degree to which the outcomes of training and education are manifested and reinforced in the workplace or other real-world settings.

     d.   *Learning*. A neuro-physical process that is the critical output of training, education, and experience. Learning takes place if one or more occurs: knowledge increases, skill improves, and/or attitudes change.[15]

6.   The distinction between training and education is an imperfect one.[16] The boundary is even more blurred when training interventions bridge the cognitive and attitudinal domains, e.g., respect training. Learning outcomes for these courses should manifest across all three learning outcomes, i.e., trainees gain knowledge about Respect (cognitive), can apply the knowledge automatically to new scenarios (skill-based), and modify their attitudes in a more positive direction (affective).[17] The learning outcomes are critical as the training is ineffective unless the learning is implemented. This process of applying taught knowledge to real-life is called learning transfer.

Evaluation overview

7.   Evaluation is the process of making a judgment as to the worth of training to the organization.[18] Like training, it is achieved through a series of complementary activities,[19] however, terms such as evaluation, assurance, validation, audit, monitoring, and inspection are often, wrongly, regarded interchangeably which can lead to organizational confusion, incoherent implementation, and inefficient practices. As a result, many organizations undertake training evaluation hesitantly, incorrectly, and often only with elemental methods.[20]

---

[13] Department of National Defence, A-P9-050-000/PT-001…, 5.

[14] British Army, *Army Command Standing Order…,* 3.

[15] Donald Kirkpatrick and James Kirkpatrick. *Evaluating Training Programmes*, 3rd ed. (San Francisco: Berrett-Koehler Publishers, Inc. 2006), 22. https://www.proquest.com/docview/199686343/fulltext/53C30DF4AA7B4352PQ/1

[16] Allison Abbe, "Evaluating Military Cross-Cultural Training Programs." *Expeditions with MCUP [on line journal],* last modified [or accessed] 8 January 2022, https://doi.org/10.36304/ExpwMCUP.2021.06

[17] Melissa K. Preusser, Lynn K. Bartels Ph.D., and Cynthia A. Nordstrom PhD, "Sexual Harassment Training: Person Versus Machine." *Public Personnel Management* 40, no. 1 (Spring, 2011): 49, https://www.proquest.com/scholarly-journals/sexual-harassment-training-person-versus-machine/docview/1664817684/se-2?accountid=9867.

[18] Ministry of Defence, *JSP 822: Defence Direction and Guidance for Training and Education Part 1: Directive.* V4.1, (UK MOD: Crown Copyright, November 2021), 27.

[19] Again, there are no universally accepted definitions for these activities. The terminology used within the paper is founded in CAF doctrine but draws on best practices from UK and US doctrine.

[20] Sten Langmann and Stefan Robert Thomas, "Rethinking training evaluation: omnidirectional evaluation, range of audience and intentional change." *The International Journal of Human Resource Management,* 30, no. 12, (2019): 1904.

Conducting evaluation

8.   Training evaluation within Defence[21] uses the Kirkpatrick four-stage model of evaluation, as illustrated in Figure 1. The model is a simple and systematic approach to training evaluation.[22]
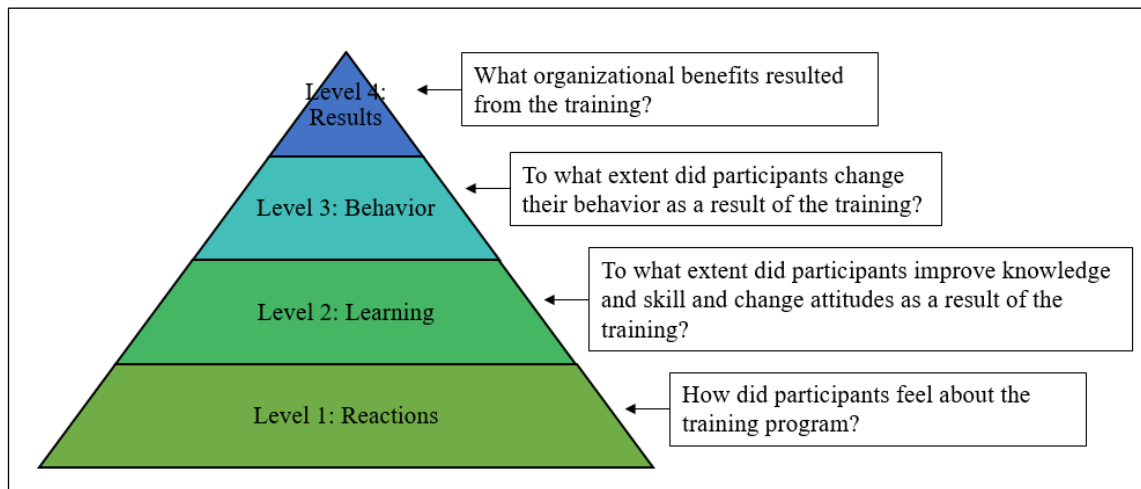

Figure 1 - Kirkpatrick's Four Stage Evaluation Model
Source: Kirkpatrick, *Evaluating Training Programmes*, 3rd ed., 19.

9.   Level one, assessment of students' reactions to the training, and level two, assessment of whether learning has taken place, are both measured during or immediately after the training intervention.[23] Level three assesses whether learning transfer has occurred and if the student is implementing the training in the real world through a change in their behavior. Level four measures the impact of training on the organization.[24] The first three levels are measured at the individual level; the fourth is an organizational assessment.

10.   Of note, the levels represent a sequence in not just evaluation, but also successive steps in any training and learning event, i.e., there are no results without behavioral change, no change in behavior without learning taking place, and there is limited learning without a positive reaction.[25] Conversely and importantly, Kirkpatrick emphasized that a favorable reaction will not guarantee that either: learning has taken place, positive

---

[21] Kirkpatrick's four-stage evaluation model is universal and is the basis for evaluation by the Canadian Armed Forces, US, and British Military to name just some.

[22] Langmann, "Rethinking training…, 1905.

[23] Abbe, "Evaluating Military…

[24] Reid Bates, "A critical analysis of evaluation practice: the Kirkpatrick model and the principle of beneficence." *Evaluation and Program Planning*, 27, no. 3 (August 2004): 341, https://doi-org.cfc.idm.oclc.org/10.1016/j.evalprogplan.2004.04.011

[25] Vit Dočekala and Miroslava Dvořákováb. "Evaluation Levels in Education." *Procedia – Social and Behavioral Sciences,* 174, no. 12 (February 2015): 3744, https://doi-org.cfc.idm.oclc.org/10.1016/j.sbspro.2015.01.1108

behavioral change will occur, or organizational results will be achieved. This is why all levels of the model should be evaluated.[26]

Evaluation in practice

11.   In theory, the CAF conducts training evaluation following Kirkpatrick's model. The Canadian Forces Individual Training and Education System (CFITES) manual outlines two phases of evaluation.[27] First, the evaluation of the student and the efficacy of the training program concerning student learning. This aligns to level-one and level-two evaluation. Second, the validation of whether the training intervention meets the verifiable, current needs of the Armed Forces. This aligns with to level-three and level-four evaluation.

12.   However, Kirkpatrick's four levels represent a sequence or continuum of complexity. The evaluation process becomes more difficult and time-consuming as it moves from one level to the next, but it also provides increasingly more valuable information.[28] The increasing level of complexity means that organizations, including the CAF, will often prioritize more simplistic evaluations measuring student reactions and competency (level one 'happy sheets' and level two end-of-training assessments) over conducting level-three behavior-based and level-four results-based evaluations.[29]

13.   Research conducted in Canada[30] revealed that organizations conduct level one evaluations 71% of the time, but only conduct level four 42.8%. Similarly, a 2016 Association for Talent Development survey found that 88% of organizations relied on level one trainee reactions as the primary means of evaluation.[31] These evaluation methodologies persist despite evidence showing that acquired knowledge and skills do not necessarily equate to behavioral changes or real-world performance.[32] The implication is that organizations primarily prove the worth of training through evidence and enjoyment of learning rather than learning transfer and application in real-world settings.

14.   Defence Research and Development Canada analyzed the CFITES for evaluation methods against Kirkpatrick's model[33] and found the data collection methods

---

[26] Thomas G. Reio, *et al.* "A Critique of Kirkpatrick's Evaluation Model." *New Horizons in Adult Education & Human Resource Development,* 29, no. 2 (Spring 2017): 37. https://onlinelibrary.wiley.com/doi/abs/10.1002/nha3.20178

[27] Blake C.W. Martin, *et al. "The Implementation of Canadian Forces Individual Training and Education System and Mission Essential Competency Training Analysis Methods."* (Defence Research and Development Canada, December 2016): 6. https://cradpdf.drdc-rddc.gc.ca/PDFS/unc265/p805133_A1b.pdf

[28] Reio, *et al.* "A Critique of…, 36.

[29] Perri Estes Kennedy *et al.* "Training professionals' usage and understanding of Kirkpatrick's Level 3 and Level 4 evaluations." *International Journal of Training and Development,* 18, no. 1 (March 2014): 1, https://web-s-ebscohost-com.cfc.idm.oclc.org/ehost/pdfviewer/pdfviewer?vid=2&sid=02ab479b-2507-4607-a99b-86f13b68ed72%40redis

[30] *Ibid.*, 2.

[31] Abbe, "Evaluating Military…

[32] Reio, *et al.* "A Critique…, 36.

[33] Martin, *et al. "The Implementation of Canadian Forces …,* 11.

recommended all correspond to level-one evaluation methods and fail to address the other levels. They also found very few objective evaluative measures recommended for level two evaluation. At level-three and level-four evaluation, the researchers found that it was an intended feature in CFITES volume 11(1) however the application of this evaluation was not evident in any of the superseding Qualification Standard (QS) documentation reviewed,[34] meaning that there is an evident disconnect between theory and practice. The CFITES also lacked evaluation methods for education programs in comparison to training programs;[35] meaning that for courses that bridge cognitive and affective domains, such as those related to Operation HONOUR, the evaluation process is additionally inadequate.

Prioritizing level-three evaluation

15.   Training interventions with outcomes primarily intended to change an individual's behavior must by their very nature be subject to level-three evaluation. Behavioral change is the only indicator that learning transfer has taken place, and noting that individual change is a requirement for successful organizational change and results, then measuring post-training behavior is essential.[36] Kirkpatrick himself stated that level-three evaluation was both the most important yet the most forgotten: it is critical to yield organizational results, yet training delivery professionals will invest their time, energy, and expense into levels one and two (as this is what they control) and the organizational hierarchy will focus on level four. Thus, there is nobody 'owning' level-three evaluation. [37]

16.   The distinction between level-two (learning) training and evaluating, and level-three (behavior) training and evaluating is stark and highlights the criticality of the learning transfer process.[38] Kirkpatrick's research looked at the relationship between the levels and found that there is a good correlation between levels one and two, i.e., positive engagement led to successful learning, and a good correlation between levels three and four, i.e., application of learned behaviors led to increased organizational results.[39] However there was no significant correlation between levels two and three, i.e., excellent training does not guarantee learning transfer.

17.   Learning transfer will only occur if certain conditions are met within the training environment and the real world, and this requires deliberate and constant reinforcement. Training does not finish when the training course ends; acquiring and assessing knowledge at the end of training will rarely be effective in the long term. Research at Columbia University found that only 24% of learning effectiveness is through level-one

---

[34] *Ibid.*

[35] Martin, *et al. "The Implementation of Canadian Forces…* 11.

[36] Kathleen D. Harrell, "Level III training evaluation. Considerations for today's organizations." *Performance Improvement,* 40, no. 5 (May/June 2001): 26, https://doi-org.cfc.idm.oclc.org/10.1002/pfi.4140400508

[37] Kirkpatrick. *Evaluating Training…, 83.*

[38] Bates, "A critical analysis of…, 342.

[39] Jim Kirkpatrick and Wendy Kayser Kirkpatrick. "The Kirkpatrick Four Levels: A Fresh Look After 50 Years 1959-2009." *Open Space Consulting,* last modified [or accessed] 9 January 2022, https://openspaceconsulting.com/wp-content/uploads/2019/06/Kirkpatrick-Four-Levels-wp-updated.pdf

and level-two activities[40] and in 2006 the American Society for Training and Development found that more than 70% of training failure comes after training is completed.[41] Thus level three is more than just evaluating; it is a comprehensive, continuous performance monitoring and improvement system.[42]

Ensuring long-term success

18.   Behavioral change, critical to ensuring long-term success post-training, will occur if four conditions are present: the individual has a desire to change, they know what to do and how to do it, they work in the right climate, and they are rewarded for changing.[43] The first two can generally be achieved from the training; however, the latter two depend on the organization, its leaders, and the immediate hierarchy. Kirkpatrick stated that the climate was the most critical factor: a preventative climate will never lead to behavioral change no matter how effective the other three factors; a neutral climate will require all three other factors to be positive; but, in an encouraging climate, change depends on the first two factors only (reward being immaterial).[44] An encouraging climate can be achieved by a culture of training follow-up and coaching,[45] and positivity by the organization's hierarchy. Learning transfer from the Operation HONOUR courses is unlikely to occur until CAF climate and culture are conducive to it.

Evaluation challenges

19.   Conducting evaluation thoroughly is difficult; level-three evaluation more so as it is the most disruptive to traditional training evaluation practices.[46] There are three main challenges to evaluating behavioral outcomes: measurement metrics, resource constraints, and training structures.[47] First, measurement, where evaluators need to decide reliable metrics, particularly challenging when learning outcomes include attitudinal change, and consider when it is best to conduct evaluation.[48] The answer will often depend on the extent of individual behavior change required (i.e., the degree of learning transfer required compared to training start-state), plus whether the organizational climate is conducive to learning transfer or not. Second, resources. Level-three evaluation requires measuring behavior consistently and systematically, plus having the training, time, and expertise to do so; all difficult to resource. Lastly, training structures. Many military courses (particularly those related to behavioral change within cognitive and affective domains) are decentralized delivery limiting the ability to conduct level-three evaluation

---

[40] *Ibid.*

[41] *Ibid.*

[42] Kirkpatrick. *Four Levels…,* 49.

[43] Kirkpatrick. *Evaluating Training…,* 23. Kirkpatrick suggests that rewards may include an increase in salary, promotion, respect, simplification of work, acceleration of work, etc. However, these rewards are more appropriate for cognitive or skill-based role-related training rather than the attitudinal training discussed in this paper.

[44] *Ibid*, 24.

[45] Dočekala. "Evaluation Levels in…, 3748.

[46] Kirkpatrick, *Four Levels…,* 49.

[47] Abbe, "Evaluating Military…

[48] *Ibid*.

properly and also creating variables that the training design cannot control nor account for.[49]

CONCLUSION

20.   Training evaluation, the systematic process of analyzing training to assess its worth to the organization, is a critical, non-discretionary activity for any organization that delivers training. Evaluation should not be viewed as a one-off event, nor just focused on the KSA acquired as a result of the training. Instead, training is only effective if learning transfer results in individual behavioral change and delivers long-term organizational benefits. Kirkpatrick's model of training evaluation describes this as level-three and level-four evaluation. However, for many organizations, including the CAF, there is either a lack of understanding of this criticality, or there is a disconnect between policy and practice. This results in organizational focus directed towards lower-level evaluation which is misleading and can give false impressions to the overall worth and success of the training.

21.   The CAF training implemented as a result of the Deschamps review, the lack of clarity over terminology, and the identified inconsistencies with CAF training policy, are examples of where organizational uncertainty over training evaluation can occur. The importance of level-three evaluation, particularly for Operation HONOUR courses, but also any training within the cognitive and affective domains, cannot be overstated. These evaluation practices must be prioritized and implemented systematically and coherently by the CAF. Thorough training evaluation must be used to set the conditions to increase the probability of learning transfer taking place. This is the most assured way of ensuring that the training outcomes are achieved; behavioral and organizational change will not occur without it.

RECOMMENDATIONS

22.   CAF leadership should consider the following:

    a.   *Increase awareness.*   Baseline knowledge of training evaluation theory and practice should not be the preserve of training professionals but instead possessed by all CAF personnel. Awareness of training evaluation and the criticality of learning transfer will assist personnel in recognizing it, applying it, and making them more intelligent customers on training events. A training needs analysis should review where this is best placed in the training pathway.

    b.   *Kirkpatrick's model in the CFITES.* The CFITES covers a rudimentary explanation of Kirkpatrick's model, yet it is the underpinning theory upon which the policy is based. The linkages between the model and the recommended evaluation processes and practices are also limited. Greater emphasis on level-three evaluation and its methodology will reinforce to training practitioners the

---

[49] *Ibid.*

importance of behavioral change and how to evaluate it. The CFITES should be reviewed and updated to incorporate more underpinning theory and methods.

c.   *Consider affective domain training.* Military training evaluation is overwhelmingly based upon evaluating skills and specific role requirements, i.e. in the psychomotor and cognitive domains of learning; this is reflected in the CFITES. Yet, increasingly training is seen as the solution to attitudinal problems and fits more broadly into 'education' rather than training. Training policy needs to be reviewed and updated to reflect evaluation for both education and affective domain training.

d.   *Expand timescales.* Training does not end when the training intervention concludes. The CAF needs to take a longer-term view to training success. Behavioral change is slow; measures of success should not be considered or reported until individuals have had time to apply their new behaviors in the workplace or real-world setting.

e.   *Redefine training success.* There needs to be a disassociation between acquired knowledge with a successful training intervention. As long as measures of effectiveness for training events are reported against attendance and pass rates then level-three and level-four evaluation will never be prioritized. The reporting and measures of effectiveness for training events to the CAF leadership should be reviewed.

f.   *Consistent terminology.* There needs to be a review of the terminology used across the CAF regarding training evaluation. Even if understanding is consistent, the interchangeable use of terms such as monitoring, evaluation, and validation can lead to inconsistent practices and confusion concerning reporting. Directions on the use of terminology should be clearly stated across all policies related to training.

BIBLIOGRAPHY

Abbe, Allison. "Evaluating Military Cross-Cultural Training Programs." *Expeditions with MCUP [on line journal],* last modified [or accessed] 8 January 2022, https://doi.org/10.36304/ExpwMCUP.2021.06

Bates, Reid. "A critical analysis of evaluation practice: the Kirkpatrick model and the principle of beneficence." *Evaluation and Program Planning*, 27, no. 3 (August 2004): 341-347, https://doi-org.cfc.idm.oclc.org/10.1016/j.evalprogplan.2004.04.011

Bloom, B.S. (Ed.). Engelhart, M.D., Furst, E.J., Hill, W.H., Krathwohl, D.R. *Taxonomy of Educational Objectives, Handbook I: The Cognitive Domain.* New York: David McKay Co Inc., 1956.

Canada. Department of National Defence. A-P9-050-000/PT-001, *CFITES Manual of Individual Training and Education, Volume 1: Interim Guidance – Introduction/Description.* Ottawa: DND Canada, 2003.

Canada. Department of National Defence. A-P9-050-000/PT-011, *CFITES Manual of Individual Training and Education, Volume 11: Evaluation of Instructional Programmes.* Ottawa: DND Canada, 2003.

Canada. Department of National Defence. A-P9-050-000/PT-Z11, *CFITES Manual of Individual Training and Education, Volume 11(1): Evaluation and Validation Techniques.* Ottawa: DND Canada, 2003.

Canada. Department of National Defence. *CAF Action Plan on Inappropriate Sexual Behaviour.* Ottawa: DND Canada, 2015. https://www.canada.ca/en/department-national-defence/corporate/reports-publications/sexual-misbehaviour/action-plan.html

Deschamps, Marie. C.C. Ad. E. *External Review into Sexual Misconduct and Sexual Harassment in the Canadian Armed Forces.* Ottawa: External Review Authority, 27 March 2015. https://www.canada.ca/en/department-national-defence/corporate/reports-publications/sexual-misbehaviour/external-review-2015.html

Dočekala, Vit and Miroslava Dvořákováb. "Evaluation Levels in Education." *Procedia – Social and Behavioral Sciences,* 174, no. 12 (February 2015): 3743-3749, https://doi-org.cfc.idm.oclc.org/10.1016/j.sbspro.2015.01.1108

Harrell, Kathleen D. "Level III training evaluation. Considerations for today's organizations." *Performance Improvement,* 40, no. 5 (May/June 2001): 26-29, https://doi-org.cfc.idm.oclc.org/10.1002/pfi.4140400508

Kennedy, Perri Estes, Seung Youn Chyung, Donald J. Winiecki and Robert O. Brinkerhoff. "Training professionals' usage and understanding of Kirkpatrick's Level 3 and Level 4 evaluations." *International Journal of Training and Development,* 18, no. 1 (March 2014): 1-21.

Kirkpatrick, Donald and James Kirkpatrick. *Evaluating Training Programmes*. 3rd ed. San Francisco: Berrett-Koehler Publishers, Inc. 2006. https://onlinelibrary.wiley.com/doi/10.1002/pfi.2006.4930450702

Kirkpatrick, James D. and Wendy Kayser Kirkpatrick. *Four Levels of Training Evaluation.* Virginia: ATD Press. 2006. https://books.google.ca/books=onepage&q=kirkpatrick&f=false

Kirkpatrick, Jim and Wendy Kayser Kirkpatrick. "The Kirkpatrick Four Levels: A Fresh Look After 50 Years 1959-2009." *Open Space Consulting,* last modified [or accessed] 9 January 2022, https://openspaceconsulting.com/wp-content/uploads/2019/06/Kirkpatrick-Four-Levels-wp-updated.pdf

Langmann, Sten and Stefan Robert Thomas. "Rethinking training evaluation: omnidirectional evaluation, range of audience and intentional change." *The International Journal of Human Resource Management,* 30, no. 12, (2019): 1902-1930. https://www.tandfonline.com/doi/abs/10.1080/09585192.2017.1303782?journalCode=rijh20

Martin, Blake C.W., Christopher Huffam, James Kerry and Stuart C. Grant. *"The Implementation of Canadian Forces Individual Training and Education System and Mission Essential Competency Training Analysis Methods."* Defence Research and Development Canada, December 2016. https://cradpdf.drdc-rddc.gc.ca/PDFS/unc265/p805133_A1b.pdf

Preusser, Melissa K., Lynn K. Bartels PhD., and Cynthia A. Nordstrom PhD. "Sexual Harassment Training: Person Versus Machine." *Public Personnel Management* 40, no. 1 (Spring, 2011): 47-62. https://www.proquest.com/scholarly-journals/sexual-harassment-training-person-versus-machine/docview/1664817684/se-2?accountid=9867.

Reio, Thomas G., Tonette S. Rocco, Douglas H. Smith and Elegance Chang. "A Critique of Kirkpatrick's Evaluation Model." *New Horizons in Adult Education & Human Resource Development,* 29, no. 2 (Spring 2017): 35-53. https://onlinelibrary.wiley.com/doi/abs/10.1002/nha3.20178

Salas, Eduardo, Laura M. Milham and Clint A. Bowers. "Training Evaluation in the Military: Misconceptions, Opportunities, and Challenges." *Military Psychology,* 15, no. 1 (2003): 3-16 https://www.tandfonline.com/doi/abs/10.1207/S15327876MP1501_01?journalCode=hmlp20

United Kingdom. British Army. *Army Command Standing Order No 3249 Conduct of Army Training.* UK MOD: Crown Copyright, October 2021.

United Kingdom. Ministry of Defence. *JSP 822: Defence Direction and Guidance for Training and Education Part 1: Directive.* V4.1, UK MOD: Crown Copyright, November 202. https://tinyurl.com/2p8pajsa

United Kingdom. Ministry of Defence. *JSP 822: Defence Direction and Guidance for Training and Education Part 2: Guidance.* V4.1, UK MOD: Crown Copyright, November 202. https://tinyurl.com/cfvdms4c