



TEACHING GENDER IN THE MILITARY

A Handbook

Beth Lape, "Assessment and evaluation as tools for improvement"

in PfPC SSRWG and EDWG, *Handbook on Teaching Gender in the Military*
(Geneva: DCAF and PfPC, 2016).



Schweizerische Eidgenossenschaft
Confédération suisse
Confederazione Svizzera
Confederaziun svizra

With the support of the Swiss Government.



DCAF

a centre for security,
development and
the rule of law

Assessment and evaluation as tools for improvement

Beth Lape (United States)¹

CONTENTS

- | | |
|---|--|
| 1. Introduction | 6. Evaluation of gender awareness in military curricula |
| 2. The theory behind evaluation | 7. Annotated bibliography |
| 3. Evaluation in the ADDIE model | Annex: Military exercise scenarios testing application of a gender perspective |
| 4. Formative and summative assessment | |
| 5. Using the Kirkpatrick model to integrate gender equality in military curricula | |

1. Introduction

Evaluation is the systematic determination of the merit, worth and significance of a learning or training process by using criteria against a set of standards. The evaluation phase is ongoing throughout the process used in designing curricula. The primary purpose of evaluation is to ensure that the stated learning outcomes will actually meet a required need, in this case for educational achievement. This chapter discusses evaluating a gender-related curriculum.²

Evaluation is performed during all phases of an instructional design process, and a model can be used when including evaluation in the review of a programme. Various questions are often asked during analysis of the requirements. Is there a performance problem related to a learning outcome? How will implementing a change to the curriculum positively affect a need or goal? What must the learners be able to do to attain the required change in performance?

Evaluation can also be taken one step further by ensuring the learners can actually meet the new performance standards once they have completed the learning process and returned to their jobs; and by ensuring that the needs or learning outcomes are actually being met by measuring performance and effectiveness in regards to a better understanding and application of more gender-sensitive practice in military operations' design and implementation.

In the process of developing an education and training programme that incorporates a gender perspective, it is important to consider how the learner will be assessed and evaluated throughout the programme, and in the end if the programme outcomes were achieved. This is not easy, as it may be necessary to determine if a change in the affective domain (attitudes, behaviour) occurred as well as a change in the cognitive domain (knowledge).

Through this chapter the theory behind evaluation is explored, as well as how it relates to learner assessment in lessons on gender. Descriptions are given of the various models and mechanisms used in developing evaluation schemes.

2. The theory behind evaluation

Evaluation is often considered the culminating feedback of a programme, but ideally the evaluation is part of a cycle with continuous oversight of the process. The evaluation results feed back into the design of the next iteration of the programme. In education, evaluation can be used to link the achieved results with the learning outcomes, in addition to providing a form of quality control for the programme.

Box 8.1 Assessment and evaluation³

Education professionals make distinctions between the terms “assessment” and “evaluation”.

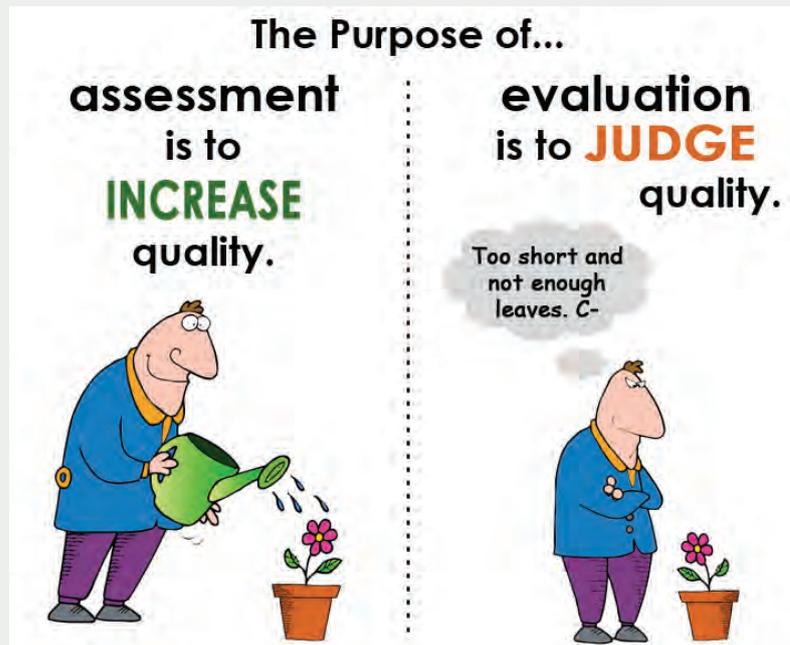
Assessment is the process of documenting knowledge, skills, attitudes and beliefs, usually in measurable terms. The goal of assessment is to make improvements, as opposed to simply being judged. In an educational context, assessment is the process of describing, collecting, recording, scoring and interpreting information about learning.

Evaluation is the process of making judgements based on criteria and evidence.

This is accomplished by making links from learning to organizational activities in addition to considering cost-effectiveness. It is important to determine the relationships between learning, training and the transfer of knowledge and skills to the job. It is also important to remember that there is a risk of evaluative data being manipulated for reasons of internal politics.⁴ For the purposes of this handbook, to help ensure the goal is actually being met it is essential for instructors to consider measures of performance (MOPs – am I doing the right things?) and measures of effectiveness (MOEs – am I doing things right?) to assist learners in striving for the new performance standards once they have completed the educational programme and returned to their jobs.

MOPs and MOEs are terms that have been recently mentioned when discussing a gender programme overall to determine if it is doing what is needed. An MOP can be easily evaluated when a person reviews the quantitative data that may be recorded. This collection could include, for example, information on how many women are in the unit, how many contacts a gender adviser had out in town, and whether extra restroom facilities were added to accommodate all the personnel in the unit. But the MOE is more difficult to evaluate, and it will take longer to determine if what is being done is having any kind of effect on the operations.

Assessments of the learner are a part of the evaluation process, as they help to measure the impact of what the instructor teaches and what the learner learns. By determining the value and effectiveness of the different parts of a curriculum, assessment tools provide useful data for the evaluation.

Box 8.2 The purpose of assessment and of evaluation⁵

A literature review suggested there is widespread underevaluation of educational programmes, and that what is being done is of uneven quality.⁶ However, in today's economy and its focus on value for money, organizations are looking to cut programmes that do not work, so this former attitude towards evaluation is changing towards ensuring that training and education support the units' needs. Evaluation provides a mechanism for decision-makers to judge whether quality standards are being met,⁷ and ensures training and education's continuing ability to produce qualified workers. A follow-up evaluation can be conducted to get a report of how well the training and education prepared learners to perform their jobs after having been on the job for a selected period of time. Evaluation is an ongoing process of assessing learner performance, identifying concerns of the instructors and initiating corrective actions for the entire programme. Data sources for evaluation often include surveys, interviews, tests, course evaluation results, service or product data and observations.

3. Evaluation in the ADDIE model

As mentioned in Chapter 4, the curriculum design process is often based on the ADDIE (analysis, design, development, implementation and evaluation) model. While each phase of the ADDIE process builds on the outputs of the preceding phases, the phases are not sequential. Any change within a phase requires review and possibly adjustments to preceding phases. Note that although evaluation is part of the review cycle, it is also the overall encompassing factor that includes all the phases. Instruction developers may move in and out of each phase. Continuous evaluation of products from each phase, along with approvals, serves to eliminate or reduce wasted effort.⁸

The designer in the *analysis* should focus on the target audience. A diagnostic tool, such as a written or oral pre-test, could be conducted to determine whether the skill level of the curriculum matches the learners' level of skill and intelligence, to make sure that what they already know will not be duplicated and instead the learning is focused on topics and lessons yet to be explored. In this phase, instructors distinguish between what the learners already know and what they have to know after completing the course. Several key components must be used to make sure analysis is thorough. In a "blended learning programme" integrating online and residential courses, the learners could all be brought up to a more common learning level before the programme begins.

Other points to consider in the analysis phase are the demographics of the target audience, to include previous experiences; the learning outcomes in terms of knowledge, skills and attributes; review of existing instructional strategies previously employed and the assessment results; available technology and time for online and residential courses; and overall resources required (both technical and support) in time, human resources, technical skills, finance, etc.

While the instruction is being *designed* there are questions to consider. What must the learners learn to enable them to fulfil the required goal? While the curriculum is being *developed*, the designers should ask about the activities that will best bring about the intended outcomes. During the *implementation* of the curriculum, the developers should ask if the skills and knowledge to perform the required tasks have been acquired.

The *evaluation* phase of ADDIE is developed in more detail in the following sections of this chapter. Every stage of the ADDIE process involves evaluation, and it is an essential component of the construct. Throughout the evaluation phase the curriculum designer will monitor the educational programme to determine if the issue is solved and whether the desired outcomes are met.

4. Formative and summative assessments

Assessments of learners are normally divided into two broad categories: *formative* and *summative*. A *formative assessment* (sometimes referred to as internal) consists of a range of formal and informal assessment procedures that are used *during* the learning process to modify teaching and learning activities to determine the level of learners' cognitive achievement. It is used in judging the status of the programme while activities are in progress, i.e. *forming*. This kind of assessment allows learners and instructors to monitor how well the goals and desired outcomes of the instruction are being met while it is in progress. The main purpose is to determine gaps in the learning so that interventions can take place immediately to allow the learners to continue to master the required skills and knowledge.⁹ The formative assessment is intended to foster development and improvement within an ongoing activity (or person, product, programme, etc.).¹⁰

A *summative assessment* (sometimes referred to as external) focuses on the outcome and determines instructional achievement at *the end* of the programme's activities, i.e. *summation*.¹¹ It is used to assess whether the results of the object being reviewed (programme, intervention, person, etc.) met the stated goals.¹² Traditional assessments at the end of a class are summative, but it is also beneficial to include some that serve as formative assessments during the instruction to make sure the information is getting across to the learners.

The various methods used to collect summative information include questionnaires, surveys, interviews, observations and testing. The methodology used to gather the data should be carefully designed and executed to ensure the data are accurate and valid. Questionnaires are the least expensive procedure for external evaluations and can be used to collect large samples of information from those who have attended the course. The questionnaires should be tested before use to ensure the recipients understand what is being asked in the way the designer intended.

5. Using the Kirkpatrick model to integrate gender equality in military curricula

One of the better-known evaluation methodology models used for judging learning processes in military education and training is Donald Kirkpatrick's four-level evaluation model, first published in his 1954 dissertation titled "Evaluating human relations programs for industrial foremen and supervisors", and then later in a series of articles in 1959 in the *Journal of the American Society of Training Directors*.¹³ However, it was not until his 1994 book, *Evaluating Training Programs*, was published that the four levels became popular. Today the model is a cornerstone in the training and learning industry.¹⁴

The four levels of evaluation consist of reaction, learning, behaviour and results. These levels are explained further in the following subsections.

Level 1: Reaction – How well did the learners like the learning process? Assessments at this level help to report how learners react to the instructional programme. Common steps in a level 1 assessment include:¹⁵

- determine what you want to find out;
- design a form that will quantify reactions;
- encourage written comments and suggestions;
- attain an immediate response from all in attendance;
- seek honest reactions;
- develop acceptable standards;
- measure reactions against the standards and take appropriate action;
- communicate the reactions as appropriate.

Assessments can be conducted to get quick feedback on the instruction using various methods:

- ABCD questions (learners hold up a card with a letter stating which answer is correct);
- open-ended questions by the instructor;
- one-minute paper written at the end of the session by the learners and evaluated by the instructor;
- self-assessments (possible reactions to new gender concepts such as “boys don’t cry” or “a woman’s place is in the home”);
- peer assessments;
- journaling by the learners.

Examples of questions that could be used in evaluating how the subject of gender was received by the audience are noted in Box 8.3.

Box 8.3 Sample evaluation questions for education that incorporates a gender perspective¹⁶

Please rate on a score of 1–5 (1 = poor, 5 = excellent).

1. Value of this topic in relation to my job
2. Usefulness of the course content
3. Presentation methods used
4. Trainer’s ability to transfer knowledge
5. Atmosphere conducive to participation
6. My opinions were taken into consideration
7. Value of the fact sheets
8. Relevance of the work sheets

Please answer the following questions in your own words.

9. Do you have any suggestions about additions to the course?
10. Is there anything you think should be dropped from the course?
11. What did you enjoy most about the course?
12. What did you dislike most about the course?
13. What aspect of the course did you find most useful?
14. What aspect of the course did you find least useful?
15. Was the course (please tick)
 - a) Too long
 - b) Too short
 - c) The right length?
16. Do you have any comments to make about the administrative arrangements for the course (e.g. room, food)?
17. Do you have any other comments to make?

Level 2: Learning – What did they learn (the extent to which the learners gain knowledge and skills)? Steps in these assessment methods include:¹⁷

- evaluate knowledge, skills or attitudes both before and after the programme;
- attain a response rate of 100 per cent;
- use the results of the evaluation to take appropriate action;
- use feedback questions such as the following:
 - What did you find interesting about this course?
 - What does it take to succeed in this course?
 - How did the instructional approaches used in this class affect your efforts to learn?
 - How would you rate the effort you put into this course compared to other courses you have taken?

Common examples of assessment methods at this level include using a written test to measure knowledge and attitudes, and a performance test to measure skills. Other examples are:

- final exams, projects, essays;
- end-of-course feedback.

Some type of a feedback form/questionnaire/survey is the most commonly used method to evaluate instruction at this level. Parts of questionnaire construction are intuitive, but there are certain points with which a course evaluator should be familiar. Concepts being asked about need to be clearly defined and questions unambiguously phrased, or else the resulting data may be seriously misleading. Here are some simple rules to follow.¹⁸

- Each question should relate directly to your evaluation objectives.
- Everyone should be able to answer every question (unless instructed otherwise).
- Each question should be phrased so that all respondents interpret it in the same way.
- The answers to each question should provide the information you need to know, not what would be nice to know.

Box 8.4 can be used when developing a questionnaire. The comments in italics give tips on good questionnaire design.

Level 2 feedback is important to help evaluate the short-term impact of gender-related education on the learners' attitudes towards a gender perspective and awareness on gender-related topics.

Level 3: Behaviour – What changes in job performance resulted from the learning process? Does the learner have the capability to perform the newly learned skills while on the job? How have the learners been able to adopt a change in attitude towards gender within their daily work? Thoughts to consider in the use of this level include the following.²⁰

- Using a control group if possible.
- Allowing enough time for a change in behaviour to take place (such as six months, enough time for the learning to be put into action to cause any change).
- Surveying or interviewing one or more of the following groups: learners, their bosses, their subordinates and others who often observe the learners' behaviour on the job.
- Choosing the appropriate number to sample (how many would be a significant number to sample?).
- Considering the cost of evaluation versus the benefits.

Examples of ways to conduct evaluation at this level could include:

- reaching out to all identified stakeholders;
- creating assessments that reach as many stakeholders as possible;
- sending out questionnaires to previous learners.

Measures of effectiveness regarding any change in behaviour towards gender would start to be evaluated at this point.

Box 8.4 General course evaluation item consideration – Questionnaire¹⁹

Consideration	Example
<p>Statements should be neutral so as not to bias the respondent</p>	<p>Indicate to what extent you agree with the following statement.</p> <p>The course was generally well presented. 1 2 3 4 5</p> <p>1 = Strongly disagree; 2 = Disagree; 3 = Neither agree nor disagree; 4 = Agree; 5= Strongly agree.</p> <p><i>The statement above is biased as it is written in a positive form. The alternate below is neutral.</i></p> <p>Indicate your response to the following statement.</p> <p>The course was:</p> <p>Very helpful Satisfactory Not helpful</p>
<p>Combined positive/negative statements</p>	<p>In cases where it is not possible to write neutral statements, positive statements must be balanced with negative statements.</p> <p>Indicate to what extent you agree with the following statement.</p> <p><i>Positive</i> – The course contained new information. 1 2 3 4 5</p> <p><i>Negative</i> – The activities were not well sequenced. 1 2 3 4 5</p> <p>1 = Strongly disagree; 2 = Disagree; 3 = Neither agree nor disagree; 4 = Agree; 5 = Strongly agree.</p> <p><i>Using an equal number of positive and negative statements balances the form and eliminates bias.</i></p>
<p>Statements should only ask respondents to evaluate one thing.</p>	<p>The course contained information, ideas, methods and techniques new to me. 1 2 3 4 5</p> <p>1 = Strongly disagree; 2 = Disagree; 3 = Neither agree nor disagree; 4 = Agree; 5 = Strongly agree.</p> <p><i>How would learners respond if they already knew the basic information but the techniques were new? Items must be singled out.</i></p> <p>Indicate to what extent you agree. The course contained the following that was new to me:</p> <p>a. Information 1 2 3 4 5</p> <p>b. Ideas 1 2 3 4 5</p> <p>c. Methods 1 2 3 4 5</p> <p>d. Techniques 1 2 3 4 5</p>

Level 4: Results – This level is the most difficult to assess, as it requires continuous follow-up with the learners and a measurement of the results of any actions they performed. In the area of gender, the results could be measured by a change in a person's affective domain in regards to how they value and include gender in any planning processes. Additionally, success at this level could be measured by how any attitudes within an area changed, or even using metrics such as improved health due to a change in the healthcare programme for a village. Overall, you need to evaluate the tangible results of the learning process in terms of reduced cost, improved quality, increased production, efficiency, etc. Points to consider include the following.²¹

- Allow enough time for results to be achieved.
- Repeat the measurement at appropriate intervals.
- Consider the cost of the evaluation versus benefits.
- Be satisfied with the available evidence if absolute evidence is not possible to attain.

Kirkpatrick's concept is quite important, as it makes an excellent planning, evaluating and troubleshooting tool, and can be applied to gender-related education.

6. Evaluation of gender awareness in military curricula

The effectiveness of learning about gender within the regular curriculum is evaluated using the standard assessment methods previously discussed. A more difficult measure is reviewing the behaviour of the learners. To achieve results an accurate assessment should be made over time, preferably several months. An accurate assessment is one that looks at integration based upon criteria determined in evaluations of the individual as well as the organization. Reviewing this kind of material is an example of the information that would be used when conducting a Kirkpatrick Level 3 review of how the knowledge transfer of the material changed the affective domain. The Committee of Women in NATO Forces has published basic metrics and drafted NATO guidelines to integrate gender-related topics into the NATO evaluation process (Box 8.5).

Box 8.5 Recommendations related to the evaluation process by NATO nations/authorities²²

The principle can only be successfully achieved if NATO nations and authorities implement, as a minimum, the following recommendations.

- Collect and analyse gender-related data, disaggregated by sex and other categories as appropriate, relevant to NATO operations.
- Monitor, review and analyse the effectiveness of gender and cultural awareness training, based on lessons learned from previous NATO operations.
- Report on identified advantages and disadvantages in respect of gender-related issues by:
 - identifying situations where having female personnel provided a specific advantage in NATO operations;
 - conducting surveys of military personnel regarding their perspectives on gender-related issues during NATO operations;
 - identifying any unintended impact on gender-related issues affecting civilian populations and military personnel.
- Evaluate the impact of actions not only on the population as a whole, but also on men and women separately.
- Integrate gender-related topics into the existing reporting system (e.g. lessons identified, lessons learned, best practices).
- Share and exchange information on gender-related best practices with other international organizations such as the United Nations, the Organization for Security and Co-operation in Europe, the European Union, the International Committee of the Red Cross and non-governmental organizations.
- Ensure that gender-related issues covered in the lessons learned are broadly shared so that the required changes can be implemented.

There are various methods for measuring the transformation of gender-aware attitudes into gender-responsive actions and behaviours – examples include the following.

- Anonymous surveys of men and women regarding sexual harassment.
- Exit surveys collecting information on people's opinions of the organization upon leaving employment, to include questions about job training and work environment.
- Evaluation forms completed by learners several months after the training programme, asking questions on relevant attitudes, perceptions and behaviours.
- Interview supervisors/managers to get their assessment of any change in behaviour of the learner when a reasonable amount of time has elapsed after return to the organization.²³

A transformative change in a person's behaviour can also be assessed through observations during military training exercises. The exercises presented in the annex to this chapter were developed by a military gender adviser conducting training on a gender perspective in the Democratic Republic of the Congo. All the information obtained at the various levels of these exercises becomes part of the "analysis" phase of the ADDIE model to review what has occurred in the field to see if there has been a change in the affective domain based upon the instruction given. Gender education and training should be transformative at all levels. Full evaluation to see if changes have occurred involves using a range of methods, from surveys that just tick a box to including observations, peer reviews and other forms of evaluation.

The availability of tools, resources and leadership support will have an effect on how easily learners will be able to apply what they have learned on gender in operations. Trained gender focal points should be available to assist with questions in the field, as well as the gender adviser on the staff. But it is also beneficial for learners to keep in contact with their gender instructor as a person to reach back to with questions on gender-related topics that have been encountered, but were not quite covered the same way in the programme. Assessments of learners and evaluation of programmes are not easy, but are important steps in ensuring that the concept of applying a gender perspective is understood and can be used in operational situations.

7. Annotated bibliography

Clark, Don, *Instructional System Design: The ADDIE Mode – A Handbook for Learning Designers*, www.nwlink.com/~donclark/hrd/sat.html.

This online handbook contains a wealth of information on the ADDIE model and a chapter specifically on evaluation, which discusses formative and summative evaluations as well as Kirkpatrick's four levels. It also contains links to other useful electronic and paper resources.

Committee on Women in NATO Forces, "Guidance for NATO gender mainstreaming", 2007, www.nato.int/issues/women_nato/cwinf_guidance.pdf.

This document produced by the predecessor to the NATO Committee on Gender Perspectives includes a section on recommendations for the integration of a gender perspective in the evaluation process. It is specifically aimed at NATO nations and authorities.

Department of National Defence and Canadian Forces, "Evaluation of instructional programmes", in *Manual of Individual Training and Education*, Vol. 11, A-P9-050-000/PT-011 (Ottawa, ON: Department of National Defence, 2003).

This volume of the *Manual of Individual Training and Education* provides guidance for the evaluation of instructional programmes given within the Canadian Forces. It provides information on evaluation models, followed by recommendations on how to create and implement an evaluation plan and then interpret its findings so they can be fed into the evaluation report.

Kirkpatrick, Donald L. and James D. Kirkpatrick, *Evaluating Training Programs: The Four Levels*, 3rd edn (San Francisco, CA: Berrett-Koehler Publishers, 2006).

This seminal book provides instructors with practical guidance on evaluation. It covers the four levels of reaction, learning, behaviour and results. Some useful resources are also available on the “Our philosophy” section of the Kirkpatrick website, www.kirkpatrickpartners.com/OurPhilosophy/tabid/66/Default.aspx.

Tönisson Kleppe, Toiko, “Gender training for security sector personnel – Good practices and lessons learned”, in Megan Bastick and Kristin Valasek (eds), *Gender and Security Sector Reform Toolkit* (Geneva: DCAF, OSCE/ODIHR, UN-INSTRAW, 2008), www.dcaf.ch/Publications/Gender-Security-Sector-Reform-Toolkit.

This publication is the twelfth tool in DCAF’s *Gender and Security Sector Reform Toolkit*. Using Kirkpatrick’s model, it provides concrete guidance for practitioners tasked with evaluating gender-related education and training courses.

Annex: Military exercise scenarios testing application of a gender perspective²⁴

When training soldiers, one method of assessing and evaluating how well lessons on gender have been assimilated is by seeing if they can apply the lesson in a practical scenario. The officer responsible for designing the exercises may benefit from liaising with a gender adviser who could assist in developing the integration of a gender perspective into the scenarios. Military exercises usually test officers through operational staff work, which includes writing orders, and tactical-level exercises which put soldiers through practical tests. The principles of UN Security Council Resolution (UNSCR) 1325²⁵ are a good foundation for planning and developing exercises which seek to implement a gender perspective, for several reasons.

- a. UNSCR 1325 seeks to promote the *participation* of women within the military and within the indigenous population where the military is operating.
- b. UNSCR 1325 encourages militaries operating overseas and at home to *protect* women, men, girls and boys from conflict-related sexual violence (CRSV).
- c. UNSCR 1325 encourages militaries operating overseas and at home to *prevent* acts of CRSV against women, men, girls and boys.
- d. UNSCR 1325 encourages militaries operating overseas and at home to assist and support women and girls in the post-conflict *reconstruction* phase. This can include ensuring soldiers know how to *respond* to survivors of CRSV.

The situations below complement UNSCR 1325 and can be woven into tactical military exercises which require the implementation of a gender perspective by the soldiers.

PARTICIPATION

Key leader engagement

Senior non-commissioned officers and commissioned officers deployed on operations overseas are frequently required to engage with the civilian population. Experience from NATO missions in Afghanistan and UN peacekeeping operations across Africa has revealed that male-dominated military and peacekeeping units are less likely to talk to women (50 per cent of the population), thus hearing only “half of the story”. Soldiers should therefore be “tested” when tasked to initiate key leader engagement to see if they consider holding meetings with civilian women. At the tactical level soldiers should be tested in running a meeting with an all-women group or with a civilian woman who has influence in the community. The *participation* of servicewomen could also be tested, as in some societies only a woman can approach another woman, and men cannot approach women to whom they are not related.

Intelligence collection

As women often have their own “intelligence networks”, it is vital that military forces operating in an area speak to women as well as men. In twenty-first-century warfare civilian women are more likely to be targeted by illegally armed groups than NATO or UN soldiers. These women will have a better understanding of when and where they are vulnerable, and so should be included when gathering information. Staff officers in HQs and soldiers on tactical exercises can be tested to see if they include this in their planning and when they go on patrol and interact with the civilian population in the exercise setting.

PARTICIPATION, RECONSTRUCTION

Civil-military quick-impact project design

Officers at the operational level work with soldiers deployed on the ground to identify projects that can “win hearts and minds”, or at least tolerance from the civilian community among whom they are operating. Military units consistently seek out ideas for projects only from influential males. In exercises at both operational and tactical levels the soldiers can be assessed on whether they think to include women and speak to women to collect ideas for projects. Both levels could be tested to see if they can suggest projects that would target the female population.

Negotiation talks

NATO and UN staff may be required to support the political process by overseeing negotiation talks. This is often conducted in a role-play scenario. Both operational and tactical-level exercises can be tested to see if they seek to include women at the negotiating table and have servicewomen on hand for searching, etc., before the talks commence.

PROTECTION, PREVENTION

At the operational level staff being exercised can be tested to see if they will order units to conduct patrols in areas where they have been told the civilian population is most at risk. Assessment can reveal if the planners do not just target armed groups but also plan patrols that are near vulnerable areas for civilians – e.g. on roads into markets or around collecting points for water and wood.

Soldiers being exercised with role plays, etc., can be approached by a woman asking them for help. How they respond and who they inform will indicate if they have understood the gender lessons.

Vehicle check points, house clearing and cordon operations can all be given a gender “flavour” by introducing tasks and tests that make soldiers think beyond their role to defeat an enemy. For example, the soldiers stop a truck full of women and girls – do they let the vehicle pass or are they concerned it is part of a trafficking ring? Do they call the civilian police or let the vehicle through, have they got female interpreters who can talk to the women, etc.?

During a clearing operation the soldiers in a house come across a room full of young boys. Do the soldiers know how to respond and who to call? Are these boys being trained as child soldiers or used for suicide bombing? Do the soldiers know what to do with children?

Notes

1. The author would like to thank Nathaly Lévesque and Jana Naujoks for providing written comments on this chapter.
2. US Department of the Army, "Army learning policy and systems", TRADOC Regulation 350-70, 2011, pp. 82–83, www.tradoc.army.mil/tpubs/regs/TR350-70.pdf.
3. ITS Training Services, "What do we mean by testing, assessment, and evaluation?", 4 November 2014, <http://tutorials.istudy.psu.edu/testing/testing2.html>.
4. P. Bramley and A. C. Newby, cited in Don Clark, "Types of evaluation in instructional design", Big and Little Dog's Performance Juxtaposition, 13 July 1995 (last updated 5 October 2015), www.nwlink.com/~donclark/hrd/isd/types_of_evaluations.html.
5. Image courtesy of Pacific Crest, www.pcrest.com.
6. Marguerite Foxon, "Evaluation of training and development programs: A review of the literature", *Australian Journal of Educational Technology*, Vol. 5, No. 2 (1989), pp. 89–104, www.ascilite.org.au/ajet/ajet5/foxon.html.
7. US Department of the Army, note 2 above, p. 83.
8. Ibid.
9. Don Clark, "Types of evaluation in instructional design", Big and Little Dog's Performance Juxtaposition, 13 July 1995 (last updated 5 October 2015), www.nwlink.com/~donclark/hrd/isd/types_of_evaluations.html.
10. M. Scriven, cited in Clark, *ibid*.
11. Ibid.
12. Ibid.
13. Kirkpatrick Partners, "Donald L. Kirkpatrick, PhD: March 15, 1924–May 9, 2014", 9 May 2014, www.kirkpatrickpartners.com/AboutUs/DonKirkpatrick/tabid/223/Default.aspx.
14. Kirkpatrick Partners, "The Kirkpatrick model", 10 January 2010, www.kirkpatrickpartners.com/OurPhilosophy/TheKirkpatrickModel/tabid/302/Default.aspx.
15. ASTD, "Implementing Kirkpatrick's four levels", handout, 5 June 2007, <http://astd2007.astd.org/pdfs/handouts%20for%20web/handouts%20secured%20for%20web%205-15%20thru%205-16/tu101.pdf>.
16. Toiko Tönisson Kleppe, "Gender training for security sector personnel – Good practices and lessons learned", in Megan Bastick and Kristin Valasek (eds), *Gender and Security Sector Reform Toolkit* (Geneva: DCAF, OSCE/ODIHR, UN-INSTRAW, 2008), p. 14.
17. ASTD, note 15 above.
18. "Course evaluation design", SmartEvals.com, 1 October 2002, <http://service.onlinecourseevaluations.com/faq/HowToDesign.aspx>.
19. Chief of Defence Staff, "Evaluation of instructional programmes", in *Manual of Individual Training and Education*, Vol. 11, A-P9-050-000/PT-011, p. 14 (Ottawa, ON: Canadian Department of National Defence, 2003).
20. ASTD, note 15 above.
21. Ibid.
22. Committee on Women in NATO Forces, "Guidance for NATO gender mainstreaming", www.nato.int/issues/women_nato/cwinf_guidance.pdf.
23. Tönisson Kleppe, note 16 above.
24. Major Rachel Grimes, British Army, former UN Military child protection and gender field adviser, Democratic Republic of the Congo.
25. UN Security Council, "Resolution 1325 (2000) on women, peace and security", UN Doc. No. S/RES/1325 (2000), 31 October 2000.