



联合国开发计划署和
联合国人口基金
执行局

Distr.: General
17 March 2009
Chinese
Original: English

2009 年年会

2009 年 5 月 26 日 6 月 5 日，纽约

临时议程项目 4

评价

开发署 2008 年年度评价报告
(2008 年 1 月 21 日至 28 日，纽约)

摘要

本年度报告涵盖期间为 2008 年 3 月至 2009 年 2 月。第一章报告了评价范围、遵守情况、质量、资源与能力；联合国评价小组采取的举措以及评价工作的运用情况。依据开发署评价办公室以及相关基金和方案机构所作评价，第二章介绍了重要的调查结果和组织学习方面的经验教训。第三章和最后一章介绍了评价办公室 2009-2010 年工作方案供执行局审批。

决定要点

执行局不妨(a) 注意本报告；(b) 请开发署处理评价工作提出的问题；(c) 请开发署加强分散评价能力并增加其使用；(d) 请开发署支助发展国家评价能力；以及(e) 核准评价办公室提出的 2009-2010 年工作方案。



目录

章次	页次
一. 评价职能	3
A. 开发署评价办公室	3
B. 联合国评价小组	6
C. 相关基金和方案机构	7
D. 方案股	9
E. 能力	13
F. 评估的使用与后续工作	15
二. 独立评价工作的主要成果和经验教训	16
A. 政策建议的实质能力	16
B. 国家能力和可持续自有权	17
C. 战略重点与一致性	17
D. 方案拟定的优势与劣势	18
三. 评价办公室 2009-2010 年工作方案	19

一. 评价职能

1. 本年度评价报告是开发署评价办公室自开发署的评价政策于 2006 年 6 月获得批准以来向执行局提交的第三份报告。在报告所述期间，评价办公室以及相关基金和方案机构的评价股的优先事项包括：提高独立评价的质量；通过扩大覆盖范围提供更多的评价依据；支持分散发展评价能力；通过参与并牵头开展联合国评价小组（评价小组）的各项举措，为联合国的改革议程做出贡献。

2. 报告突出强调了做出的努力和取得的进展，并概要介绍了本组织在执行评价政策的原则和准则方面遇到的挑战。本报告及前两份年度报告的成果将作为对首次评价政策独立审查的关键投入提交给执行局 2009 年第二届常会。

A. 开发署评价办公室

质量

3. 2007 年年度评价报告把质量、时间安排和标准化程序确定为评价办公室在开展独立评价时必须加以改进的几个领域。在报告所述期间，评价办公室侧重于提出多项举措解决这些不足之处。

4. 国家一级的方案评价方法，又称发展成果评估法，已经得到加强，彼此更加协调一致。例如，采用了一套统一的标准就开发署对联合国的协调工作所做贡献进行了评估。对发展成果程序与方法指南的最新评估中反映了这些改进。这增加了评价方法的严谨性，并能系统地提炼经验教训与良好做法。这些改进亦使评价办公室能够在更大范围内评估发展成果。评价办公室已经整理了良好做法，并确认了各项措施改进专题评价的技术质量。来自联合国、多边与双边开发机构以及各国政府的评价专家为编写新的专题评价指南提供了宝贵投入。

5. 为提高工作人员的技能与知识，评价办公室组织了知名专家讨论会。讨论会的主题包括：影响评价、概念绘制和案例研究方法。通过这些学习机会获取的知识与技术投入有助于推动评价方法的发展。

6. 评价办公室成立了外部质量保证体系。针对发展成果的每一次评估，均有两名专家审查员对初稿提供专业、客观的评论意见。专家审查员专业、客观的评论对评价工作产生了积极影响。所有的专题评价均设有顾问小组，负责审查评价计划和报告草案。

7. 为改进评价小组的两性平等和地理均衡情况，评价办公室开始协同努力，从方案国征聘更多女性评价人和专业人员。然而，目前仍有改进的余地。2008 年征聘的小组组长中，80%来自捐助国。在小组成员的征聘方面取得了更大的成功——35%的国际顾问来自方案国。联合评价联合国系统在南非的作用和贡献为评价小组吸收高比例的本国评价人树立了榜样。小组组长和两位副组长中的一位均为

南非人。此外，5 位专题专家中有 4 位是南非人。因此，与大多数外部小组相比，评价小组对国情的了解更为深入。评价办公室正致力于确认和吸引评价领域的合格国家机构和专业人士参与。

8. 在两性平等方面，男女国际顾问的人数基本持平。然而，在小组组长的职位上，妇女的在职人数严重偏低（28%），更多的女性是作为小组成员征聘的（68%）。在国内顾问中，男性占绝大多数（92%）。有必要确认合格的女性国内顾问以及能够担任小组领导职务的妇女。

9. 2008 年，发展成果评估时间表有所改进。通过与区域局的广泛磋商，仅选择那些将于 2009 年向执行局呈交新方案的国家作为评价对象。评价办公室及时向国家办事处及其国家合作伙伴通报了发展成果评估草案，以便编写新的国家方案。¹在 2009 年 4 月之前，在这一进程结束时举行的关于 2008 年开展的所有评估的利益攸关方讲习班将完成，这样就能在年度会议之前提供最后报告，年度会议将对新的国家方案进行审查。

10. 为最大限度地发挥利益攸关方讲习班的影响，评价办公室鼓励开发署高级别代表参与讲习班。开发署高级管理人员的开放态度展示了对方案改进、透明化和成果问责的承诺。区域局高级别官员的参与对确保后续支持和指导国家办事处至关重要。

11. 作为环境与能源专题评价的后续举措，在世界银行的协作下，评价办公室吸引了利益攸关方和公众参与讨论多方机构在环境与可持续性发展领域中的作用。

覆盖范围

12. 评价办公室对执行局批准的全球方案、区域方案和国家方案以及开发署战略计划中确定的成果进行了独立评价。为支持知情决策，评价办公室大幅拓展了 2008-2009 年工作方案。尤其是，发展成果评估的覆盖范围扩大了一倍以上。在阿富汗、阿根廷、巴巴多斯和东加勒比国家组织、波斯尼亚和黑塞哥维那、博茨瓦纳、危地马拉、菲律宾、塔吉克斯坦和乌兹别克斯坦完成了 9 项评价，相比之下去年只完成了 4 项评价。

13. 评价办公室对一项阿拉伯国家区域合作框架、开发署的第三次全球合作框架以及对开发署在环境与能源方面的作用与贡献进行的专题评价的成果之一进行了评价。

¹ 菲律宾发展成果评估被纳入两年期国家方案扩展部分，利益攸关方讲习班将于 2009 年中期举办。

向开发署的评价文化提供支持

14. 作为监督评价职能的机构,评价办公室支持开发署旨在改善评价文化的各项努力。评价办公室编写了年度简报,提醒开发署的管理部门留意评价过程中出现的问题。开发署承认加强评价监管的重要意义,并制订了一项指令,强化区域局在国家办事处方面的作用和监管职责。简报的主要成果是一份行动计划,由业务支助组和评价办公室共同制订。行动计划中提议改进开发署的方案拟订指南和管理行动,以加强评价文化以及对评价文化的使用。

15. 评价办公室与业务支助组以及发展政策局(发展局)紧密协作,订正了《成果监测和评价手册》(2002年),以反映评价政策的要求和原则以及发展评价方面的新趋势。由于成果规划是有效监测和评价的前提条件,因此订正手册将规划、监测和评价纳入了同一项指南。2008年底,评价办公室举办了一场讲习班,与高级管理人员、方案干事和来自国家办事处和区域局监测和评价专家共同审议了草稿初稿。

16. 为支持评价方面的管理问责,评价办公室维持了一个以网络为基础、公众可进入的评价数据库,又称为评价资源中心。²各方案股³直接向数据库上传其评价计划、参考术语、报告和管理部门的回应。2008年,评价办公室对数据库进行了扩展,纳入了联合国资本发展基金(资发基金)和联合国志愿人员方案(志愿人员方案)。在本报告陈述期间结束时,数据库中已包含1 060多份评价报告和400个参考术语,与前一年相比,分别增长了33%和51%。

17. 评价办公室与管理局携手合作,在开发署高级管理部门版面中纳入了一项管理部门回应合规指标。每天运用各方案股上传给评价资源中心的数据对这一版面进行更新。评价办公室希望这一举措能够强化本组织管理部门在评价工作中的自律。

18. 评价办公室主办了开发署的一项“电子知识”网络,名为“评价网”,以支持评价工作中的知识管理和参与评价的群体。2008年,236名新成员加入“评价网”,目前成员总数为1 226人。该网络征聘了一名全职网络管理员,以支持该网络与其他的企业知识网络交叉互动。通过相互粘贴信息,监测和评价领域的挑战 and 良好做法得以分享。

19. 评价办公室应请求向各个方案股提供咨询服务,并加强了对区域和国家一级人数不断增多的评价专家的支持,为其初步开展工作提供便利。评价办公室为150多名开发署工作人员进行了评价培训。2009年,评价办公室正与相关的总部办公室携手合作,加强对国家办事处在规划、监测和成果评估方面的咨询服务。

² 网址如下: erc.undp.org。

³ 包括国家办事处、区域局以及各个总部的政策和实践股。

资源

20. 2008 年，评价办公室的预算和人员有所增加。2008-2009 两年期的核准预算为 1 760 万美元，其中分配给 2008 年的预算为 772 万美元。分配给 2009 年的预算仍为 917 万美元，使 2008-2009 年的分配款达到 1 689 万美元。2008 年共支出 640 万美元，与前一年相比增加 45%。评价办公室现有 23 名工作人员，与前一年相比增加了 3 位专业人员和 2 位支助人员。为扩大发展成果评估的范围，这一增长是非常有必要的。

B. 联合国评价小组⁴

21. 评价办公室继续强力支持联合国评价小组（UNEG）的严谨议程，提高联合国全系统在评价工作中的一致性、质量与创新。2008 年是评价办公室的主任作为选举产生的主席任职的第二年。开发署主持联合国评价小组的秘书处工作，由评价办公室的副主任作为联合国评价小组的执行协调员负责管理秘书处。联合国评价小组 2008 年工作方案由 11 支工作队执行。开发署是关于培训、南非联合评价工作和国家一级评价工作的各个工作队的共同主席。评价办公室的工作人员参加了以下举措，涉及：“一体行动”试点的可评价性、联合评价联合国在南非的作用、影响评价以及提高评价质量等内容。联合国评价小组的举措被纳入了评价办公室工作人员的工作计划和个人业绩计划。

22. 联合国评价小组对“一体行动”的 8 个试点国家进行了可评价性研究，为联合国的改革做出了有益的贡献。可评价性研究已于 2008 年完成，是三个阶段研究进程中的第一阶段，是对试点国家的评价参数的技术评估。研究确定了试点举措在设计和执行方面最初存在的优势与劣势，以促进制定早期矫正措施，开展注重成果的监测工作。开发署是管理小组的成员，并作为联合国评价小组的秘书处为这一实践提供支持。

23. 开发署是共同协调员，它和南非政府均为联合国评价小组的联合评价管理组成员，对联合国系统在南非的作用与贡献进行评价。评价工作对联合国系统在南非的长期发展中的有效性和贡献进行了评估。2008 年 11 月，在关键的国家行为者以及联合国各组织的驻地国和非驻地国的高级管理部门之间广泛分发了报告草案，以征求评论意见。报告将于 2009 年中期最终完成，其中应提供经验教训，以指导未来由国家领导的评价活动，并成为联合国与各国机构携手共担评价责任和自主开展评价的典范。

24. 为协调和加强评价职能，联合国评价小组制定了道德指南，扩展了联合国评价小组的评价行为准则。在评价做法交流讨论会上，开发署介绍了一项用于分散

⁴ “联合国评价小组”是一个专业网络，由联合国系统各组织的 43 个评价股组成。

评价的实验性工具，“质量评估工具”，从而促成创建了一个新的工作组，以开发适用于整个联合国评价小组的评价报告质量评估工具。

25. 作为培训工作队的共同主席，开发署对评价工作的专业化做出了贡献。开发署发挥了重要作用，使评价介绍课程成为联合国系统的制度。在本报告所述期间，35 名开发署工作人员参加了四次培训课程。开发署牵头界定联合国评价小组与联合国系统职员学院之间的伙伴关系框架，以制定和执行全面的评价工作培训方案。工作队正对该方案进行“学习需求评估”。

C. 相关基金和方案机构

联合国资本发展基金

26. 联合国资本发展基金的评价股直接向执行秘书汇报，以保持其评价职能的独立性。评价工作由一名评价顾问和一名评价干事管理，并由顾问团队执行。2008 年，联合国资本发展基金为评价工作划拨了 62 万美元，该款项由各个项目预算供资。

27. 评价股授权对联合国资本发展基金在布基纳法索、海地和老挝人民民主共和国的当地发展方案进行了 3 次评价。11 月，通过对利比里亚、马达加斯加和多哥的 3 个案例研究，联合国资本发展基金启动了对区域性方案“在非洲建立包容性的金融部门”的评价工作。越来越关注组织进行联合评价：在海地，与加拿大国际开发署联手开展评价，与该署共同为该方案供资；在布基纳法索，与比利时生存基金共同开展了评价工作。

28. 联合国资本发展基金通过了统一的程序，努力推动所有评价咨询机会透明化与顾问遴选竞争上任。在尝试将咨询人管理工作外包给一家商业公司后，联合国资本发展基金安排利用联合国项目事务厅的服务来提高成本效益，压缩评价工作外部咨询支助费用。评价股仍负责指导所有的评价工作。新的安排将于 2009 年底接受审查。

29. 联合国资本发展基金开始使用“评价资源中心”管理部门回应系统。管理部门对布基纳法索评价工作的回应已上传至该数据库。联合国资本发展基金已经把阐明和后续跟踪管理部门的回应作为决策、知识管理和问责等方面的优先事项。联合国资本发展基金正采取措施，确保发展成果评估中包含对“联合国资本发展基金对其资助国家所做贡献”的评估。

30. 联合国资本发展基金是一个关于评价职能与其他职能之间的区别的工作队的共同主席，该工作队制订了一份立场文件以提交联合国评价小组 2009 年年度会议。联合国资本发展基金还参加了工作队编写评价顾问名册的工作，并对制定该名册的框架和内容做出了贡献。

联合国妇女发展基金

31. 2008 年，联合国妇女发展基金（妇发基金）评价股的工作人员包括：评价顾问——直接向执行主任报告，1 名新工作人员和 3 名咨询人。评价股完成了招聘 1 名评价分析师和 2 名区域评价专家的工作，他们将于 2009 年加入妇发基金。2008 年的预算总额为 908 132.40 美元。

32. 继公开的竞争招标过程之后，评价股发起了 4 次企业评价（其中 1 次于本报告所述期间完成）；1 次对 26 项评价的总评价；9 次审查；2004-2008 年间完成了另外 8 项评估。评价股参与了 2 次联合评价。

33. 妇发基金 2008-2011 年评价战略于 2008 年制定，该战略确立了妇发基金评估工作的四项主要目标：开展高品质的评价工作；加强工作人员和合作伙伴的评价能力；参与联合国的评价进程；促进在评价做法中采用两性平等观点；确立妇发基金战略计划评价程序。制订了一项评价政策以供 2009 年批准。

34. 评价股继续致力于制定与评价资源中心兼容的评价计划，为妇发基金在 2009 年中期之前实现一体化提供便利。妇发基金实行了分散评价计划和指导文件，支持制定评价计划，完成对 3 项评价的管理部门回应。经过以上努力，妇发基金开展了 37 项有计划的分散评价和评估。2008 年完成了 7 项分散评价。

35. 开发了多项评价工具，例如，评价预算计算模板、利益攸关方分析模板、评价小组采访指南，以及关于制订参考术语以便在分散评价工作中融入两性平等观点和人权原则的指导说明草案。评价股开展了针对工作人员和合作伙伴的能力建设评价项目的设计工作，包括与加拿大卡尔顿大学合作开发以网络为基础的评价手册和培训方案。妇发基金正在开发一个系统跟踪其内部评价能力的改进情况。

36. 评价股是联合国评价小组关于人权与两性平等、培训、“一体行动”评价、名册评价，以及影响评价的 5 个工作队的积极成员。评价股参加了发展援助委员会/联合国评价小组共同开展的关于内部监督事务厅的同行审查。作为工作队的共同主席，妇发基金为制订联合国评价小组的《关于在联合国系统的评价工作中融入人权与两性平等观点的指南》提供了资金和技术支持，该《指南》于 2009 年 3 月呈交至联合国评价小组的年度大会。

联合国志愿人员方案

37. 联合国志愿人员方案的评价股现有工作人员 4 名，预算增长了 31%，达到 202 150 美元，其中 94% 用于开展评价工作，6% 用于业绩计量。联合国志愿人员方案的评价职能在很大程度上由特别志愿人员基金供资，该基金是一个支持重要研究和尝试创新方法以加强发展工作的自愿性的一般捐助基金。为提高能力，尤其是开展更专业的评价工作的能力，评价股增设了评价专家员额，使 2009 年的工作人员总人数达到 5 人。

38. 联合国志愿人员方案完成了 6 次分散评价：5 次期终评价，1 次中期评价。其中 1 次期终评价——评价拉丁美洲将两性平等观点纳入预算的情况——是与联合国妇女发展基金共同完成。其他 5 项评价如下：在斯里兰卡、苏丹、多哥和乌克兰的国家级方案、东南欧的一项区域方案。联合国志愿人员方案评价股尚未完成独立评价；它的重点是通过培训和制订指南来发展能力以及促进从评价工作中汲取经验教训。

39. 《志愿主义对发展所做贡献评估方法》手册吸收了从 2007 年的“自愿主义促进发展成果”讲习班中汲取的经验教训，手册已经完成，并将于 2009 年初发行。

40. 2009 年，联合国志愿人员方案将依据其即将完成的“企业计划”和方案拟订战略制订一项监测和评价战略。评价股支持方案工作人员拟订项目一级的成果框架和评价工作参考术语以及最终完成评价报告，以增加评价知识，扩大这种知识在方案拟订方面的战略意义。为加强从评价工作中汲取经验教训的能力，联合国志愿人员方案编写了知识产品，例如，2006 和 2007 年评价工作综合报告、评价简报模板等。联合国志愿人员方案正在准备简报和 2000-2008 年评价工作综合报告。评价股借鉴联合国评价小组的培训举措和新的手册，正在与方案部门协作制定工作人员培训方案，以培养他们的主人翁责任感和归属感。

41. 联合国志愿人员方案关于 2008 年评价工作的管理部门回应已编写完毕并上传至评价资源中心，对后续行动进行了跟踪，系统中也体现了这一点。为支持应用管理部门回应系统，评价股为各项讨论提供了便利，以提高对该系统在问责和借鉴评价工作方面的重要性的认识。该系统为编写管理部门回应提供了广泛的同行支持，联合国志愿人员方案今后的规划、监测和评价培训中将纳入这一主题。

42. 联合国志愿人员方案是联合国评价小组关于人权和两性平等、国家一级的评价以及名册评价的各工作队的积极成员。联合国志愿人员方案自愿对其评价职能开展了联合国评价小组同行审查。这一进程有望产生有助于联合国志愿人员方案加强自身评价进程的信息。

D. 方案股

覆盖范围

43. 由方案股委托进行的评价，又称“分散评价”，对于支持改进方案拟订情况的知情决策而言至关重要。此外，这些评价也为管理部门的监测工作以及对利益攸关方和合作伙伴的问责提供了支持。这项评价政策要求方案股为这一方案期间制定评价计划，授权外部评价人规划成果评价、项目评价和其他类型的评价，编写管理部门回应，并在评价资源中心予以公布。方案股管理人负责确保评价工作的质量，并提供委托评价所需的适当资源。

44. 在本报告所述期间，50%的国家办事处开展了 158 次评价，其中包括 24 次成果评价。尽管评价总数出现下降，但至少开展过一次评价的国家办事处所占比例未发生较大变化。这就意味着，平均而言，每个国家开展的评价次数只是稍有减少。在至少开展过 1 次评价的国家中，各国的平均评价次数在各区域之间相当一致——介于 1.8-2.6 之间。去年的数字介于 2.2-4.6 之间，与今年明显不同。与前几年相同的是，大部分已完成的报告均为项目评价报告（77%）。

45. 在亚洲和太平洋区域以及欧洲和独立国家联合体（独联体）区域，开展过 1 次评价的国家办事处的比例最高（分别为 71%和 72%），这与去年的情况大致相仿。欧洲和独立国家联合体区域开展过成果评价的国家比例最高（24%），而非洲区域的比例最低（9%）。

表 1. 国家办事处已完成的评价

2007 年 全球 (137)		2008 年 全球 (137)	区域（具有评价计划的国家数量）				
			非洲 (45)	阿拉伯国家 (18)	亚洲和太平洋 (24)	欧洲和独联体 (25)	拉丁美洲和加勒比 (25)
183	2008 年评价总次数	158	33	18	44	43	20
137	项目评价次数	121	23	16	33	36	13
28	成果评价次数	24	5	2	8	6	3
18	其他评价次数	13	5	0	3	1	4
15%	成果评价在所有评价中所占百分比	15%	15%	11%	18%	14%	15%
51%	至少开展了任何一种评价的国家所占百分比	50%	33%	39%	71%	72%	44%
14%	至少开展了一次成果评价的国家所占百分比	15%	9%	11%	21%	24%	12%
1.3	每个国家的平均评价次数	1.2	0.7	1.0	1.8	1.7	0.8
2.6	每个至少开展了任何一种评价的国家的平均评价次数	2.3	2.2	2.6	2.6	2.4	1.8

46. 各项评价的专题覆盖范围与去年类似。评价次数最多的工作领域⁵为能源与环境领域，占 34%，而专题领域仅占 2007 年方案支出的 9%。⁶这种情况可能是因为全球环境基金（全环基金）对资助项目提出严格的评价要求造成的。在所有的

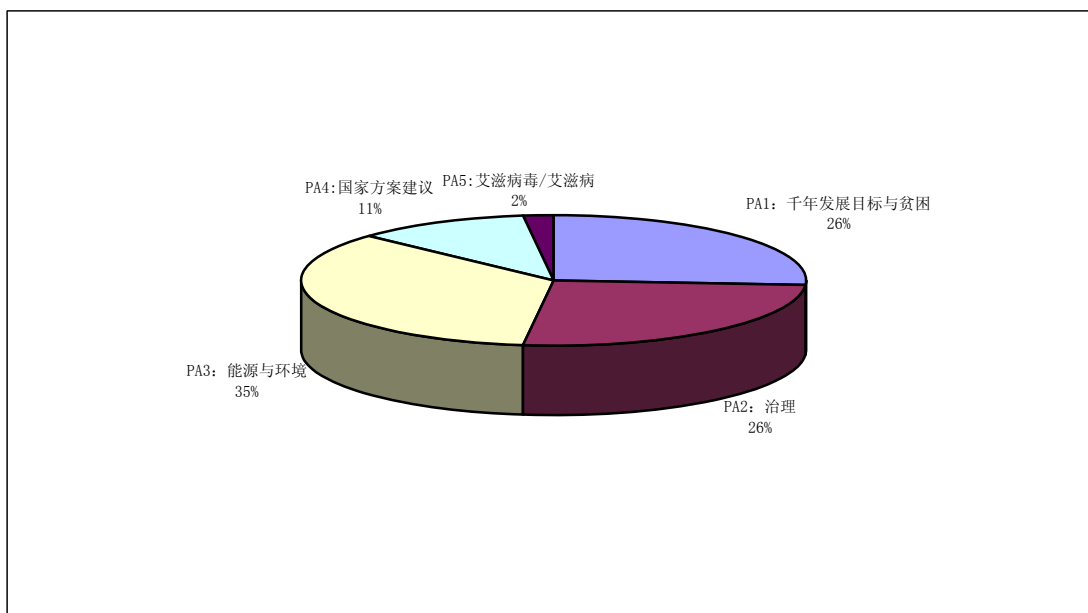
⁵ 由于在多年筹资框架（MYFF）期间所有的评价活动都对开发署的干预措施进行了评估，因此为便于报告，参考了多年筹资框架的工作领域，这与战略计划的成果领域不同。

⁶ 《开发署 2008 年年度报告》（<http://www.undp.org/publications/annualreport2008>）。

评价活动中，有 52%是在《千年发展目标》与贫困及民主治理相结合的领域中开展的，共占方案支出的约 72%。

47. 区域局以及政策和实践局也必须对各自的政策和方案开展评价。在本报告所述期间，它们仅开展了 2 次评价：发展政策局完成了对民主治理专题信托基金的评价，阿拉伯国家区域局对其区域合作框架开展了 1 次成果评价。

图 1：2008 年 145 次各类评价的专题范围⁷



遵守情况

48. 评价政策授权方案股在自己的方案周期内开展各自评价计划中确定的成果评价。在 15 个 2008 年结束的国家方案中，有 9 个国家方案得到了全面遵守（见表 2）。

⁷ 仅包括具有评价资源中心列明的多年筹资框架工作领域之一的评价活动工作领域。

表 2：遵守成果评价的情况

区域	有遵守义务的国家数量	遵守*	部分遵守**	不遵守***
非洲	8	5	1	2
阿拉伯国家	3	2	1	-
亚洲和太平洋	1	1	-	-
欧洲和独联体	0	-	-	-
拉丁美洲和加勒比	3	1	1	1
全球	15	9	3	3

* 完成了所有已规划的成果评价。

** 至少完成了一次评价，但没有完成所有已规划的成果评价。

*** 没有进行任何已规划的成果评价。

49. 审计和调查处的审计结果突出表明规划和执行成果评价方面的遵守情况较差。在 2008 年提出的共 701 条审计建议中，有 13 条直接针对缺乏已规划的成果评价或延迟执行的情况。2008 年，随着发展成果评估的覆盖范围不断拓宽，缺乏成果评价的情况对评价小组提出了挑战。评价小组不得不把有限的时间用于收集基本的评价材料，而不是专注于进行分析。

50. 近期的评价规划工作有了一些改进。为了提供足够的问责评价证据，建议方案股在每个工作领域至少规划和执行 1 次成果评价。这就意味着，从整体而言，每个国家方案应规划 2 至 4 次成果评价。⁸ 2007 年批准的已规划的国家方案成果评价次数介于 0-15 次之间。2008 年，已规划的成果评价次数更加一致，更贴近事实，介于 2-6 次之间。这表明提高了对成果评价遵守和监测活动的理解。

质量

51. 评价办公室对 2008 年完成的 18 次成果评价进行了质量评估。达到“满意及以上”级别的报告所占比例从 34%增至 39%，这表明与去年相比稍有进步。然而，质量参差不齐仍是一个主要问题，因为 61%的报告仍未达到“满意”级别。调查结果表明，当评估人在报告中参考《评估办公室指南》时，报告的质量较高（见表 3）。

⁸ 大部分国家方案在 2-4 个工作或成果领域开展。

表 3：按质量标准分列的成果评价等级摘要 9

	等级		质量标准						
2007 年 总数		2008 年 总数	报告 结构和表述	评价 主题和内容	评价 目标和标准	评价 设计和方法	调查结果	结论	建议和教 训
3 (17%)	非常满意	2 (11%)	3 (17%)	1 (6%)	2 (11%)	2 (11%)	5 (28%)	1 (6%)	3 (17%)
1 (6%)	满意	3 (17%)	2 (11%)	4 (22%)	2 (11%)	3 (17%)	1 (6%)	4 (22%)	5 (28%)
2 (11%)	较为满意	2 (11%)	--	1 (6%)	--	--	--	1 (6%)	--
2 (11%)	不太满意	1 (6%)	6 (33%)	7 (39%)	5 (28%)	2 (11%)	3 (17%)	5 (28%)	6 (33%)
8 (44%)	不满意	8 (44%)	7 (39%)	5 (28%)	3 (17%)	2 (11%)	7 (39%)	5 (28%)	3 (17%)
2 (11%)	极不满意	2 (11%)	--	--	6 (33%)	9 (50%)	2 (11%)	2 (11%)	1 (6%)

52. 与去年相似，通过分析每项评估标准对应的等级，表明到目前为止，评价报告在“评价设计和方法”方面最为不足，其次是在“调查结果”方面。在“评价设计和方法”这一标准上，有 72% 的报告的等级为“不太满意”。尽管大多数报告确定了数据来源，但似乎并没有建立在完善的评价方法和设计上，在回答评价问题时也缺乏明确的理由。

53. 在开发署评价办公室的协助下，瑞典发展评价署对开发署各项分散评价的质量开展了一项独立研究，其中呼吁开展机构努力，以改进评价工作的质量不均衡问题。需更多地重视规划工作，因为劣质的方案设计会导致评价工作软弱无力。报告的调查结果在企业承诺发展评价能力方面引发了关于“评价网”的讨论。作为更进一步的措施，《订正的规划、监测和评价手册》中纳入了一个全面介绍评价方法和设计的章节，以指导开发署的工作人员和评价人确保专业的评价标准。

E. 能力

开发署

54. 评价办公室从各个国家办事处收集关于现有能力的信息。数据证实了他们在监测与评价职能专业化方面的努力。一年来，专门的监测和评价专家的人数由 38 人增至 46 人。尤其是，非洲区域出现了显著增长，由 8 人增至 13 人。只有 78% 的专家支持仅对国家方案进行监测和评价。大部分专家致力于不同层面的工作：52% 在联合国发展援助框架（联发援框架）层面上为联合国国家工作队提供支持；

⁹ 本表列明了按等级分列的符合各项质量标准的报告的数量和比例。

50%在特定主题领域、项目或信托基金的背景下开展工作。¹⁰专门的监测和评价股的数量也略有增长，从 30 个增至 31 个（见表 4）。

表 4：2008 年各个国家办事处的评价能力

区域和国家的数量	2008 年 全球 (137)	非洲 (45)	阿拉伯国家 (18)	亚洲和太平洋 (24)	欧洲和独 联体 (25)	拉丁美洲 (25)
专门的监测和评价专家在各区域所有国家办事处中的人数和所占比例	46 (34%)	13 (29%)	6 (33%)	13 (54%)	2 (8%)	12 (48%)
监测和评价股在各区域所有国家办事处的数目和所占比例	31 (23%)	8 (18%)	2 (11%)	11 (46%)	0 (0%)	10 (40%)

M&E = 监测和评价

55. 随着区域一级对监测和评价咨询服务的需求增多，区域局已承诺在各区域设立专门的专家员额。然而，目前的进展并不均衡。非洲区域局是唯一拥有这一能力的机构，目前有 3 名顾问（1 名在总部，2 名在达喀尔和约翰内斯堡的区域中心）。拉丁美洲和加勒比区域局和阿拉伯国家区域局正为各自的区域办事处征聘 1 名评价专家。欧洲和独联体国家区域局已取消了顾问员额。

56. 预防危机和复原局在日内瓦设有 1 名专家，由另 1 名干事为其提供支持，后者把 25%的时间用于在纽约开展监测和评价工作。两人一起制定和实施该局的监测和评价系统，并支持各个国家办事处在危机预防和复原方案中加强监测和评价。根据其评价计划，该局正对 2002 年以来开发署的减少灾害风险干预措施在一些高风险国家的影响进行评价。2008 年，其他政策和实践局，包括发展政策局、管理局、伙伴关系局等，在各自的方案或战略中都没有专门的监测和评价专家，也没有评价计划。

57. 2008 年中期关于“评价网”的讨论突出强调了国家一级对评价能力的诸多约束。尽管各个国家办事处的监测和评价专业化趋势逐渐上升，但主要还是由协调中心来安排评价工作。在设有专门的监测和评价专家的办事处里，专家在办事处的地位以及报告路线会影响其有效性；那些在汇报结构上远离高级管理部门的专家对决策的影响极为有限。聘用监测和评价人员时往往使用非核心供资短期合同，因而很难建立和持续保持一个有效的监测和评价体系。

58. 一些国家办事处已采取措施提高监测和评价工作的质量和能力。依据评价办公室提供的指导，尼泊尔和秘鲁的监测和评价干事为从事规划、监测和评价工作的方案工作人员开发了一个资源包，并为各自的工作人员和政府合作伙伴开办了培训课程。

¹⁰ 这些职责并不总是相互排斥。

59. 向国家一级监测和评价工作分配的资源通常不足，这是因为，方案设计中不一定会包含适当的评价活动预算。保证为成果评价供资仍是一个难题，共同开展项目的捐助方或政府伙伴经常要求或赞成开展项目评价而不是成果评价。要确保评价所需的适当的机构能力和财政资源，各个国家办事处的高级管理部门就必须具备较强的领导能力。此外，国家办事处的核心职能应当是监测和评价，并应相应地获得资源与支持。

为国家评价能力提供支持

60. 开发署的各个股已经对方案政府和合作伙伴关于支持发展国家评价能力的需求做出了回应。开发署的各个国家办事处通过方案干预提供了支持，其重点是建立国家监测和评价体系；开发监测和报告技能以跟踪和评价该国在实现《千年发展目标》和国家发展目标方面的进展；以及创建关于发展干预措施的监测和评价体系。在喀麦隆和乌干达，开发署支持国家努力监测和报告《千年发展目标》。开发署的塔吉克斯坦办事处正在支持经济部和统计署发展能力，以监测和评价《减贫战略文件》和国家发展战略中反映的各项指标。

61. 评价办公室为开展了评价工作的国家发展评价文化与能力做出了贡献。例如，在贝宁开展“发展成果评估工作利益攸关方讲习班”期间，与负责规划、制定和评价公共政策与方案的国务大臣开展的一次讨论导致贝宁政府请求开发署为发展评价能力提供支持。作为第一步措施，设于达喀尔的区域评价顾问正牵头开展一项需求评估。评价办公室还通过正在与联合国评价小组开展的活动做出了贡献。2008年，约有10个政府和国家合作伙伴从联合国评价小组的培训方案中受益。对“一体行动”卢旺达试点的可评价性研究促成了在卢旺达政府建立了第一个独立的评价实体。

62. 尽管正在开展各种举措，但不断增长的需求仍超出了现有的支助和开发署的能力。为更好地协调各种努力并最好地发挥其影响，开发署有必要制定一项支持发展国家评价能力的战略。

F. 评估的使用与后续工作

63. 评价办公室与评价人紧密合作，努力提高评价建议的清晰度和战略性质。“发展成果评估工作利益攸关方讲习班”吸引了合作伙伴参与广泛讨论主要的评估结果、结论和建议，已经证明这种讲习班对于确保把评价工作有效地运用于设计新的国家方案方面极为有益。利益攸关方讲习班上的讨论还有助于评价办公室提出更加尖锐的建议，使其更具战略性和可操作性。利用讲习班吸引利益攸关方参加，并改善建议的质量，这都有利于澄清管理部门对发展成果评估的回应。

64. 开发署已经采用了管理部门回应制度，以透明化的方式跟踪该组织应用评价工作的情况。评价政策要求开发署的管理部门编写对每一次评价工作的管理部门

回应并将其上传到评价资源中心。编写并向执行局呈交关于独立评价工作的管理部门回应已成为一项制度。然而，对 2006 年以来完成的独立评价工作，仅有 54% 的管理部门回应进入评价资源中心或可以在该中心跟踪查到。开发署有必要更加关注问责制的改进和后续评价工作的透明化。

65. 在分散评价方面，编写并公布管理部门回应的做法尚未制度化。在 2006 年以来完成的分散评价中，仅有 31% 向评价资源中心呈交了管理部门回应。报告所述期间完成的所有评价中，只有 35% 获得了管理部门回应。尽管去年有所改善（21%），但距离完全遵守这一目标仍相差很远。

二. 独立评价工作的主要成果和经验教训¹¹

A. 政策建议的实质性能力

66. 近年来，开发署采取了协调一致的行动，以加强内部能力，提供专业的政策建议。虽然在这一领域取得了重大进展，如一些发展成果评估所示，例如关于阿根廷的发展成果评估，但评价表明还有很多工作要做。

67. 9 次独立评价¹²得出结论认为，开发署通常被视为是一个有效的捐助资金管理人而不是实质性的发展伙伴。对于大部分预算来自方案国的国家，例如阿根廷、博茨瓦纳和危地马拉，评价指出开发署倾向于向政府提供服务，这造成了一种风险，也即偏离提供实质性政策建议的做法而宁愿通过提供支助服务来立即赚取金融收益。在阿富汗，考虑到运作大型方案所需的报告和行政要求，开发署正在征聘工作人员管理资金而不是提供政策建议。

68. 评价发现，开发署应将重点放在建设其工作人员的实质性能力方面，并征聘更多的政策专家与资金管理人共同工作。¹³在阿富汗、波斯尼亚和黑塞哥维那以及刚果共和国开展的发展成果评估表明，为了在国家一级建设实质性能力、为决策者开展政策调查并加强与政府和其他伙伴的战略伙伴关系，必须提供更多资源。

¹¹ 见附件 1 “评价办公室以及相关基金和方案机构的评价股已完成的独立评价清单”。《2007 年年度报告》中载有关于贝宁和刚果共和国的发展成果评估，但从这些评价中吸取的经验教训载于本报告。

¹² 关于阿富汗、阿根廷、巴巴多斯和东加勒比国家组织（OECS）、贝宁、波斯尼亚和黑塞哥维那、刚果共和国、危地马拉、塔吉克斯坦的评价，以及环境和能源评价。

¹³ 关于阿富汗、阿根廷、波斯尼亚和黑塞哥维那、博茨瓦纳、菲律宾、乌兹别克斯坦的评价，以及环境和能源评价。

B. 国家能力和可持续自有权

69. 开发署的核心任务是加强国家解决人类发展挑战的能力。然而，多种多样的方案模式和做法可对国家伙伴的能力发展产生负面影响。

70. 当政府的能力不够强大时，开发署可能必须在短期内参与管理公共资金。然而，开发署并非总是将国家伙伴和机构的能力发展问题纳入这些方案。由于过度依赖开发署的直接管理，尤其是在冲突后国家，例如波斯尼亚和黑塞哥维那、刚果共和国以及塔吉克斯坦，一旦开发署撤销支助，就会妨碍发展所需的能力，不能保持各项举措的效益。特别是，开发署使用商业服务的做法，包括在采购和征聘方面，存在取代而不是加强国家机构执行能力的风险。

71. 几乎每次发展成果评估都提到在规划各项举措时没有制定撤离战略。如果没有撤离移交战略、界定明确的国家伙伴能力发展目标以及强有力的国家所有权，包括高级别的高层政府承诺，各项成果就不会持续很久。在刚果共和国，评价发现，当地机构需要具备接管对各方案的长期供资和人力资源责任的意愿和能力。在布基纳法索，评价发现，开发署/资发基金联合项目的核心是建设当地的能力，¹⁴在该项目中，社区的捐助平均占当地发展项目资金的 10.5%。

72. 多项发展成果评估¹⁵表明，需改进战略，以更有效地与民间社会组织合作，加强国家的自有权和成果的可持续性。在菲律宾等国家，开发署与民间社会组织在方案设计、执行和政策宣传方面携手合作，大幅提高了国家对开发署方案的自有权。尽管一些方案与民间社会组织合作良好，¹⁶但仍需开发更多的与私立部门合作的领域。

73. 通过评价妇发基金关于在东南亚促进执行《消除对妇女一切形式歧视公约》的方案，发现妇发基金对报告进程的支持尤为有效。它使政府和民间社会伙伴能够积极参与编写报告，保证了对这一进程的自有权，并取得了成果。

C. 战略重点与一致性

74. 尽管大部分方案符合开发署的授权和国家的优先事项，然而，它们往往范围太广或者缺乏重点。尤其是，第三次全球合作框架评价发现，如果把重点放在高需求领域和开发署具有相对优势的领域，可以强化它对发展成果的贡献。

75. 开发署方案在很大程度上与各国面临的挑战有关。然而，评价发现，在广泛的战略目标和成果以及相关方案和投入之间往往缺乏联系。开发署的规定目标与其在国家层面上的实施之间存在实质性差距。

¹⁴ 纳门滕加省农牧资源支持项目。

¹⁵ 阿富汗、巴巴多斯和东加勒比国家组织（OECS）、博茨瓦纳和刚果共和国。

¹⁶ 危地马拉、塔吉克斯坦和阿拉伯国家区域合作框架。

76. 环境与能源评价发现，全球环境基金的供资可用性对开发署确定优先事项有重大影响。开发署很好地管理着全球环境基金组合，但除了全球环境基金的全球重点之外，各个方案国还有当地重点。在塔吉克斯坦和刚果共和国，捐助重点影响了开发署在当地资源调集方面的方案拟订工作，导致方案分散，缺乏重点。此外，非核心供资难以预测，使得难以进行规划，菲律宾和阿富汗的评估工作就是这方面的例子。

D. 方案拟订的优势与劣势

77. **中立性。**开发署被视为是中立的发展伙伴，其工作范围几乎辐射全球，这使其在促进政策对话、知识共享和协调发展成果等方面占有战略优势。对发展成果评估进行的 7 项评价以及 1 项区域合作框架评估¹⁷突出强调了这一事实，即这一可感觉到的中立性使开发署能够在人权和宪法改革等敏感领域发起对话。在阿根廷，开发署利用其声望对总统发起的高级别论坛“对话阿根廷”进行协调，该论坛旨在制定战略并达成共识解决 2001-2002 年的金融、社会与政治危机。目前正在宣传“阿拉伯国家区域合作框架”中增值性最高的组成部分之一；这一组成部分让决策者和民间社会了解了治理和两性平等领域的一些紧迫的敏感问题。此外，许多评价都把开发署的中立性视为该署之所以能够在阿富汗和塔吉克斯坦等政府从冲突中恢复时为它们充当可靠的重要融资渠道的重要原因。

78. **性别平等主流化。**开发署在乌兹别克斯坦的经验是与各级政府合作制定两性平等战略的成功典范。由此，越来越多的妇女开始在开发署项目的规划与执行过程中担任管理人和专家。然而，7 项评价¹⁸突出强调，两性平等问题并未适当成为方案工作的主流。第三次全球合作框架评估指出，尽管已经加强了内部能力和对两性平等问题的理解，但尚未在框架内的所有实质工作领域实现性别平等主流化。菲律宾的发展成果评估发现，两性平等原则在开发署方案中的应用情况不均衡，往往是在方案一开始就没有进行分析，个人工作计划中也没有这方面的内容。正如 2006 年的开发署性别平等主流化工作评价所指出的，多项评价表明，各个国家办事处高层仍然缺乏高级别两性平等问题专家。这种状况妨碍了在方案拟订过程中考虑两性平等问题。

79. **联合国的协调。**开发署有权通过驻地协调员办公室协调联合国各组织在当地的各项工作，开发署通常被视作是能够有效召集当地有分歧的伙伴的召集人。尽管开发署正通过联合国发展援助框架有效地促进联合规划工作，然而，发展成果

¹⁷ 阿富汗、阿根廷、巴巴多斯和东加勒比国家组织（OECS）、贝宁、危地马拉、菲律宾、塔吉克斯坦，阿拉伯国家区域合作框架。

¹⁸ 巴巴多斯、贝宁、博茨瓦纳、危地马拉、菲律宾、塔吉克斯坦，第三个全球合作框架。

评估¹⁹表明，几乎没有通过联合方案协作执行的实例。关于波斯尼亚和黑塞哥维那的发展成果评估发现，开发署应请求参与了联合国各机构的项目制定过程，但在执行过程中却几乎没有合作的空间。在关于危地马拉的发展成果评估中，联合国各机构之间经常出现工作重复，因此有必要在业务层面上开展更多的规划。

80. **基于成果的管理原则。**许多评价都有一个共同的主题，即基于成果的管理原则在方案制定、监测和评价中的运用不均衡。²⁰国家方案中的大多数个体项目都成功地取得了既定成果，但不太清楚开发署是否为实现这些成果做出了有效贡献。共同的问题包括：缺乏建立可靠的评价所需的清晰标准、指标或基准；对成果的定义太过宽泛或夸大其辞；成果、产出和指标之间的界定模糊；以及在常规监测中制定、检测和使用指标的情况不佳。刚果共和国的发展成果评估发现，成果往往是指产出，而指标是从质量上进行界定的，致使监测工作极为主观化。必须在规划阶段清楚地阐明成果、产出和指标，以有效地取得可持续的成果，使基于成果的管理方式得到最充分的利用。

三. 评价办公室 2009-2010 年工作方案

81. 评价办公室工作方案符合开发署的战略计划。将开展各项评价工作，以评估执行局批准的全球方案、区域方案、国家方案和其他方案所规定的产出，以及在战略计划中的各项目标、关键成果领域、交叉问题以及组织原则中列明的产出。

82. 评价范围将具有选择性和战略性。每年至少将开展 15 次发展成果评估，并将对全球、区域和南非方案进行评价。每项目标和关键成果领域均将被覆盖，并将关注与方案国的强烈需求有关的产出以及开发署的高水平投资。这将提供一个重要的组织学习机会，并解决至关重要的新出现的全球发展问题。

83. 评价将提供充分的评价证据，以便在战略计划周期结束时对其进行总结评估。2009-2010 年工作方案如下：

(a) 已批准的工作方案

- (一) 16 次发展成果评估；
- (二) 1 次区域合作框架评价：欧洲和独立国家联合体第三次区域合作框架评价（2006-2010 年）；
- (三) 对开发署为加强《千年发展目标》等发展成果方面的国家管理能力所做贡献的评价；

¹⁹ 阿富汗、巴巴多斯和东加勒比国家组织、波斯尼亚和黑塞哥维那、博茨瓦纳、刚果、危地马拉、塔吉克斯坦和乌兹别克斯坦。

²⁰ 巴巴多斯和东加勒比国家组织、波斯尼亚和黑塞哥维那、刚果、菲律宾、塔吉克斯坦，联合国妇女发展基金关于在东南亚促进执行《消除对妇女一切形式歧视公约》情况的评估。

- (四) 对开发署为权力下放和地方治理所做贡献的评价；
- (五) 对开发署为“促进减贫的环境管理：贫穷与环境之间的联系”所做贡献的评价；
- (六) 对开发署为受自然灾害影响的国家复原所做贡献的评价；
- (七) 关于开发署/联合国工业发展组织合作协议的联合评价；
- (八) 对开发署在有效促进全环基金用于实现发展成果方面的评价；
- (九) 对开发署的支助在加强选举制度和进程方面的作用与贡献的评价；
- (十) 对开发署的区域化程序的评价。

(b) 拟议工作方案：

- (一) 15 次发展成果评估；
- (二) 3 次区域合作框架评价，分别为非洲、亚洲及太平洋、拉丁美洲和加勒比；
- (三) 对开发署在减贫领域所做贡献的评价；
- (四) 对 2008-2011 年战略计划的有效性的评价。

(c) 通过以下举措支持评价文化：

- (一) 通过关于订正评价手册的区域讲习班和培训，建设开发署工作人员和国家伙伴的评价能力；
- (二) 管理和改进评价资源中心；
- (三) 管理“评价网”，开发评价知识产品；以及
- (四) 主持和管理联合国评价小组秘书处，并为评价小组的工作方案做出贡献。