



经济及社会理事会

Distr.: General
25 March 2013
Chinese
Original: English

方案和协调委员会

第五十三届会议

实质性会议

2013 年 6 月 3 日至 28 日

临时议程* 项目 3(b)

方案问题：评价

审查联合国难民事务高级专员公署的评价能力

内部监督事务厅的报告

摘要

内部监督事务厅(监督厅)根据 2008 年开展的一次风险评估确定对联合国难民事务高级专员公署(难民署)进行评价。方案和协调委员会选择在其 2013 年第五十三届会议上审议关于难民署的方案评价(见 A/66/16, 第 66 段)。大会在其第 66/8 号决议中认可了这一选择(见第 6 段)。

在与难民署高级管理层进行协商和讨论后,并根据审计委员会和外部审查最近得出的结论,各方一致认为,在对方案进行评价之前先审查难民署的评价能力将更为及时和有益。

监督厅的审查目的是要评估难民署当前的评价能力并确定如何改进这一能力,以便更好地为组织成果、学习和问责制作出贡献。监督厅采用了定量和定性方法,包括审查文件、访谈、问卷调查、实地考察和对难民署的评价报告进行元评价。

难民署的评价职能未得到明确界定。虽然现行评价政策采用联合国评价小组认可的有关评价的基本定义,但却未充分阐明这一职能的明确和独特的作用,也

* E/AC.51/2013/1。



未把这一职能与难民署的法定任务、目标和战略优先事项明确地联系在一起。对 2010 年和 2011 年期间的 28 份报告进行评估后发现,这些登载在政策制订和评价处网站上被称为难民署评价报告的报告实际上是一个评价报告、政策文件、激发思考的文章和学术研究文章的混杂体。此外在难民署的整体预算中,专用于评价职能的资源一直较少。

不过,目前负责执行难民署评价任务的政策制订和评价处在其既定任务和资源的范围内仍然起到了有益的作用。该处在只有 5 名工作人员和咨询预算有限的情况下在开展评价工作方面维持了高生产率。总体而言,难民署各利益攸关方就该处提供的政策事项信息所作的反馈一直是正面的。

现行评价职能也未完全符合联合国评价小组关于评价工作独立性、可信性和有用性的准则和标准。将该职能置于一个同时负责政策制订和研究的单位内的做法限制了独立性,并且评价工作具有向高级专员和副高级专员报告的双重报告线,虽然这两人都在执行办公室。工作规划流程在很大程度上是由需求驱动的,并不是基于对组织风险的评估结果,而高级专员方案执行委员会不审查或讨论评价报告,虽然参与问卷调查的许多成员表示,他们愿意这样做。

此外,难民署评价职能的可信性因在评估组织成果的方法方面有限制而有局限性。利益攸关方报告说,对难民署在其法定任务所有方面(包括其保护和方案工作)的业绩和影响的评价不够充分,可用于战略决策的评价证据很少。同样,除了政策制订和评价处处长外,未规定负责评价工作的工作人员要具备评价经验或技术能力。缺乏用于讨论评价结果和追踪评价建议的常规和系统性的后续流程,从而进一步限制了难民署评价工作的有用性。

与联合国其他具有强大实地存在的实体不同,难民署没有一个功能性的中央及分散评价结构,关于难民署国家方案业绩的评价证据也很有限。在联合国其他具有强大实地存在的实体中,中央评价工作负责处理全组织范围的战略问题,通常是在独立于负责方案业务的那些人的情况下实施的,并要对结果进行评价。而分散评价则范围较窄,往往嵌入在方案业务之内。难民署可以更有战略性地利用方案支助和管理司和国际保护事务司的现有资源和活动,以使其分散评价工作系统化和规范化,而难民署在某种程度上已经开始这样做了。

监督厅提出了下列五项得到难民署赞同或部分赞同的重要建议:

- 在总部设立一个专职和单列的评价单位,负责进行难民署的方案和交叉战略评价工作
- 修订难民署的评价政策
- 加强难民署方案评价工作的严格性和有用性
- 为评价建议的后续行动拟订一个常规和系统性流程
- 为加强在外地的分散评价拟订一项战略

目录

	页次
一. 导言	4
二. 范围和方法	4
三. 背景	5
四. 联合国系统的评价工作	6
五. 审查结果	7
A. 难民署未明确界定评价职能	7
B. 难民署当前的评价能力：未完全符合联合国评价小组关于独立性、 可信性和有用性的关键准则和标准	8
C. 缺乏联合国其他具有强大实地存在的实体通常都有的一个功能性的 中央及分散的评价系统	13
六. 结论	15
七. 建议	16
附件	
2013 年 3 月 20 日联合国难民事务高级专员公署副高级专员给内部监督事务厅 检查和评价司的备忘录	19

一. 引言

1. 内部监督事务厅(监督厅)根据 2008 年开展的一次风险评估确定对联合国难民事务高级专员公署(难民署)进行评价。方案和协调委员会选择在其 2013 年 6 月第五十三届会议上审议关于难民署的方案评价(见 A/66/16, 第 66 段)。大会在其第 66/8 号决议中认可了这一选择(见第 6 段)。
2. 在与难民署高级管理层进行协商和讨论后,并根据审计委员会和外部审查最近得出的结论,各方一致认为,在这个时候,在对方案进行评价之前先审查难民署的评价能力将更为及时和有益。
3. 这次审查的目的是:
 - (a) 评估难民署当前的评价能力;以及
 - (b) 确定如何改进这一能力,以便更好地为组织成果、学习和问责制作出贡献。

二. 范围和方法

4. 监督厅从 2012 年 6 月至 12 月开展了这项审查工作。这不是一次全面的方案评价,审查工作遵循了监督厅和难民署商定的职权范围。在赞赏难民署根据其历届高级专员自 1999 年以来所作决定而整合其评价和政策制订工作的同时,监督厅将其审查仅限于该组织的评价职能。在本报告中,监督厅所指的“方案评价”包括难民署法定任务的所有范围,包括在某一特定国家业务中的保护、解决办法和援助活动。方案评价包括在国家、区域和全球一级的评价。
5. 在评估难民署的评价能力时,监督厅采用了联合国评价小组认可的方法,而难民署是该小组的成员之一。这个方法是为《联合国各组织评价职能专业同行审查框架》制订的,来源于基于该小组评价准则和标准的三项标准:¹
 - (a) 独立性: 独立性决定了评价的公正性;
 - (b) 可信性: 建立在质量、方法的严格性、透明性和协商性质之上的可信性会促使评价具有针对性和有效性;以及
 - (c) 有用性: 有用性包含了评价产品的针对性、使用和成本。
6. 审查工作采用了下列定量和定性数据收集方法:
 - (a) 严格审查相关文件,包括联合国内部文件和公开文件及关于难民署及其评价职能的外部文献;

¹ 见《联合国各组织评价职能专业同行审查框架》,UNEG/REF(2011)1。

(b) 以非随机抽样方式对 12 个国家办事处进行自我管理的电子问卷调查。问卷调查是在 2012 年 11 月进行的，有 9 个办事处作了答复(答复率为 75%)；²

(c) 对高级专员方案执行委员会的全部 87 位成员进行自我管理的电子问卷调查。问卷调查是在 2012 年 10 月和 11 月进行的，有 29 位成员作了答复(答复率为 33%)；³

(d) 对难民署工作人员和利益攸关方以单人、小组或通过电话的方式进行了 87 次半结构化访谈；

(e) 对在肯尼亚内罗毕和卡库马的办事处进行实地考察，以了解外地的监测和评价职能；

(f) 与类似的联合国组织，如联合国开发计划署(开发署)和世界粮食计划署(粮食署)，作了评价职能的比较；以及

(g) 对政策制订和评价处编写并在其网站上被认作评价报告的 2010 年和 2011 年期间的 28 份难民署报告作了元评价。⁴

对所有分析都用来自多重来源的数据作了三角核证，以加强审查结果的可靠性。

7. 在审查的一些关键点上，监督厅都征求了难民署的意见，并对难民署提供合作与协助表示感谢。本报告附件一载有难民署的答复。

三. 背景

8. 难民署总部的评价职能是于 1973 年设立的，当时设立了一个直接向高级专员报告的单独评价员额。自那时以来，这一职能经历了几次变更：从 1994 年至 1999 年是“检查和评价股”，从 1999 年至 2006 年被称为“评价和政策分析股”。目前担负这一职能的是政策制订和评价处。该处除负责评价工作外还担负其他一些责任并整合了政策制订、研究和知识管理等职能，以及兼管外部研究和出版物工作。该处设在执行办公室内，处长的报告线有两个分叉，一是就政策事项向高级专员提交报告，一是就评价事项向副高级专员提交报告。

² 根据下列标准挑选了 12 个办事处：难民人口的规模和类型、业务的持续年数、法定任务及地理区域。并不是所有 9 个答复者都回答了问卷调查上的每一个问题，因此所报告的调查数据只代表对所报告的那个问题作了回答的答复者的数字的公分母。

³ 答复率过低限制了从调查结果中得出一般性结论的能力。在计算调查结果时未将“无意见”答复考虑在内。

⁴ 监督厅雇用独立的咨询人进行元评价，依据四项标准：质量、可信性、独立性和有用性；这些标准的来源是联合国评价小组的《联合国系统评价准则》、《评价报告质量核定表》和《联合国各组织评价职能专业同行审查框架》。

9. 该处的职权范围还要求该处与副高级专员和助理高级专员就类似实时评价(即总部和国家办事处高级管理人员就整体相关问题提出的“快照”式快速审查要求)等问题进行密切协作,并就政策和评价问题提供咨询意见。该处的工作计划是以滚动的方式编制的,每6个月更新一次,并由执行办公室进行审查。这反映该处职能的多层面责任及其动态的、以需求作为驱动力的性质。

10. 政策制订和评价处还负责监督“难民新问题研究”工作;这是一个基于万维网、以难民、人道主义和移徙问题为重点的系列研究文件项目,该处负责为该项目接收学术及政策专家、工作人员和相关人员提交的文件。

11. 除该处进行评价外,难民署一些较大的国家办事处会以临时权宜方式将其方案评价委托出去。但是,在区域或国家一级未进行系统性和定期的分散评价。

12. 此外,监察主任办公室也对难民署在世界各地的业务和办事处进行定期检查,以确保在运作过程中做到遵守组织规则和条例。在2009年至2012年期间,该办公室进行了28次标准检查并派出了3个合规检查团。通过定期合规报告的方式对检查团的后续行动进行了监测。该办公室自2011年7月以来发布了涵盖2008年至2012年期间的27个检查结束备忘录。

四. 联合国系统的评价工作

13. 在评估难民署当前的评价能力时,应首先说明一个强有力和有意义的的评价职能究竟是什么?联合国系统的方案评价有两个基本目的:(a)向捐助方、高级领导层和受益者说明方案的问责情况;以及(b)确定方案的实效及所获得的经验教训以改善方案。

14. 在评价类似难民署这样的机构时所采用的全面方法需要回答三个关键问题:

(a) 我们在做正确的事情吗?这个问题涉及方案干预和业务的相关性、性质和范围。具体而言,提出这个问题的目的是要说明有关方案的设计是否恰当、是否能充分说明方案试图要解决的那些问题的性质;

(b) 我们在以正确的方式做事情吗?这个问题涉及方案干预和业务的质量和效率。具体而言,提出这个问题的目的是要说明哪些做得好、哪些做得不好,以及如何改进执行工作;

(c) 我们做这些事情时的规模是否大到足以促成变化?提出这个问题的目的是要说明方案的覆盖范围、实效和影响,包括方案的成本效益和增值情况。这是个最难回答的问题,需要采用严格的混合方法,以结构化方式对多重数据源的分析结果进行三角核证。这个问题的目的是要确定方案在多大程度上实现了既定目标并对更大范围的影响作出了何种贡献。

15. 一个实体首先必须确立一个全面的监测和评价框架，以便对这些基本问题作常规及定期评估。必须收集关于实体业绩的可靠及可信信息，以帮助作战略性决策及预算和资源分配。

16. 当干预及方案业务得到很好的界定、执行并持续了一段时间之后，同样需要对质量和可持续性作定期及持续评估，以确定如何进行改善。一个全面的评价组合应包括着眼于国家、区域及全球一级的评价。评价结果将就方案的整体业绩情况提供互补信息并利用各项指标来追踪机构产出、成果和影响。

五. 审查结果

A. 难民署未明确界定评价职能

(a) 难民署的现行评价政策未把评价与组织任务和目标充分联系在一起

17. 难民署的评价政策在 2010 年 8 月作了修订，界定了用于指导难民署评价工作的总框架，就为该职确立明确及强有力的特征而言至关重要。难民署的现行评价政策把评价称为是“对本组织的政策、方案、做法、伙伴关系和程序作尽可能系统化及客观的分析和评估，重点是规划、设计、执行和影响”。但是，该政策未阐明难民署评价工作的明确及独特作用，特别是与其他组织职能(如检查和考绩)相关的作用。该政策既未说明如何在难民署的具体情况下开展评价工作，也未将评价与该组织的具体任务、目标和战略优先事项充分联系在一起，而联合国评价小组的评价政策准则和标准是有这样的要求的。⁵ 例如，开发署的政策将评价的目的定义为：评估开发署对方案国在消除贫穷、经济增长和可持续发展方面的影响；难民署的政策与难民署的具体及独特任务之间缺乏这样的联系。最后，该政策未明确阐述难民署评价工作的作用和责任。虽然政策制订和评价处被称作“评价工作的协调中心”，但是该政策几乎没有提及其具体任务和责任，也没有说明希望各区域局和国家办事处发挥何种作用和作出何种贡献。

(b) 被称作评价报告的文件里有各种不同类型的出版物

18. 在政策制订和评价处网站上被认作 2010 年和 2011 年评价报告的 28 份报告里有各种题材和性质的报告。一些报告明显是对难民署干预工作的评价。但另有一些则是政策文件、激发思考的文章和学术研究成果，与对难民署干预工作的评价无直接关系。看不出有何明显理由要把这些文章归类为评价报告，而在该网站上却作出了这样的归类。

(c) 评价资源太少

19. 目前专用于难民署评价工作的资源太少。在 2012-2013 两年期中，确认供政策制订和评价处使用的资源估计为 360 万美元，比 2010-2011 两年期的 300 万美

⁵ 《联合国系统评价标准》，标准 1.2。

元和 2008–2009 两年期的 210 万美元有所增长，但不如联合国类似实体的评价资源那样多。例如，粮食署 2012 年的评价咨询预算为将近 500 万美元，而难民署在同一时期的咨询预算则仅为 37.5 万美元。总体而言，与难民署工作的规模和广度相比，专用于支付评价职能方面的人事费用的资源非常少。评价资源占难民署总预算的 0.05%至 0.08%，低于为评价工作确立的标准基线，即方案总预算的 3%至 5%。

(d) 政策制订和评价处发挥了有用的职能

20. 但政策制订和评价处在其目前的任务和资源范围内在该组织中发挥了有益作用。该处短小精悍，5 名工作人员(1 个D-1、1 个P-5、2 个P-4 和 1 个一般事务人员)担负起评价、政策制订和研究这三项综合职能，在 2006 年至 2011 年期间，共产出 61 份评价/审查报告、5 份政策报告和 106 份万维网上的“新问题”出版物。⁶ 难民署各利益攸关方对其工作普遍反应良好。在对执行委员会调查作出答复的人中，绝大多数(80%)对来自该处的信息表示非常满意或比较满意。同样，接受访谈的总部高级管理人员也指出，该处就关键政策问题提供了有用的信息。在国家办事处的调查答复者中，绝大多数(回答该问题的 7 个中有 6 个)报告说，他们利用了该处的报告，还具体提到对他们在各自方案业务中面临的问题的分析结果和有关该组织内部良好做法的报告特别有用。

21. 特别是，该处在政策问题上在该组织内颇具影响力。该处经常被要求为高级专员的讲话稿提供资料，近些年来，该处还应高级专员关于提供定向信息的要求而担当起内部“智囊团”的角色。该处还负责监督“难民新问题研究”工作；这是一个基于万维网、以难民、人道主义和移徙问题为重点的系列研究文件项目，该处负责为该项目接收学术及政策专家、工作人员和相关人员提交的文件，并为最重要的出版物——《世界难民状况》提供资料。此外，该处被视为支持发展实时评价的先行者，而实时评价已在联合国系统内外的人道主义工作中得到普遍应用。至于战略性、整体层面的政策咨询，正是该处帮助促使难民署考虑到类似城市难民和在方案制定和执行中采用年龄、性别及多样性方法这样的问题。

B. 难民署当前的评价能力：未完全符合联合国评价小组关于独立性、可信性和有用性的关键准则和标准

22. 尽管未作明确界定，但是监督厅仍对难民署目前确实具有的评价能力作了评估。这一能力受到了限制，因为在来源于联合国评价小组准则和标准的三项关键标准(独立性、可行性和有用性)方面存在限制因素。下文将详细讨论这三项标准中的每一项标准。

⁶ 资料来源：监督厅所作的难民署网站上的报告汇编。

(a) 独立性：难民署的评价职能限制了独立性

23. 独立性标准源自于联合国评价小组的准则 6；该条准则规定，评价流程应公正并在其实施过程中独立于与一个组织的决策、方案交付和管理有关的流程。⁷ 该条准则还要求采取系统性措施，以确保评价的必要客观性和公正性。

24. 这条标准很重要，因为它要确保在理论和实践上都能以独立的方式施行评价职能，并使那些开展评价工作的人没有偏见或看上去没有偏见。这项职能应置于组织一级，从而使其独立于在其管辖之下的其他管理和方案职能。通常该职能应放在高级管理层的执行办公室内或作为一个单独的单位独立存在。

25. 此外，独立评价职能的负责人应直接向有关组织的理事会或负责人报告（见联合国评价小组标准 1.1）。在实践中，这就意味着向与作为评价对象的方案无关的高级管理层报告。这通常是有关组织的最高层管理人员。同样，评价职能可全权决定将其报告直接提交给与评价对象有关的适当的决策层审议。此外，必须建立一套制度，以便对评价结果、结论和建议进行系统性审议，同时还能将相关信息用于组织的决策（见联合国评价小组准则 2.6）。

26. 在难民署，评价工作被整合进并嵌入一个较大的负责评价、政策制订和研究的部门（政策制订和评价处）之内；难民署的评价政策规定，该处还应协助执行办公室确定是否需要发表新的或经修订的政策声明，并应与其他司、部门和区域局密切协商，为政策声明的形成、传播和宣传作出贡献。⁸ 该处工作人员负责开展所有这三项活动，并未指定专人负责评价工作。该处的政策责任范围很广，负责拟订全球业务政策（涉及财务、行政、采购或人力资源管理的政策除外）、前瞻性政策分析和发展能力，以及与各司协商确定是否需要发表新的或经修订的政策声明。⁸ 虽然难民署有意想通过把这些职能都置于该处而试图把政策与评价联系起来，以确保将评价结果直接用于决策进程，但由于该处担负着制订政策和评价该政策执行情况的双重角色，因此就有可能产生实际或据认为的利益冲突。在难民署，在与其他实体开展联合评价方面还存在着其他利益冲突风险，因为在这种评价工作中，该处同时以评价伙伴和负责代表难民署协调作出管理层答复的难民署内的实体的身份参与评价。

27. 此外，虽然将政策制订和评价处置于高级专员办公室内的做法在促进结构独立方面是合适的，但该处处长就政策事项向高级专员提出报告，就评价事项向副高级专员提出报告。⁹ 这种报告线把评价置于直接的资源管理报告线内（虽然副高级专员确实担负着监督责任），从而有可能会损害独立性。这种做法也不符合

⁷ 见联合国评价小组的《联合国系统评价准则》。

⁸ 见难民署评价政策，2010 年 8 月，第 9 页。

⁹ 见难民署的组织结构和评价政策。

联合国其他实体的良好做法。例如在粮食署和开发署，评价主任向执行主任和执行局报告，不存在向其他高级管理人员报告的双重报告线。

28. 难民署目前的工作规划流程也对该职能的独立性构成了风险，因为这一流程在很大程度上是由需求驱动的，虽然政策制订和评价处在设计其评价方法时有完全的独立性。目前不存在一个基于风险的战略工作规划系统，在与关键利益攸关方协商后，该系统本可为独立评估难民署的效率、实效和影响及独立评估对该组织构成最大风险、因而值得进行评价的主题打下基础。

29. 虽然高级专员方案执行委员会不是难民署的理事会，但是该委员会担当着重要的咨询角色，负责核准该组织的预算。政策制订和评价处已经形成了一种向执行委员会提供关于相关年份活动的年度报告的良好做法(例如，见 A/AC.96/1115)。但是，评价报告本身则未直接提交供讨论或审查，而仍然必须经过行政和预算问题咨询委员会和大会第五委员会，执行委员会仅起咨询作用。难民署必须在向其执行委员会和向其他机构提交报告之间实现同样的正确平衡。

30. 来自执行委员会成员的反馈进一步表明，必须加强评价职能的独立性。在执行委员会的调查答复者中，绝大多数人(64%)报告说，该委员会应负责确保以独立和公正的方式进行评价；占同样比例的调查答复者赞同该委员会应负责确保评价工作有利于决策和管理。执行委员会的调查答复者还表示，应安排更多的时间来讨论评价问题，并要求政策制订和评价处向他们提供更多关于难民署整体评价流程的信息。执行委员会的调查答复者最后建议，政策制订和评价处的年度报告应包含更多关于评价工作的信息，例如关于实效、效率和影响的评价结果，并说明如何将评价结果纳入及用于改善难民署的工作方案。

(b) 可信性：难民署评价职能的可信性因在评估成果的方法方面有限制而有局限性

31. 可信性标准来源于涉及评价的公正性、质量、能力和道德操守的联合国评价小组准则 5、8、9 和 11。可信性依赖于评价者的专门知识及独立性和评价流程的透明度。可信的评价应是公正和在方法上严格的评价，既要说明方案成绩，也要指出方案挑战。要考虑到所有利益攸关方的看法，如果有关各方有不同看法，则在评价分析和报告中应有所反映(见联合国评价小组准则 5)。可信的评价同样应以适当的方法进行设计和规划。

32. 虽然政策制订和评价处工作人员有各种相关业务经验，但他们的评价能力则较为有限。在联合国，评价职能以确保工作人员具备进行评价和管理外聘评价员(咨询师)所需的评价经验和技能的方式来解决可信性问题。采用轮换政策的机构通常会确保保持一些合格的核心工作人员，有些机构，例如粮食署，则维持评价工作人员和方案工作人员各占一半的混合比例。这就确保不管从评价角度还是从方案角度来看，评价都是可信的。共同导则和一致性的评价方法同样会加强可信性。

33. 虽然关于政策制订和评价处处长的岗位说明要求具备评价经验，但未要求该单位其他专业人员具备这类经验。后者是 4 年或 5 年轮换一次的内部工作人员，通常未受过评价培训，不具备与评价有关的经验或能力。由于难民署的人力资源规则要求进行工作人员轮换，因此该处在总部没有一支负责执行该职能的稳定队伍。因而该处的现任处长必须对专业人员进行有关基本评价方法及方式的培训。

34. 调查和访谈数据提供了关于难民署因缺乏合格评价工作人员而出现问题的进一步证据。在对国家办事处调查作了答复的 7 个国家办事处中有 4 个办事处主动指出，为加强评价职能，需要在国家一级进行更多评价培训和提供更多专门知识。同样，接受访谈的总部关键工作人员确认需要加强评价方法，以便能更有效地评估难民署外地工作的影响。

35. 为了补充工作人员资源，政策制订和评价处还利用身为有经验的评价员的咨询师来协助落实其工作计划，虽然这方面的资源一直很少。该处的咨询预算近年来一直维持不变，大约在 35 万至 37.5 万美元之间，与难民署整体预算相比没有增长。因此，利用具有评价技能的专家咨询师的做法并不能完全弥合因评价工作人员的评价背景有限而产生的差距。

36. 监督厅对政策制订和评价处在 2010 年和 2011 年期间发表的 28 份报告进行了元评价，目的是确定这些报告是否符合联合国评价小组的准则和标准并评估整体优点和弱点。¹⁰ 该处报告的整体优点包括说明进行评价/研究的情况及总的理由、确定报告的主题、列入不同的观点，以及各项建议具有可管理性。总的来说，这些报告写得不错，逻辑结构也很好。例如，该处发表了关于持久性难民状况和关于混合移民现象的两个系列评价报告。这些报告大多数都采用标准化格式，因此读者易于熟悉其结构并知道能从这些系列文件中得到什么。报告还明确说明进行评价/研究的理由及其对难民署战略利益的重要性，并经常援引高级专员确认的某种必要性。

37. 但是，被审查的报告在质量上的得分则要低一些，在被评估的所有报告中有很大差别，这主要是因为在评估方案结果和影响时存在方法方面的限制；应当承认，鉴于难民署法定任务的特性，要评估其方案结果和影响确实是一项挑战。这些报告往往未对难民署干预措施的结果作出明确结论，而仅强调了一般性的挑战、限制因素或问题。即使在明显是评价报告的报告中，也往往看不出该组织的工作结果是什么。这些报告缺乏关于评价标准、数据收集方法的信息和分析框架。特别是政策制订和评价处工作人员和混合小组进行的评价更是如此。在属于评价报告的报告中，由外聘咨询师在负责敲定评价结果的该处工作人员管理下编写的评价报告被认为比该处工作人员和混合小组产出的报告的质量要高一些。

¹⁰原评价是“对评价的系统性审查，以确定评价流程和结果的质量”(Cooksy 和 Caracelli, “Quality, Context and Use: Issues in Achieving the Goals of Meta-evaluation” (2005 年))。

38. 同元评价的结果相似，难民署各利益攸关方也报告说，对难民署在法定任务所有方面的业绩和结果的评价不足，包括其保护及方案工作。在执行委员会调查的答复者中，大多数人(55%)报告说，政策制订和评价处在为决策提供循证信息方面不太有效或非常无效，有的答复者还建议，需要把评价结果纳入难民署的工作方案。其中一些答复者还主动提出，难民署的评价报告应更清楚地阐明其评价结果和结论的依据。此外，总部的高级管理人员确认，在开展强有力的评价以确定难民署的外地成果和影响方面存在差距，同时还缺乏评价证据来说明难民署的成本效益和实效情况。在7个国家办事处调查答复者中，有2个办事处指出，难民署的评价报告只在很小的程度上或根本没有向高级管理人员提供对所取得的成果的评估结果，有4个办事处则指出，只在某种程度上提供了。同样，大多数(7个中的4个)办事处指出，评价报告只在很小的程度上向利益攸关方提供了让他们提出对难民署法定任务活动的评估意见的机会。

(c) 有用性：对评价结果的利用和对建议采取的后续行动有限构成了对学习和问责制的重大风险。

39. 有用性标准来源于涉及评价的定义、责任、质量、透明度、咨询及后续行动的联合国评价小组准则1、2、8、10和12。有用性依赖于被认为是相关和有用的评价结果是否清楚和简洁，是否能对决策产生影响，包括通过采取系统性的后续行动。为确保最低限度地利用评价报告，应确立对建议采取后续行动及落实建议的清晰程序，同时确立检查落实及进展情况的时间表。在联合国内，包括开发署和粮食署在内，评价职能的一个通常做法是，为管理层对评价报告作出答复和为某个论坛(例如在理事会里)讨论这类报告确立清晰的程序。

40. 目前难民署没有一个用于讨论评价结果和追踪评价建议的常规及系统性的后续流程。¹¹ 这就破坏了评价报告在总部和外地的有用性，并对学习和问责制构成了重大风险。难民署管理层、工作人员或执行委员会不对评价报告采取系统性的后续行动，管理层也不对这类报告作正式答复。不存在一个负责审议评价结果和结论的正式机制。此外，虽然在难民署网站上刊登了这些报告，但是缺乏一个从这类评价报告中吸取经验教训并将这些经验教训纳入组织的工作流程和方案的正式机制。在执行委员会调查答复者中有将近一半答复者(46%)指出，委员会应确保对评价结果、结论和建议进行系统性审议。

41. 难民署还缺乏一个用于追踪建议落实情况的系统性方法。没有一个可用于评估建议是否被公布的正式机制，也没有任何个人或实体承担公布这类建议的责任。接受访谈的总部及外地工作人员普遍认为，需要改进这一不足之处。此外，在对该问题提出意见的执行委员会调查答复者中，大多数人(56%)报告说，难民

¹¹ 难民署报告说，自监督厅审查后，难民署已要求管理层作出答复。但是在进行本次审查时这种要求尚未到位。

署在提供关于如何将结果、结论和建议纳入难民署工作的后续信息方面不太有效或非常无效。大多数人(56%)还指出,他们对政策制订和评价处就建议采取的后续行动不太满意或非常不满意。事实上,评价的后续行动程序是在目前评价职能所有组成部分中被执行委员会成员打分打得最低的一个组成部分。如果不对评价采取足够的后续行动,就很难评估组织的影响和资金效益。

42. 2012年6月成立的难民署内部合规和问责制委员会具有解决在后续行动方面的这一不足之处的潜力。成立该委员会的目的是为了加强组织及个人问责制,并使之成为负责对审计、评价、检查及临时报告提出的建议排出优先秩序并对落实情况监测的内部中央机关。接受访谈的总部及外地工作人员还确认,内部合规和问责制委员会可以成为一个负责将从评价中获得的经验教训用于政策和方案编制工作从而促进知识管理的机构。鉴于该委员会刚成立不久,因而尚待观察它是否有能力解决本次审查所查明的在对评价的后续行动方面存在的一些差距。

C. 缺乏联合国其他具有强大实地存在的实体通常都有的一个功能性的中央及分散的评价系统

43. 鉴于难民署大部分业务都在远离总部的外地,一个包含中央及分散评价的评价职能将有助于评估其在各级(外地、区域和总部)的工作。中央评价的范围较广,负责评估全球、全组织一级的工作。中央评价的实施独立于那些负责保护及方案业务的人,独立于作为评价对象的结果。中央评价要解决全组织范围、事关整个组织的交叉战略问题。中央评价还要独立评估分散评价的质量并确立对组织内分散评价活动进行评估的标准。一般而言,这一职能应在总部所在地实施。

44. 分散评价的范围则相对要窄一些。分散评价应在外地一级(包括国家办事处和区域办事处)实施,并由一个总部部门负责支助和协调分散评价活动,并酌情综合评价结果和提出报告。分散评价应嵌入方案业务、作为评价对象的结果内,并应解决对国家及区域方案编制工作具有具体相关性和优先性的问题。分散的做法能确保所获得的证据对当地的情况是有意义的,并有助于为制订国家政策提供信息。一般而言,这一职能应在业务一级由区域和/或外地办事处负责实施,同时与在总部的同评价对象相关的一个总部司保持密切联系。

45. 为使分散评价职能可信,并遵循联合国其他业务实体的良好做法,必须达到一些标准。首先,需要有质量保证,通常这是中央评价职能的责任,例如开发署和联合国人口基金就是这样做的,它们每年编写一份关于对国家评价的质量评估报告。同样,需要确立共同的组织标准和程序,以及对外地工作人员进行基本评价行为方面的培训,以便促进不同国家办事处和区域办事处之间的一致性和可靠性。还需要有一个工作规划流程,以确保有一个常规及有协调的分散评价方案;通常这部分工作由负责外地业务的总部办公室实施。

46. 具有中央及分散评价职能的联合国其他实体面临的两项挑战是：(a) 需要清楚界定中央及分散评价活动的范围，包括各自的作用和责任；以及(b) 需要确立一种综合和全面的评价职能，以避免对组织内的评价工作采用脱节的方法。通过确立清晰和准确的评价政策、协调机制，以及通过定期和公开的交流，解决了这两项挑战。

47. 难民署缺乏常规及系统性的分散评价职能，导致关于难民署国家方案(包括其保护、解决和援助活动)执行情况的评价证据很有限。正如对在肯尼亚的难民署办事处的实地考察结果所显示的，目前没有对各国家办事处所取得的成果作常规及系统性评估，而本来这类评估可就未来的国家方案编制工作提供有益的决策信息，特别是提供关于整体重点、战略优先事项、关键活动及对受益者的影响的信息。

48. 难民署那些确实得到实施的分散评价一般都是临时性质的，并且是根据国家办事处管理人员的自由裁量权进行的，在范围、时机和类型上都有所不同。在 9 个国家办事处调查答复者中有 4 个答复者报告说，在过去两年中进行了一次评价。最值得注意的是，政策制订和评价处在以需求驱动的基础上，为应对紧急情况而在主管业务助理高级专员的指导下进行了实时评价。现在已开始重视要为这类活动拟订正规协议的必要性。在某些情况下，这类评价是作为机构间活动展开的，目的是向管理层提供一个“快照”，以说明进展情况并指出需要立即得到业务关注的限制因素。

49. 目前在国家一级进行的方案执行情况报告有可能可以提供信息并向更系统性的分散评价提供支助。这一活动的很大部分属于年度方案编制周期及其监测和报告要求的组成部分。方案执行情况报告提供了关于方案编制活动及在实现成果方面的进展情况的有用信息。例如，在肯尼亚的卡库马外地办事处开展了年度业务审查；这是一个参与性进程，目的是与执行伙伴和受益者一起审查难民署的成果。

50. 此外，可更好地将方案支助和管理司和国际保护事务司的现有能力用于加强分散评价。现在有机会在现有活动和资源的基础上进行战略性建设，通过加强这两个司里的现有评价部门，为分散评价建立一个更系统性的结构。

51. 难民署在发展中央及分散评价系统时可参考联合国里的一些模式。例如，粮食署有一个中央和分散评价职能；¹² 中央评价职能设在总部，负责以全球视野作战略评价；分散评价职能遵循与中央评价同样的标准，由区域局和国家办事处负责实施，评价重点是国家业务。¹³ 粮食署依据其中央和分散评价职能实施组合评价，包括：

¹² 见粮食署 2008 年评价政策，第 18 段。

¹³ 见同上，第 13 和第 21 段。

- 采用全球视野和着重于整个粮食署战略问题的战略评价(中央评价)
- 用于检查粮食署政策在多大程度上实现其目标的政策评价(中央和分散评价)
- 用于帮助国家办事处了解其组合及这类组合促成了哪些变化的国家组合评价(分散评价)
- 用于深度分析粮食署的工作对受益者带来的影响的影响评价(中央评价)
- 用于了解业务相关性、效率和实效的业务评价(中央和分散评价)
- 与其他组织就上述任一类型评价联合实施的联合评价¹⁴

52. 另一个与中央和分散评价相关的模式就是开发署。开发署评价政策第 17 段确认了该组织内的两类评价：由评价办公室实施的独立(中央)评价和经方案单位委托由外聘独立专家实施的分散评价。该政策还把中央评价办公室定义为开发署评价职能的监管单位，负责规划独立评价、实施专题及国家方案评价(称为“发展成果评估”)，以及确保其评价能对开发署的成果作战略性覆盖并用于决策。所有独立评价报告都会随同国家方案文件一起提交给执行局。中央评价职能还为分散评价的规划、实施和利用设立了标准，并负责评估这类评价的质量。分散评价则由区域局和国家办事处委托实施，评价重点是外地一级的方案成果。评价办公室建立了评价资源中心数据库，里面包含为国家方案制订的评价计划。所有根据这些评价计划编写的报告都由评价办公室进行评估，在其向开发署执行局提交的年度报告里会就分散评价在遵守评价计划方面的情况及其质量提出报告。开发署的所有评价报告和管理层的答复都会在开发署的评价资源中心网站上公布，供公众查阅。

六. 结论

53. 鉴于难民署在其工作中面临关键问题并对个体生命产生直接影响，因此尽管充满挑战性，但还是必须提供关于方案结果和成果的循证评价，以指导在方案设计和执行方面的决策。难民署必须了解其所开展的工作是否有成效和是否具有成本效益，是否可改变其行事方式以产生更大的影响。为实现这个目标，难民署就需要有一个强大、强有力和具有充分全面性的评价职能。虽然政策制订和评价处在该组织起到了关键和有用的作用，但鉴于其目前的法定任务和资源，该处无法充分满足该组织不断变化的评价需求。

54. 难民署需要回答的一个关键问题是：它希望评价在组织中发挥什么作用？一旦有了答案，高级管理层就需要倡导该职能并确保为实施评价营造足够强大和有

¹⁴ 见 <http://www.wfp.org/about/evaluation> 和粮食署 2008 年评价政策，WFP/EB. 2/2008/4-A。

利的环境。其中的关键就是要在难民署内建设并支持一种评价文化，即要将评价视作是在为该组织的工作增加价值，以及视作是在展示难民署的成果、对受益者生活的影响及资金效益。

55. 难民署可采取一些重大步骤来加强其评价职能。就评价工作与难民署总体任务、预期成绩和目标之间的联系展开更为明确和详细的讨论，界定评价对实现成果的预期贡献，从而促进加强评价活动。在评价政策中应对此作出具体规定。难民署的评价应回答三个关键问题：正在做正确的事情吗？正以正确的方法做事吗？做这些事时的规模大到足以促成变化吗？

56. 难民署还可从一个同时包含中央及分散评价的强有力的评价职能中受益。中央评价将解决全组织范围的战略问题，在实施时应独立于那些负责方案业务的人和作为评价对象的结果。分散评价的范围要窄一些，应嵌入在方案业务之中。难民署可以更战略性地利用方案支助和管理司和国际保护事务司的现有资源和活动，以加强其分散评价并使之常规化。

57. 对成果制进行投资将有可能振兴作为综合监测和评价系统组成部分的评价职能。难民署重视并自 2006 年以来为建立一个广泛的成果制投入了数以百万美元计的资金，这个对规划、管理、数据收集及对其年度方案周期的循证评估采取逻辑方法的成果制是一个重要的组成部分，应纳入难民署范围更广的评价系统之内。

58. 加强难民署评价职能以便更好地为组织成果、学习和问责制作出贡献，这不是一件一蹴而就的事情；要做到这一点就需要有长远的目标和战略，以及强有力的领导。监督厅为启动这一进程提出了五项重要的建议。

七. 建议

建议 1 (成果 A)

59. 难民署应在总部设立一个专职和单列的评价单位，负责实施难民署的方案和交叉战略评价。在设立这个单位时，应满足三项标准：

- (a) 该单位应独立于任何其他方案业务；
- (b) 该单位应仅开展评价工作，不应承担任何其他责任(如政策拟订)；以及
- (c) 该单位应在评价规划、实施和报告方面享有足够的组织独立性。

60. 难民署在执行本建议时有几个选择，包括：将评价职能留在政策制订和评价处内，但使之成为一个单独的单位，其工作人员只负责评价工作；将评价职能划归监察主任办公室；以及新设一个直接向高级专员报告的完全单列的单位。

61. 难民署对本建议表示部分赞同。

建议 2(成果 A 和 C)

62. 难民署应修订其评价政策，以便为在组织内如何实施评价确立一个适当的框架，包括在本次审查中所确认的差距。具体而言，难民署应：

(a) 充分明确地解释评价在难民署具体组织范围内的概念和作用；

(b) 确立一个包含在总部一级实施的中央评价和在外地(区域和国家)一级实施的分散评价在内的系统，并制订一个以这两个职能之间的互补性和协同效应为基础的明确战略；

(c) 更明确地界定和划分中央评价和分散评价各自的作用、责任和目标；

(d) 明确评价的报告线，包括考虑向执行委员会直接提交关键、战略性的评价报告；以及

(e) 根据本组织的方案逻辑制订一个常规和系统性的评价工作规划流程，要顾及难民署开展的所有类型工作及临时要求、法定任务或与联合国其他机构或办事处开展的联合评价。

63. 难民署对本建议表示赞同。

建议 3(成果 B)

64. 难民署应确立用于评估难民署成果的程序和方法，以加强其评价的严格性和有用性。具体而言，应包括：

(a) 拟订评价的职权范围，要明确提出评价标准、问题及数据收集方法，以及评价设计；

(b) 拟订能评估难民署外地干预措施的实效和影响的评价方法；以及

(c) 提交能清楚和直接说明难民署所取得的成果的报告。

65. 难民署对本建议表示赞同。

建议 4(成果 B)

66. 难民署应为评价建议的后续行动拟订一个常规和系统性的流程。具体而言，应包括：

(a) 为拟订管理层对所有评价的答复设立明确的机制；

(b) 要求对所有评价制订行动计划，说明如何和何时落实各项建议；

(c) 为就落实建议而采取的定期和常规后续行动制订程序；以及

(d) 为落实建议确立明确的问责制。

67. 难民署对本建议表示赞同。

建议 5 (成果 C)

68. 难民署应为加强外地的分散评价拟订一个战略。具体而言，该战略应考虑到：

- (a) 区域和国家一级常规和系统性评价所需的流程和机制；
- (b) 如何更好地利用方案支助和管理司和国际保护事务司现有的能力；
- (c) 在拟订国家一级评价职能时要考虑到国家方案活动的规模和类型；
- (d) 如何在外地的方案管理人员中建设一种评价文化；
- (e) 中央评价和分散评价各自的作用和责任；以及
- (f) 应在何处和何时将中央及分散评价纳入难民署的方案编制周期。

69. 难民署对本建议表示赞同。

主管内部监督事务副秘书长

卡门·拉普安特(签名)

2013 年 3 月 22 日

附件^{*}

2013 年 3 月 20 日联合国难民事务高级专员公署副高级专员 给内部监督事务厅检查和评价司的备忘录

难民署对监督厅关于难民署评价能力的审查的答复

1. 难民署对监督厅关于难民署评价能力的审查表示欢迎，但对该报告的以下几点有不同看法：

难民署当前的评价能力未完全符合联合国评价小组关于独立性、可信性和有用性的关键准则和标准；特别是，难民署评价职能的可信性因在评估成果的方法方面有限制而有局限性。

答复：

在联合国系统里，评价职能的结构不尽相同，但监督厅的报告似乎表示只有一个标准模式。在难民署，政策制订和评价处是一个统一的结构，享有充分的方法上的独立性。此外，难民署的评价政策可预防高级管理层干预评价结果和建议。难民署认为，对评价采用一种由需求驱动的方法并不会损坏独立性。恰恰相反，需求驱动法更有可能确保评价专注于对组织具有战略重要性和风险较高的业务和问题。如果对政策制订和评价处实施的评价作彻底审查就会发现，高级管理层并未要求该处评价那些被认为成功的业务。事实上，往往是因为看到了组织方面的困难和弱点才启动评价工作的。

本组织拒绝接受关于其目前的评价缺乏严格性和未专注于难民署干预措施的实效和影响的说法。此外，政策制订和评价处和执行办公室从执行委员会成员、联合国其他机构、非政府组织和人道主义界其他成员那里获得的积极反馈意见和互动并不支持监督厅的难民署评价职能缺乏可信性这一评论意见。特别是主要捐助方始终对该处及其工作表示支持。

2. 难民署赞赏审查报告提出的各项建议，并就各项建议提出下列意见：

建议 1 (成果 A)

难民署应在总部设立一个专职和独立的评价单位，负责实施难民署的方案和交叉战略评价。在设立这个单位时，应满足三项标准：

- (a) 该单位应独立于任何其他方案业务；
- (b) 该单位应仅开展评价工作，不应承担任何其他责任(如政策拟订)；以及

^{*} 在本附件中，监督厅全文抄录了来自难民事务高级专员办事处的评论意见。这一做法的根据是大会第 64/263 号决议，也是独立审计咨询委员会所建议的。

(c) 该单位应在评价规划、实施和报告方面享有足够的组织独立性。

难民署在执行本建议时有几个选择，包括：将评价职能留在政策制订和评价处内，但使之成为一个单独的单位，其工作人员只负责评价工作；将评价职能划归监察主任办公室；以及新设一个直接向高级专员报告的完全单列的单位。

建议 1：难民署对本建议表示部分赞同，并提出如下意见：

难民署认为评价和政策分析是整体职能。此外，难民署重视在评价规划、实施和报告方面的组织独立性，并同意需要有与其年度预算相称的能力和资源来实施评价。但这需要作进一步内部分析和思考，还应考虑到难民署内部今后的监督工作和监督厅在本次报告里提出的各项选择，以便规划前进的道路。

建议 2(成果 A 和 C)

难民署应修订其评价政策，以便为在组织内如何实施评价确立一个适当的框架，包括在本次审查中所确认的差距。具体而言，难民署应：

(a) 充分明确地解释评价在难民署具体组织范围内的概念和作用；

(b) 确立一个包含在总部一级实施的中央评价和在外地(区域和国家)一级实施的分散评价在内的系统，并制订一个以这两个职能之间的互补性和协同效应为基础的明确战略；

(c) 更明确地界定和划分中央评价和分散评价各自的作用、责任和目标；

(d) 明确评价的报告线，包括考虑向执行委员会直接提交关键、战略性的评价报告；以及

(e) 根据本组织的方案逻辑制订一个常规和系统性的评价工作规划流程，要顾及难民署开展的所有类型工作及临时要求、法定任务或与联合国其他机构或办事处开展的联合评价。

建议 2：难民署对本建议表示赞同，并将继续就战略性关键评价向执行委员会定期提交报告。

建议 3(成果 B)

难民署应确立用于评估难民署成果的程序和方法，以加强其评价的严格性和有用性。具体而言，应包括：

(a) 拟订评价的职权范围，要明确提出评价标准、问题及数据收集方法，以及评价设计；

- (b) 拟订能评估难民署实地干预措施的实效和影响的评价方法；以及
- (c) 提交能清楚和直接说明难民署所取得的成果的报告。

建议 3：难民署对本建议表示赞同。

建议 4(成果 B)

难民署应为评价建议的后续行动拟订一个常规和系统性的流程。具体而言，应包括：

- (a) 为拟订管理层对所有评价的答复设立明确的机制；
- (b) 要求对所有评价制订行动计划，说明如何和何时落实各项建议；
- (c) 为就落实建议而采取的定期和常规后续行动制订程序；以及
- (d) 为落实建议确立明确的问责制。

建议 4：难民署对本建议表示赞同，并于最近根据本建议作出了一项要求管理层对评价作出答复的新规定，并由新设立的内部合规和问责制委员会负责监督实施。

建议 5(成果 C)

难民署应为加强外地的分散评价拟订一个战略。具体而言，该战略应考虑到：

- (a) 区域和国家一级常规和系统性评价所需的流程和机制；
- (b) 如何更好地利用方案支助和管理司和国际保护事务司现有的能力；
- (c) 在拟订国家一级评价职能时要考虑到国家方案活动的规模和类型；
- (d) 如何在外地的方案管理人员中建设一种评价文化；
- (e) 中央评价和分散评价各自的作用和责任；以及
- (f) 应在何处和何时将中央及分散评价纳入难民署的方案编制周期。

建议 5：难民署对本建议表示赞同，并提出下列意见：

必须将难民署内的分散评价职能视作为长期流程，因为该职能具有重要的资源和能力影响。这将包括增强专门知识和能力，以便向分散评价职能提供质量控制并进行定期监测。