

Evaluation methodology

16. The evaluation constraints and procedures outlined below were discussed with the PLP Evaluation Steering Committee and with a cross-section of UNHCR managers, and other persons of concern who were invited to a meeting to comment on the evaluation methodology. In addition, the evaluation team reviewed *Impact Evaluation of the Protection Learning Programme: Methodological Considerations*, prepared by Hugh Massey of UNHCR in September 2004, and considered his recommendations.

Guiding principles of the evaluation

17. The PLP evaluation focused on identifying the impact of the PLP programme, assessing its current state and suggesting how it could be improved. The evaluation was based primarily on the Context, Inputs, Processes, and Products (CIPP) Model developed by Daniel Stufflebeam. The model is based on the following operational definition of evaluation: 'a process of delineating, obtaining, reporting, and applying descriptive and judgmental information about some object's merit, worth, probity, and significance in order to guide decision making, support accountability, disseminate effective practices, and increase understanding of the involved phenomena.'³

18. Below are the core concepts of the CIPP Model:

- 'Context evaluations assess needs, problems, assets, and opportunities to help decision makers define goals and priorities and help the broader group of users judge goals, priorities, and outcomes.
- Input evaluations assess alternative approaches, competing action plans, staffing plans and budgets for their feasibility and potential cost-effectiveness to meet targeted needs and achieve goals.
- Process evaluations assess the implementation of plans to help staff carry out activities and later help the broad group of users judge programme performance and interpret outcomes.
- Product evaluations identify and assess outcomes – intended and unintended, short term and long term.'⁴

19. Thus, the evaluation team aimed to assess needs, identify problems and outline alternatives that decision makers could employ. The team also sought to suggest alternatives that were feasible and could address intended and unintended outcomes of the evaluation.

³ D. L. Stufflebeam. (2003). *The CIPP Model for Evaluation*, presented at the 2003 Annual Conference of the Oregon Program Evaluators' Network (OPEN). Portland, Oregon. p. 10.

⁴ *Ibid.*, p. 3.

20. The CIPP Model adheres to the standards set by the American Joint Committee on Standards for Educational Evaluation. Thus, the model meets the following conditions:

- 'Utility – serving the information needs of the intended users
- Feasibility – keeping evaluation operations realistic, prudent, viable, and frugal
- Propriety – conducting evaluations legally, ethically, and with due regard for the welfare of participants and those affected by the results
- Accuracy – revealing and conveying technically sound information about the features that determine the evaluand's merit, worth, probity, and/or significance'⁵

21. The PLP evaluation followed the above parameters as well as the principles of transparency, independence, consultation and relevance established by UNHCR's Evaluation and Policy Analysis Unit.

22. The evaluation involved UNHCR stakeholders throughout the process. In order to confirm that the evaluation was proceeding in the right direction and that appropriate evaluation questions were being asked, the evaluation team kept the stakeholders informed and asked for their input as the evaluation proceeded. For example, various members of the PLP Steering Committee were asked to review and pilot the questionnaires prior to their distribution to the intended audiences.

23. The CIPP Model advocates the use of a wide range of qualitative and quantitative methods in order to triangulate the evaluation outcomes. Therefore, the evaluation team did not rely on just one data source, but gathered data from several sources and by different means including but not limited to the following:

- Individual interviews were conducted with members of the PLP Steering Committee, PLP graduates, non-completers or in-progress learners, supervisors, peers, trainers/facilitators, UNHCR persons of influence, UNHCR staff who have not enrolled in the PLP and personnel who work for UNHCR partners.
- Six different online and Microsoft Word-based questionnaires were developed and distributed to six of the groups identified above.
- Focus groups were conducted with PLP workshop attendees.
- The self-study material was compared to distance learning best practices. The results of this comparison led to suggestions that could be used to improve the material.
- One PLP workshop was observed.
- Existing evaluation forms were reviewed.

⁵ Joint Committee on Standards for Educational Evaluation. (1994). *The Program Evaluation Standards* (2nd ed.). Thousand Oaks, CA: Sage. p. 10.

24. A large amount of information was gathered from different people and sources. Thus it was possible to examine ambiguities, crosscheck findings and discern trends from different perspectives.

25. The evaluators were determined not to fall into the trap that captures many evaluators – the ‘popularity contest’. They would present their recommendations based on their best judgment. They would be guided by all the information received, but would not make a recommendation just because it was favoured by most respondents. After critical examination of the issues and options, the evaluators would recommend only those measures most likely to improve the content and process of the PLP.

26. All of the above concepts were applied to the evaluation of the PLP. Members of the Steering Committee and other members of UNHCR acted as a sounding board and provided guidance as the evaluation proceeded. They were actively involved in the identification of the evaluation questions and the methodology to be employed.

Evaluation constraints

27. All evaluations are conducted under constraints, usually those associated with time, resources and the availability of the participants. This evaluation was conducted under constraints that were discussed at length with the PLP Evaluation Steering Committee. Despite these constraints, this evaluation has credibility because all PLP graduates, in-progress participants, and in-completers were invited to express their views and a 33.8% response rate was attained.

28. During the planning of the evaluation, a number of constraints were taken into account. These are outlined below.

29. **Time limitations.** As DIP’s Protection Capacity Section had established a 2-month timeframe for gathering data, the evaluation team had to design triangulation procedures that could be conducted within the specified time and involve participants whose time was limited due to pressures of work and humanitarian emergencies.

30. **Randomness.** The analysis of statistical or quantitative data is based on randomness – that each participant has an equal chance of being selected. In the PLP evaluation plan, every participant who had completed the PLP received a questionnaire, as did all non-completers and in-progress participants of the programme. However, this procedure could not be followed with all groups that were investigated, such as peers, supervisors and refugees, as these individuals were not all identifiable and it was difficult to associate them with specific PLP graduates. Therefore, the evaluation team received permission from the PLP Evaluation Steering Committee to sample some groups based on the availability of their members.

31. **Numbers.** Ideally, in order to ensure valid and reliable results, information from a large number of individuals is required. However, it was not practical within the evaluation plan timeframe to expect input from everyone who could offer views. Thus, the evaluation team estimated the number of responses expected from various groups. This response rate was approved by the PLP Evaluation Steering Committee and is outlined in the next section of this report.

32. ***Geographic Locations.*** UNHCR has refugee operations involving over 5,000 staff members in 120 countries. However, only individuals located in Bosnia, Hungary, Kenya and Switzerland participated in face-to-face interviews. Individuals from Europe, Africa and the Americas were involved in the focus group sessions. The evaluation team also used teleconferencing to reach selected individuals in Bosnia and Herzegovina, Colombia, Hungary, Pakistan, Sierra Leone, Sri Lanka, Thailand and Turkey. Attempts were made to reach individuals in Ecuador, Egypt, Jordan and Congo by telephone prior to the completion of the data-gathering stage.

33. ***Assessing Refugees.*** The thrust of the PLP is to improve the knowledge, skills and attitudes of those who work to protect refugees. Ideally, the evaluators would need to ask refugees whether they had noticed differences between the knowledge, skills and attitudes of the protection officers who took the PLP and those who did not. Further, for the results to be considered reliable, a large number of refugees would have to participate in the study. The large size of the sample would take into account that some refugees might want something and provide a positive review, while others might consider the PLP evaluation as an opportunity to strongly criticise the assistance they received, even though their complaints might have nothing to do with the protection officer.

34. Although the evaluation team recognised the importance of refugee input, it was unable to ascertain how to obtain it effectively within the evaluation plan timeframe. For example, the team did not have access to any records that could show which PLP graduates worked with which refugees; therefore, the evaluation team could not ask specific refugees for their input. The evaluation team asked others within the UNHCR to suggest how information could be obtained from refugees, but did not receive any feasible suggestions that could be implemented within the evaluation timeframe. The UNHCR has numerous assistance and protection indicators, but few suggestions as to how to measure the more qualitative indicators. Although it is a challenging task, the UNHCR could develop mechanisms that it could use to measure its direct impact upon refugees. Refugees, the recipients of UNHCR services, should have an opportunity to express their opinions about the services they receive.

35. It was proposed by members of the PLP Steering Committee that the PLP evaluation team should observe how PLP graduates and non-PLP participants work with refugee committee representatives and refugee leaders. It was also suggested that the team interview women and children in two refugee camps. Due to the disruption of air service to the refugee camps, the evaluators could not visit Kakuma and the visit to Dadaab was restricted.

36. ***Pre- and Post-assessment.*** One of the most common methods for assessing changes in people's knowledge, skills and attitudes is to assess them at the beginning of the training and compare this assessment with their behaviour after training has been completed. Due to the short timeline, this could not be done. Also, the evaluators would need access to personal records, which would raise concerns about confidentiality. However, if the UNHCR is serious about evaluating the contribution training makes to its humanitarian efforts, it should assess individuals before and after training and formally follow their actions in the field after training has been completed. It could use the information contained in the Performance Appraisal Reports (PARs) database. However, many UNHCR personnel indicated during

interviews that they did not feel that their supervisors had maintained this database or included useful information other than the fact that individuals had attended particular training sessions.

37. With the above constraints in mind, the evaluation team was able to devise evaluative instruments and procedures that were acceptable to the PLP Evaluation Steering Committee. Everyone agreed that it would be possible to evaluate the impact of the PLP on individual staff, but given the time, resource and access limitations of the study, it would be more difficult to ascertain the impact on refugees. One would need to infer that if the PLP graduates acquired new knowledge, skills and attitudes or enhanced existing ones, they would use them as they worked with refugees. By asking supervisors and peers if they observed changes in those who took the PLP, the evaluation team would obtain an indication as to whether graduates were retaining and applying their new or enhanced skills. Thus, if responses from supervisors and peers yielded positive comments about the effects of the PLP, the Protection Capacity Section could have some confidence that the PLP had an impact upon the lives of refugees.

Data gathering instruments and procedures

38. The evaluation procedures that were followed are briefly described below. They provided a balance between breadth and depth and were selected to provide a valid and reliable assessment of the impact of the PLP on the participants and an indication as to whether the PLP had an impact on the effective protection of refugees.

39. ***Review of feedback from participants*** who completed self-study units. This activity helped the evaluators design the interview and questionnaire questions.

40. ***Interviews.*** The purpose of these interviews was to yield an in-depth assessment of the PLP and its impact as well as to ascertain how the programme could be improved. The following groups were interviewed:

- PLP Evaluation Steering Committee
- UNHCR persons of influence
- PLP graduates*
- Non-completers or in-progress PLP participants*
- Trainers/facilitators*
- Supervisors and peers
- Other persons of interest to UNHCR
- Non-involved (those who have not taken the PLP)*

41. A series of questions was developed for each group above that is marked with an asterisk (*) and used as a guide during the semi-structured interviews. These questions are provided in Appendix A.

42. In order to elicit their true feelings about their involvement with the PLP, interviewees were assured confidentiality – their comments would not be linked to their names. Note that every attempt was made to include staff members located in French- and Spanish-speaking posts. In addition, those in ‘hot’ zones were targeted. Initially, face-to-face interviews were to be conducted in Bosnia, Kenya, Pakistan and Switzerland. However, the Protection Capacity Section limited travel and the evaluators were not able to travel to Bosnia and Pakistan to conduct interviews. Interviews were conducted by telephone. The face-to-face interviews were usually 60–90 minutes in length, while the telephone interviews tended to be 30–60 minutes in length.

43. **Questionnaires.** Online questionnaires using Web-based Zoomerang software, as well as Microsoft Word-based versions, were distributed to the groups of individuals below.

- PLP graduates
- Non-completers or in-progress PLP participants
- PLP workshop attendees
- PLP trainers/facilitators
- Supervisors and peers
- Other persons of interest to UNHCR

44. Six different questionnaires were developed and distributed – one for each group listed above. These questionnaires are provided in Appendix B. Via e-mail, individuals were requested to complete either the online or Word-based questionnaires. All graduates and non-completers who still worked for UNHCR and had an e-mail address were contacted and invited to participate in the study. Note that more than 50% of the respondents submitted items in a Word format even though many had access to the Internet, thus the evaluators had to enter these Word responses into Zoomerang. Zoomerang allows individuals to submit their comments anonymously.

45. ***Review of self-study material relative to distance learning best practices.*** The criteria used for this activity were based on Criteria for Evaluating the Quality of Online Courses.⁶ These guidelines are used at universities such as Central Queensland University in Australia to assist lecturers in evaluating the effectiveness of their Blackboard courses. They are also used in institutions such as the University of Calgary in Canada to teach educators how to design distance education courses. A copy can be found in Appendix C.

46. ***Review of existing workshop evaluation forms.*** This activity helped the evaluators to decide which questions they would use during the focus group sessions held in Budapest.

⁶ C. R. Wright. (October, 2003). *Criteria for Evaluating the Quality of Online Courses*. Edmonton, Alberta: Alberta Distance and Training Association.

47. *Focus group sessions* with participants in the PLP workshop in Budapest, Hungary.

48. Observation of one PLP workshop and distribution of detailed questionnaires to the participants. The questionnaire normally used by the PLP team was rather basic, thus a more detailed form was developed that would help the evaluation team answer the study questions. As the evaluators were able to attend only one PLP workshop, their observations may not be generalised to all PLP workshops.

49. *Review of field study projects* in order to access their impact.

50. *Examination of the PLP relative to other UNHCR learning programmes.* This activity looked at coordination between programmes, possible linkages and similarities among the programmes. Available literature regarding these programmes was reviewed and meetings were held with those who offered several of these programmes such as the Middle-Management Learning Programme, now referred to as the Management Learning Programme.

51. The table below lists the number of individuals who were to be targeted by the evaluators. Note that the total population of PLP completers and non-completers or in-progress participants received questionnaires. As the Department of International Protection was responsible for arranging the interviews, the evaluation team requested that the individuals selected by the department represent a wide variety of staff. The selection criteria took into consideration gender, academic and professional background, geographic coverage and language.

Targeted audience for PLP evaluation			
Category	Questionnaires	Interview	Comments
Steering Committee	-	5	Based on availability prior to Oct. 11
PLP Graduates	320	32	From as many regions as is possible
Non-completers/In-Progress	161	10	Selected based on availability
Supervisors and Peers	60	10	Selected from various regions
Other Persons of Interest	60	10	Selected from HQ, Kenya & Hungary
PLP Trainers/Facilitators	10	5	Selected based on availability
Workshop Attendees	30	-	Focus groups instead of interviews
Non-involved	-	10	Random informal interviews
Total	641	82	