



第六十六届会议

暂定项目表* 议程项目 136 和 143

方案规划

关于内部监督事务厅活动的报告

加强评价的作用并执行关于方案设计、交付和政策指示的评价结果

内部监督事务厅的报告

“来自评价的信息有各种重要用途，但是评价还没有成为一个获全面接受的管理功能”

摘要

本报告是 1988 年以来，根据《方案规划、预算内方案部分、执行情况监测和评价方法条例和细则》条例 7.4，通过方案和协调委员会，每两年向大会提出的一系列研究报告的第十一份。其目的是：

- 审查联合国秘书处评价活动的现状；^a
- 概述 2008-2009 两年期完成的评价报告的主要结果、结论和建议；
- 评价 2008-2009 年内部监督事务厅(监督厅)从各方案收到的自我评价报告的质量；
- 研究如何在联合国秘书处执行关于方案设计、交付和政策指示的评价；

* A/66/50。

^a 就本研究报告来说，“联合国秘书处”指业经确定在监督厅任务规定范围内的 31 个方案。这 31 个方案的全部清单见附件一。



- 提出监督厅 2012-2013 年的评价工作计划。

在 2008-2009 两年期，评价的重点仍然是方案实施、产出和成果方面的问题，主要依循业绩评价功效标准。本报告说明评价如何有助于改进方案活动，使组织更有效率。

所有方案都从事评价活动。总的来说，自我评价报告的质量令人满意，但各方案的报告质量参差不齐。2010-2011 年用于评价的资源与前一个两年期相比有所下降，仍低于建议的基准。

评价政策的存在对评价能力和质量产生有利的影响，不过尽管在某些领域取得进展，评价还没有成为一个完全获接受的管理功能。2010 年高级主管同秘书长的契约都删除了自我评价方面的预期成果和相关目标。

监督厅在本报告提出了中央评价职能工作计划。监督厅还报告跟进收自方案和协调委员会第四十九届会议的各项建议的执行情况，并建议提高评价报告质量的行动，同时提出一个问题供委员会审议。

目录

	页次
一. 导言	4
二. 方法	4
三. 研究成果	5
A. 2008-2009 年评价范围和重点	5
B. 2008-2009 年对本组织各优先领域的评价结果、结论和建议 提供了一个改进路线图	8
C. 内部评价报告总体质量令人满意，但方案之间和之内差别很大	10
D. 评价为多种重要的用途服务	11
E. 秘书处方案一级评价做法和能力概况	13
F. 内部监督事务厅的评价活动和工作计划	22
四. 结论	27
五. 加强评价报告质量的行动	28
六. 供方案和协调委员会审议的问题	28
附件	
一. 两年度研究包括的部门/办事处名单和 2008-2009 年评价报告的数目	30
二. 质量评价成果	31
三. 秘书处各实体对报告草稿的评论	33

一. 引言

1. 本报告是 1988 年以来，根据《方案规划、预算内方案部分、执行情况监测和评价方法条例和细则》(ST/SGB/2000/8)，通过方案和协调委员会，每两年向大会提出的一系列研究报告的第十一份。秘书处各部门审查了报告的草稿，适当时报告纳入了它们的意见。

2. 这项研究有五个目标：

- (a) 审查联合国秘书处评价活动的现状；¹
- (b) 概述 2008-2009 两年期完成的评价报告的主要结果、结论和建议；
- (c) 按照方案和协调委员会第四十九届会议的建议(A/64/16, 第 42 段)，评价 2008-2009 年内部监督事务厅(监督厅)从各方案收到的自我评价报告的质量；
- (d) 研究如何在联合国秘书处执行关于方案设计、交付和政策指示的评价；
- (e) 提出监督厅 2012-2013 年的评价工作计划。

二. 方法

3. 为了确定联合国秘书处 2008-2009 年两年期的评价活动的范围和质量，首先从综合监测和文件信息系统抽出一些内部评价报告和外部评价报告，之后监督厅再请各方案提供在两年期内最后完成的其他评价报告的副本。监督厅审查了收到的所有 279 份报告，以确定它们是否符合监督厅对评价的业务定义。

4. 监督厅定义评价为根据某一方案的任务或目标来确定该方案的任何业绩方案的相关性、效率、实效、影响和/或可持续能力的一种有系统、谨慎而尽量客观的方案。评价可以用来促进问责、学习和/或决策。一份评价报告应当说明评价的目标、所用方法、循证评价结果、结论和建议(以适用者为准)。在 279 份报告中，有 155 份符合这些标准。附件一按方案列出这些评价报告。监督厅进行的评价不列入这份清单；下文第三节 F 将单独报告监督厅进行的评价。

5. 在 155 份评价报告中，抽出了 71 份，作为有目的的抽样，以研究报告的各个属性(即范围、目标、各类结果、结论和建议)。该抽样最多选用每个方案的五份报告。如一个方案产生的报告超过五份，则根据下列因素选取样本：

- (a) 两年期的每一年抽取差不多一半；
- (b) 平衡兼顾期中与最后评价；

¹ 就本研究报告而言，“联合国秘书处”指业经确定在监督厅任务规定范围内的 31 项不同的方案。这 31 项方案的全部清单见附件一。

- (c) 搭配项目、次级方案，方案和其他各级的评价；
- (d) 均衡搭配各种专题；
- (e) 覆盖广泛的地理范围。

6. 另一个抽样是从属性样本的 71 份报告中抽出 45 份，以便使用 27 个标准去评价报告的质量。² 每个方案有 2 至 4 个评价报告接受评价，份数多少看方案完成的评价数目而定。上面所列标准都一一适用；不过，质量评价抽样不包括少于 11 页的报告(因为初步审阅报告发现这种报告的数量不够，不足作为审查材料)。在此基础上，管理部、政治事务部、法律事务部、非洲问题特别顾问办公室和联合国日内瓦办事处的所有报告都被排除在外。为了使质量评价尽可能公正，监督厅与一个独立评价专家订约，请他负责审查和评价报告。

7. 监督厅还在 71 份报告中抽出 42 份非随机抽样，参照本组织八个战略优先领域，审查各项结果、结论和建议。³

8. 为了进一步查明评价做法，监督厅代表对研究范围的 31 个方案的所有协调人进行了网上调查。调查在 2010 年 11 月 19 日至 2011 年 1 月 6 日之间进行，有百分之百的回应率。监督厅还分析了二级数据来源，包括两年期评价计划、监督厅在七个秘书处部门的监测和评价做法报告、以及监督厅评估联合国秘书处评价能力和需求后提出的各项建议的实施情况(检查和评价司第 IED-2006-006 号报告)。

9. 本研究有一个主要限制因素，就是监督厅可能没有收到所有在 2008-2009 两年期内最后完成的评价报告。虽然监督厅在综合监测和文件信息系统列出了秘书处的评价报告，并向各协调人提供关于提交评价报告的指引，但监督厅可能没有获得所有报告。此外，非随机性的抽样过程和相对较小的样本规模限制了将抽样结果推而及于整个秘书处的程度范围。一有可能，即采用三角测量法验证结果。

三. 研究成果

A. 2008-2009 年评价范围和重点

10. 两年期内，秘书处所有方案均进行评价活动。有二十二个协调人报告它们的方案已完成内部和外部评价。五个方案只进行了外部评价，这些评价都是由联检

² 这些标准是 2009 年制订来评价监督厅检查和评价司评价报告质量的 42 个综合标准的一个子集。制订这些标准的参考来源包括联合国评价小组在联合国内进行评价的标准、监督厅检查和评价司内部标准、世界银行和美国国际开发署。

³ 这八个领域列于大会第 61/254 号决议第 9 段。抽样不包括大会和会议管理部、管理部、新闻部和联合国日内瓦办事处的评价报告，因为这些部门支持所有八个领域的工作。

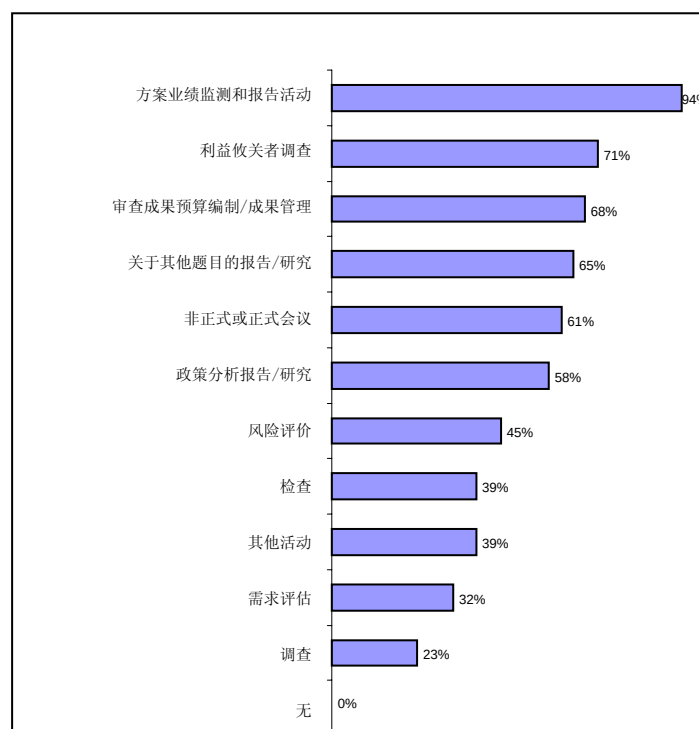
组和监督厅等外部实体负责管理和进行的。⁴ 另五个方案只进行了由方案负责管理和/或进行的内部评价。⁵

11. 与前一个两年期相比，这些评价更为分散于秘书处各个部门。以前，有五个方案占了所有评价报告的 65%(见 A/64/63，第 45 段)，而在 2008-2009 年，五大方案只占了 46%。

12. 图一显示了除外部和内部评价以外其他各种各样的评价活动。监督厅定义“其他评价活动”为：(除上文第 4 段定义的评价以外的)任何活动，目的是与收集数据，以评价或确定方案或组织的业绩，并作出方案规划和实施、问责和/或学习方面的决策。

图一

2008-2009 年各方案进行的“其他评价活动”



注：“其他活动”包括外部联合国评价小组同行审议、内部审议、吸取经验教训活动、自我评价举措等。

来源：监督厅对各方案协调人的调查。

⁴ 管理部、安全和安保部、秘书长办公厅、最不发达国家、内陆发展中国家和小岛屿发展中国家高级代表办事处和联合国内罗毕办事处。

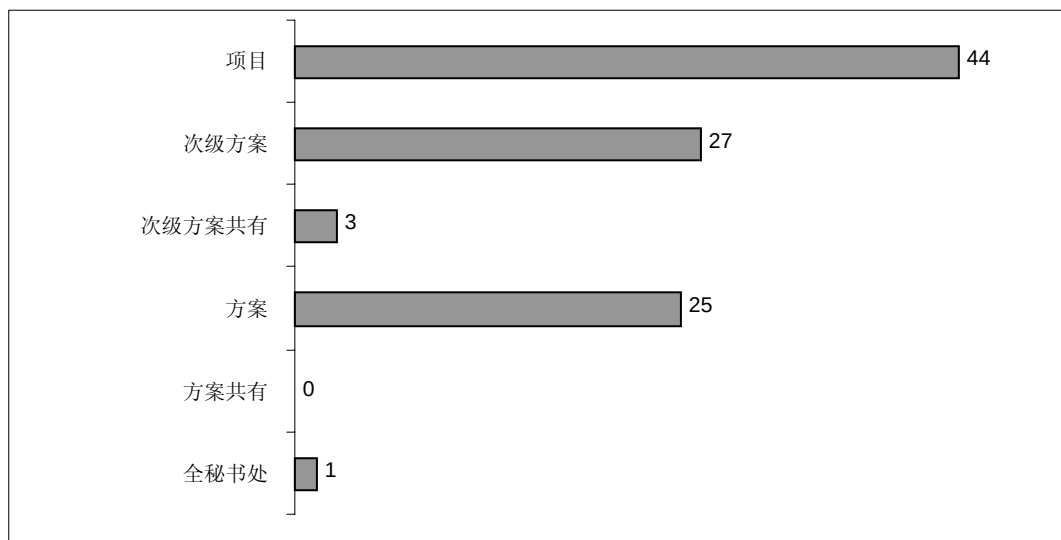
⁵ 西亚经济社会委员会、维持和平行动部/外勤支助部、人道主义事务协调厅、法律事务厅和联合国人类住区规划署。

2008-2009 年侧重方案实施、产出和成果的评价，许多在项目一级进行

13. 经审查的 71 个评价中，将近四分之三处理方案或部门之下各级的问题(图二)。最常见的评价是在项目一级(44%)。⁶ 在上一个两年期，经审查的 43%的报告其评价对象都在项目一级(见A/64/63，第 7 段)。

图二

2008-2009 年评价范围(百分比)



来源：监督厅对 2008-2009 年评价报告属性的分析。

14. 经审查的 71 个评价报告中，将近一半(44%)以区域为重点(包括同一区域几个国家)，37%侧重于全球性问题，20%集中于个别国家的问题。

15. 经审查的 71 个评价报告中，96%集中在方案产出，方案实施情况和方案成果各有 94%。在上一个两年期，经审查的评价报告只有四分之一的评价针对结果或影响(同上，第 8 段)。在 2008-2009 年，经审查的评价报告中有 85%以方案设计为重点，而 2006-2007 年为 30%(同上)，38%则针对政策指令。然而，由于分析是质量性的，不同个人负责评估不同两年期的报告的重点，因此一些差异可能反映的是方法工具方面的变化。

16. 关于用来评价 2008-2009 年业绩的标准，功效标准指导了所有经审查的评价，比上一个两年期增加了 5 个百分点。此外，在 2008-2009 年，93%用影响作为评价标准。⁷ 使用效率标准的评价从以前的 73%(A/64/63，第 9 段)增至 2008-2009

⁶ 这符合方案规划、预算内方案部分、执行情况监测和评价方法的细则和条例第 107.3(b)条。

⁷ 这符合方案规划、预算内方案部分、执行情况监测和评价方法的细则和条例第 107.3(c)条。

年的 99%。89%使用相关性作为评价标准，比 2006-2007 年的 59%增加很多(同上)。2008-2009 年经审查的评价有 85%使用可持续发展能力标准来评估业绩。

17. 一般来说，在 2008-2009 年，方案一级的评价是在政府间机构建立的框架内进行的。但是，评价报告看来并没有支持实施关于性别平等和人权的全秘书处任务规定。71 份评价报告中只有 9 份(13%)在方法、结果、结论和/或建议中提到两性平等方面。甚至更少(7 份)提到人权问题。这个结果与监督厅在方案一级监测和评价的检查结果相同。⁸

B. 2008-2009 年对本组织各优先领域的评价结果、结论和建议提供了一个改进路线图

18. 监督厅审查了 42 个评价报告样本，以查明 2008-2009 年评价本组织八个优先领域的结果、结论和建议。下面列出了一些例子。但应当指出，在裁军领域，没有收到任何报告。

维护国际和平与安全

19. 维持和平行动部对非洲联盟-联合国达尔富尔混合行动的评价发现没有既定的区域政治战略；军事单位和各组织单位的任务和职责没有界定；警察和军队之间也没有既定的业务程序。

20. 对外层空间事务办公室的评价发现，该办公室面临能见度不足、积极主动不够以及在联合国系统内联系不多等挑战。

促进持续经济增长和可持续发展

21. 联合国人类住区规划署(人居署)对贫民窟改造基金的中期审查发现，它已吸引到改造贫民窟的私人资金，并为可持续地解决城市贫民住房问题提供了一个模式。审查注意到必须使贫民窟改造基金更好地纳入国家和地方的政策框架和体制环境。

22. 亚洲及太平洋经济社会委员会(亚太经社会)对泛亚铁路方案的发展和正式运营的评价指出该铁路方案对内陆发展中国家的重要性。铺设新线、调整现有的路线以及修订技术标准都有助于提高该方案的可持续能力。

23. 联合国环境规划署(环境署)对联合国全球环境监测系统水方案的评价发现，它已完成其大部分任务。它的主要挑战是要确保有足够持续供应的资金，以便继续维持该方案，使另外 35 个发展中国家能够积极参与。

⁸ 见检查和评价司关于下列部门监测和评价的检查报告：联合国毒品和犯罪问题办公室(IED-09-006，第 15-16 段)；管理部(IED-10-002，第 25-26 段)；维持和平行动部/外勤支助部(IED-10-001，第 42-44 段)；联合国贸易和发展会议，(IED-10-010，第 38-40 段)；安全和安保部，(IED-11-001，第 53-57 段)。

非洲发展

24. 非洲问题特别顾问办公室对大会关于“非洲的发展需要：各种承诺的履行情况、挑战和今后的道路”的高级别会议的评价表明，它使与会者有机会重申其承诺，促进非洲的发展和解决各项挑战。

促进人权

25. 联合国人权事务高级专员办事处(人权高专办)对一项加强实施人权条约建议的倡议进行的评价发现，其影响和效果有赖于国家行为人的作用以及与外部机构的合作。该倡议的总体目标和具体目标是明确的、可行的和相互关联的，外部和内部监测工具均已改善。面临的挑战包括缺乏财务可持续能力，人手不足，资源限制造成延误，内部的职责和任务没有界定。

有效协调人道主义援助工作

26. 联合国难民事务高级专员办事处(难民高专办)对努力防止和应对人口贩运的评价发现，其关于贩运问题的既定政策没有详细处理预防问题。它发现，各办事处在实施政策方面很不一致、缺乏培训、而且参考材料过时。此外，来自几个国际组织和国际非政府组织的代表认为，他们同难民高专办之间有尚未实现的合作潜力。

27. 人道主义事务协调厅对回应纳尔吉斯气旋的活动的活动的评价发现，总的来说，人道主义反应是积极正面的——食品、收容所、医疗保健和其他重要部门都有广泛的覆盖面。然而，一些具体领域，包括分组间的协调，发现有弱点。评价报告建议人道协调厅加强与灾害发生后受影响社区的沟通和磋商。

28. 联合国近东巴勒斯坦难民救济和工程处(近东救济工程处)的评价描述它的保健方案为一个“运行良好的提供综合初级保健的发动机”，它可以提供更广泛的服务，更好地利用其现有的资源。为了可持续起见，需要有更多的资金和更好的预算编制和财务管理工具。

促进司法和国际法

29. 法律事务厅/联合国训练研究所(训研所)关于条约法和惯例及国内落实条约义务问题能力建设讲习班的参与者给予该讲习班积极的评价。36名参与者中，妇女占了将近一半。总的来说，参与者对扬声器、信息和材料均表满意，并感谢创造机会让他们结交来自各区域的参与者和联合国专家。评价为改善今后的讲习班提出了具体建议。

毒品管制、预防犯罪和打击一切形式和表现的国际恐怖主义

30. 联合国毒品和犯罪问题办公室(禁毒办)对巴基斯坦药物滥用问题妇女的咨询及治疗方案的评价发现它实现了大部分目标。方案的设计和执行显示资源

获透明和有效利用，并证明有专业和承担精神。该方案又推广了药物滥用方面的知识。该方案的需求超过预期。应为该方案争取长期资助，以提高其可持续能力。

C. 内部评价报告总体质量令人满意，但方案之间和之内差别很大

31. 平均来说，经审评的评价报告总体质量是尚可，即 5 分制的中点，5 分制反映从优秀(1)到非常差(5)这五个可能的质量范围。方案之间和之内的质量差别很大。16%的报告被评为优秀，而 20%则质量差。质量好坏，部分取决于评价范围和资源有无着落。有些报告是全面的评价报告，而其他则范围相当有限(例如，列出与会者调查结果)。各方案之间的高质量和低质量报告分布较为均匀。有 6 个方案的七份报告被评为“优”。另 6 个方案的九个报告被评为“差”。

32. 平均来说，在引言、方法、背景、结果和评价报告格式各方面的质量都被评为“好。”不过，有一些具体的共同弱点。报告的作者往往没有说明选择数据的理由，也没有解释数据收集和分析方法。关于方法上的挑战和/或限制以及评价结果的可靠性和有效性方面，报告所述不多，属较弱环节。

33. 许多评价结果都被认为是太过描述性的叙述。通常主要评价结果并没有在报告清楚浮现。许多报告没有适当地汇总和综合各种证据(数字和意见)。虽然有些报告提出很多证据来支持各项评价结果，但其他报告只选择性地提出证据，或只在附件提出。一些评价结果只有有限的表面证据支持(例如，一次访谈)。

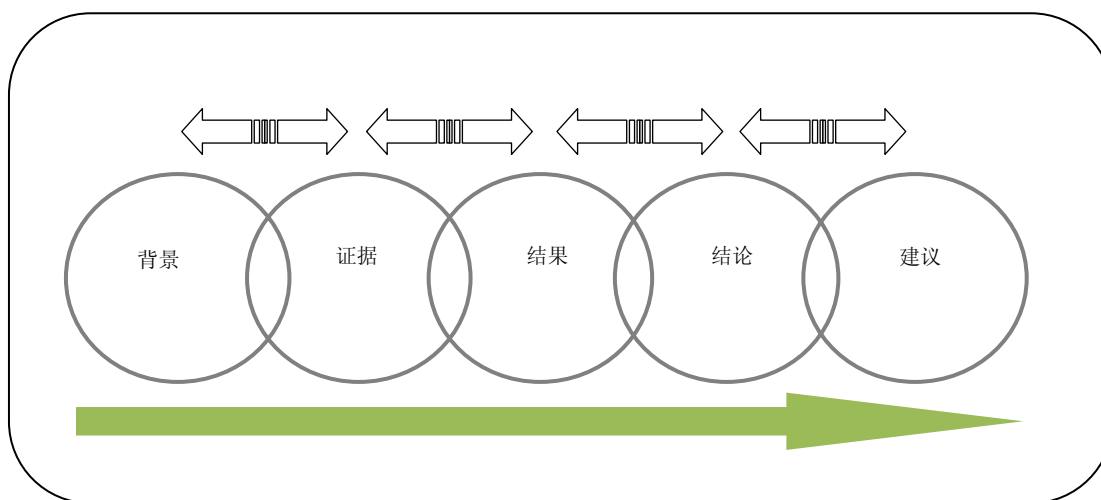
34. 分析深度也相差很大。有时，评价结果的一些内容只表达评价员证据不足的意见。在某些情况下，评价结果的得出似乎没有对事实的意义和影响作出过充分的思考。有些报告没有解释事件或变化的根本原因。

35. 平均来说，执行摘要、结论和建议的质量被评为“尚可”。结论部分是经审评的评价报告中最弱的环节：有些报告没有结论，而其他报告的结论基本上是扼要说明评价结果，增加值不多。

36. 这些报告的一项挑战是在概念上区分开：背景资料和证据；证据和评价结果；特别是，评价结果和结论。虽然大家承认这各个类别并不总是有清楚的分野(见图三)，质量高的评价报告对于这些报告内容都有清楚明确的概念。

37. 有些报告提出太多过于规范性的建议，而其他报告只提出几条很短的建议，看来不像是对所有评价结果和结论作出过系统分析得出的成果。建议应该可以操作，应当在适当战略层面上提出。

图三
区分类别是评价员的一项挑战



D. 评价为多种重要的用途服务

38. 经审查的 71 份评价报告全部有方案实施的结果和结论。97 份报告全都有关于方案产出、方案成果、结果和影响方面的评价结果。经审查的 2008-2009 年评价报告中，89%有方案设计方面的评价结果，而只有不到一半(45%)有政策指示方面的结果。一般来说，政策问题比针对具体实施过程或交付具体产出的问题较为广泛，因而出现于评价报告中的次数可能较少。

39. 在审议各项报告提出的建议时也发现类似的模式出现。几乎所有经审查的报告都载有方案产出和方案实施的建议。90%经审查的报告载有关于成果、结果和影响的建议，86%有关于方案设计的建议。但只有不到一半(45%)经审查的报告载有关于政策指示的建议。

40. 协调人调查结果证实了这些评价结果。65%的答复者表示，在他们的方案中，评价报告被用来改善目前的方案。在 62%的方案中，评价资料用来供未来的方案和/或预算规划学习参考。大约有一半(48%)利用评价报告于政策性决定，而只有 10%利用它们来制订公共宣传或媒体材料。62%方案利用评价资料向高一级管理层汇报，55%用来向政府间机构汇报，39%用来向捐助者汇报。

41. 从 2008-2009 年各项评价报告产生的后果可以看出，大家都侧重于目前方案的实施情况。最经常提到的后果(55%答复者)是，这些报告促使人们对目前方案的实施情况作出更知情、更有效和更相关的决策。第二经常提到的后果(52%答复者)是提高组织/运作效率。只有 45%指出改进方案业绩(实效)，而更少(29%)提到评价报告所产生的后果促使政策的改变。

42. 协调人提到 2008-2009 年具体评价报告在方案活动、方案管理和评价方法等方面产生的后果如下：

(a) 大会和会议管理部根据联合国总部成功的经验向其他工作地点(日内瓦办事处、维也纳办事处、内罗毕办事处)颁布各种统计方法。大会部所有工作地点的政府间机构调查均已标准化；

(b) 政治事务部报告利用评价结果作为分析的基础，通过多年拟议经常预算大大加强了该部；

(c) 新闻部的部门评价一直强调增强互动在线存在的重要性(例如，2008 年对《世界人权宣言》六十周年的年度方案影响审查和对重新设计联合国网站的 2009 年度方案影响审查)。该部的反应是努力改变互动能力有限的网站的使用政策，有系统地在通信规划中纳入社交媒体工具，以促进互动能力，这是联合国的主要网站所无法提供的；

(d) 在非洲经济委员会(非洲经委会)，评价有助于促进成果管理文化，促使方案主管更多地参与。评价活动还改进了方案/项目的说明和重点(例如发展账户项目)，并改进与总体工作方案的协作；

(e) 拉丁美洲和加勒比经济委员会(拉加经委会)，评价结果被用来促使加勒比次区域办事处的工作更为相关。更好地将拉加经委会的工作成果和影响记录于文件，使它变得更有重点和有效。

(f) 亚洲及太平洋经济社会委员会(亚太经社会)在 2012-2013 两年期方案预算规划中转用“方案办法”，更加重视能力发展办法，同时整合预算外和经常预算工作。这是评价建议的一个直接成果；

(g) 人道主义事务协调厅发现，机构间实时评价有助于支持实地学习，改善人道主义业务，缓解协调方面的瓶颈。对中央应急基金的评价被用来改善基金的运作，并进一步建立其对捐助方的信誉。2009 年，人道协调厅元评价被用来设计人道协调厅 2010-2013 年战略计划；

(h) 联合国人权事务高级专员办事处(人权高专办)指出，评价活动帮助它改善工作流程，形成关于两性平等和地区办事处的新政策。评价还导致人权高专办各部分之间更多的信息共享和协调统一，数量较少但更有重点的战略优先次序，并改进人道主义领域的伙伴关系；

(i) 在最不发达国家、内陆发展中国家和小岛屿发展中国家高级代表办事处，一项关于精简和合并机构间协调努力的建议有助于该办事处继续召开机构间协商会议，作为其工作方案的经常特点；

(j) 在法律事务厅某些领域，评价有助于改善 2010-2011 年和 2012-2013 年的战略框架；

(k) 联合国人类住区规划署的中期战略和体制计划中所述的着重成果的做法增加了部门间的协作。

(l) 在联合国环境规划署，评价建议遵守制度提供了一个改善业绩的直接链接。评价办公室整理了环境署各司遵行建议的资料，执行主任用这些资料来衡量各司司长的业绩；

(m) 在联合国难民事务高级专员办事处，一份评价报告成为了难民高专办和合作伙伴举办研讨会以重新调整方案方针的基础。另一份报告通报了全面改革组织结构的执行决定。结果创建了一个更强大的供应链管理服务，将它从日内瓦搬到布达佩斯。

E. 秘书处方案一级评价做法和能力概况

秘书处用于评价的资源停滞不加

43. 监督厅 2007 年评估联合国秘书处的评价能力和需要的报告发现，秘书处在中央、方案和次级方案三级的评价能力不够，原因是财政和人力资源不足，能力参差不齐，而且缺乏高级领导层的支持。⁹

44. 评价预算继续偏低。2008-2009 年情况改善，但在 2010-2011 两年期没有得到维持(见图四)。2010-2011 年为秘书处编列的监测和评价活动资源共达 5 100 万美元。然而，所增加的 640 万美元监测和评价经费低于两个两年期之间资源总额的相对增加。¹⁰ 因此，预算中用于监测和评价的资源比例下降，从 0.42%(2008-2009 年)降至 0.37%(2010-2011 年)。目前所占比例约为建议一般评价基准下限的三分之一，根据一般评价基准，至少要将方案经费总额的 1%至 3%用于评价活动(IED-2006-006，第 30 段)。

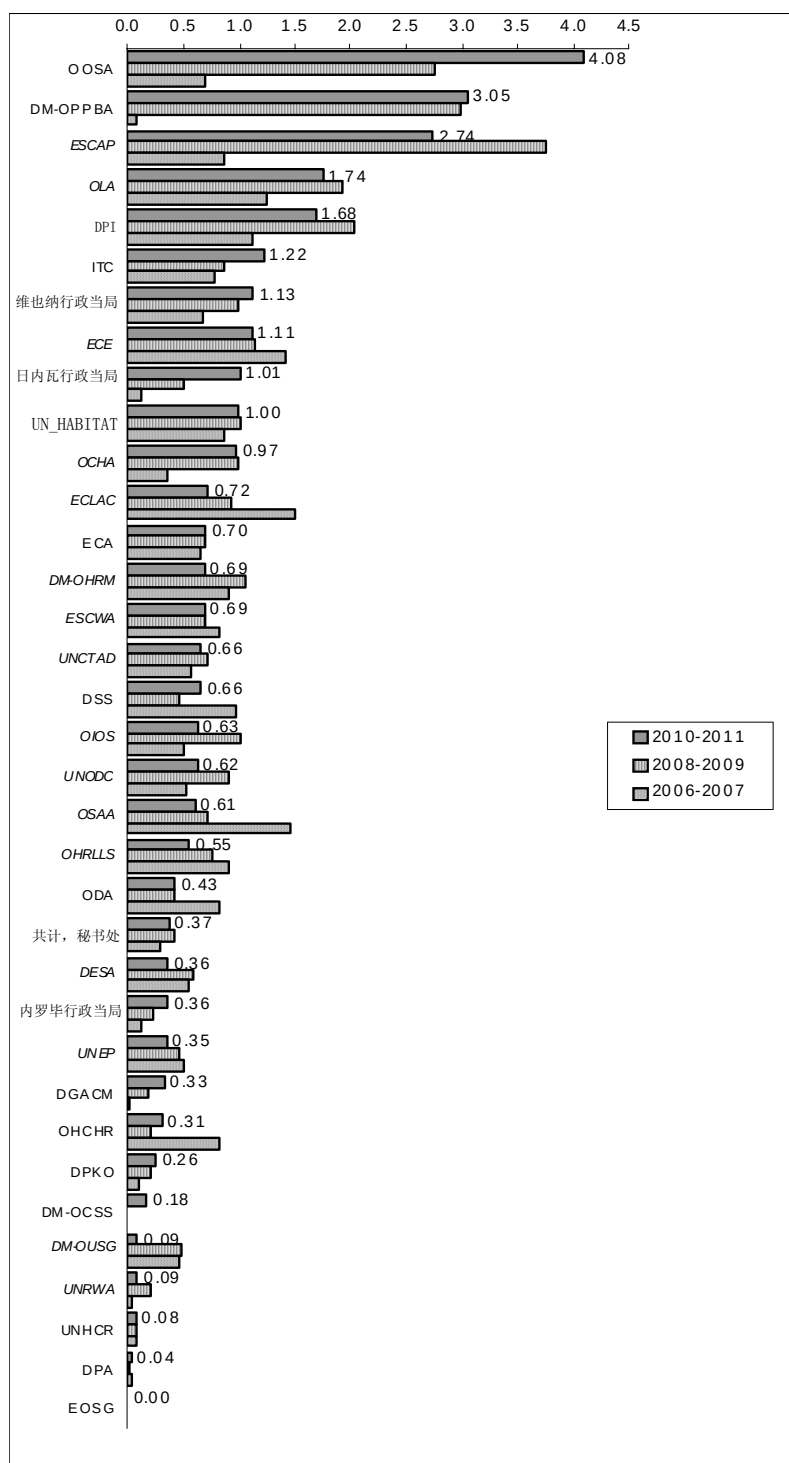
45. 2008-2009 年，9 个方案指定用 1%或以上的资源于监测和评价活动，而在 2010-2011 年，有 10 个方案达到这个水平。许多协调人(15，占答复者的 56%)报告，缺乏足够的资源是他们的方案在管理和进行评价方面的最大障碍。

⁹ A/64/63，第 43 段；又见《评估联合国秘书处的评价能力和需要》(IED-2006-006)。

¹⁰ 这些数字所系根据检查和评价司对个别方案预算专册的审查，与 A/64/6(导言)第 61 段载列的数字不同。

图四

用于评价的资源占总方案资源的比例(百分比)



注：斜体显示的方案在 2008-2009 年和 2010-2011 年期间的评价资源占预算额的比例下降。

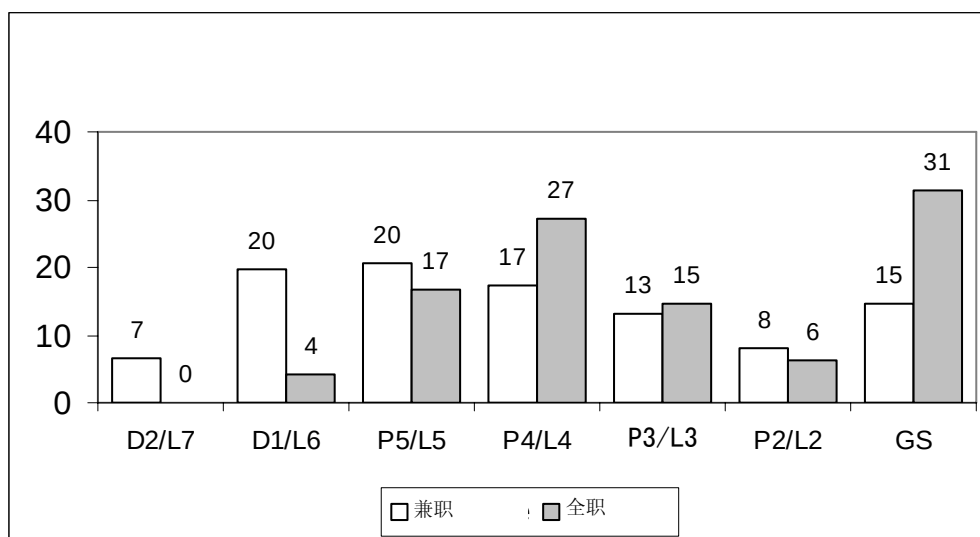
简称：DESA，经济和社会事务部；DGACM，大会和会议管理部；DM，管理部；DPA，政治事务部；DPI，新闻部；DPKO，维持和平行动部；DSS，安全和安保部；EOSG，秘书长办公厅；ITC，贸发会议/世贸组织国际贸易中心；OCHA，人道主义事务协调厅；OCSS，中央支助事务厅；ODA，裁军事务厅；OHRM，人力资源管理厅；OLA，法律事务厅；OOSA，外层空间事务办公室；OPPBA，方案规划、预算和账户厅；OSAA，非洲问题特别顾问办公室；OHRLLS，最不发达国家、内陆发展中国家和小岛屿发展中国家高级代表办事处；OUSG，副秘书长办公室；UNODC，联合国毒品和犯罪问题办公室

来源：监督厅对 2010-2011 年、2008-2009 年、2006-2007 年预算专册的分析 (A/64/6、A/62/6 和 A/60/6)。

46. 有 5 个方案(经济和社会事务部、管理部、秘书长办公厅、联合国日内瓦办事处和联合国内罗毕办事处)报告没有工作人员专门负责评价活动，其他方案表示有 1 至 49 名工作人员负责进行和/或管理评价。秘书处总共有 170 名工作人员进行评价活动，不过其中 122 名只是兼职。只有 11 个方案有全时专职的评价人员。通常，这些方案有独立评价单位。有两个方案雇用全职和兼职工作人员，而在 13 个方案，负责评价的所有工作人员还有其他职责。图五按职等列出兼职和全职评价人员的分布情况。

图五

各级兼职和全职评价人员(百分比)



来源：监督厅对各方案协调人的调查。

评价计划的贯彻实施没有保证

47. 在审评 2009 年秘书长与 23 名高级主管的契约中，实现“实施自我评价计划”目标方面被认为“满意”的占 78%。然而，在提出的详细程度方面，这些契约与 2008-2009 年各方案评价计划有很大差异。

48. “经常对所有方案和次级方案进行有效的自我评价”这个预期成果再也不是2010年秘书长与高级主管所订契约的内容。这就引起人们关注，因为只有42%的协调人报告他们的部门规定“正式跟踪和/或监测(两年一次)的评价计划”。监督厅对方案一级的监测和评价检查还发现，管理人员的职务说明不包括监测和评价职能。¹¹

方案一级的评价面临多重挑战

49. 对方案一级监测和评价的检查显示，秘书长两年期方案执行情况报告对产出交付的汇报非常准确。¹² 然而，在汇报包括交付产出的已取得的成果方面，各方案面临重大的挑战。对于所有的方案，大会规定在预算期终了时用来衡量实现程度的绩效指标。在迄今经检查的方案中，各方案需要收集来衡量其业绩指标和评估其预期成果实现程度的数据有11%至44%没有正确列认。¹³ 此外，在各方案描述其实现预期成绩的成果说明中，只有44%至85%的绩效指标被称为尺度。¹⁴ 因此，在15%至56%的案例中，各方案汇报预期成果的实现程度时没有提及大会规定的标尺。大会制定的评价系统并没有按原意被充分利用。

50. 协调人调查的答复将数据方面的困难列为他们的方案在管理和进行2008-2009年自我评价时所面临的三大障碍之一。这些障碍包括：

- (a) “数据不一致、不可得和不可靠；
- (b) 难以得到充分的反应；
- (c) 无法轻易访问/提取不同系统的数据；
- (d) 用来收集有意义数据工具不足。

51. 一个方案的评价文化可造成另一个障碍。各协调人特别列出了以下挑战：

- (a) 来自同事对评价活动的阻力；
- (b) 来自管理层的阻力；
- (c) 克服方案管理人员对评价的恐惧；

¹¹ 例如关于以下各部门的审查报告：贸发会议(IED-10-010，第27和28段)；安全和安保部(IED-11-001，第45段)和非洲经济委员会(IED-09-005，第27段)。

¹² 准确度在90%或以上，但对经济和社会事务部和非洲经济委员会监测和评价活动的审查除外，它们采用的方法不同，得出的百分比较低。

¹³ 见以下部门的审查报告：联合国贸易和发展会议(IED-10-001，第17段)和维持和平行动部/外勤支助部(IED-10-001，第17段)。

¹⁴ 见以下部门的审查报告：维持和平行动部/外勤支助部(IED-10-001，第15段)和安全和安保部(IED-10-001，图2)。

- (d) 高级主管缺乏自我评价的决心；
- (e) 高级主管完全不支持 and 无知；
- (f) 高级主管不愿处理困难的结构性问题；
- (g) 方案和项目主管的漠不关心。

52. 最后，评价职能本身的特点为进行和管理自我评价造成障碍。协调人提到的这些问题包括：

- (a) 没有全时评价职能；
- (b) 缺乏指定的评价办公室；
- (c) 缺乏自我评价的专用能力；
- (d) 缺乏评价政策；
- (e) 缺乏独立性和公正性；
- (f) 不断依赖外部评价；
- (g) 制订评价专用政策和程序方面的挑战；
- (h) 缺乏实务指引。

评价政策和独立评价单位加强方案一级的评价职能

53. 各协调人报告指出，在 2008-2009 两年期，有 15 个方案订有评价政策，其中 11 项在 2008 年至 2010 年间制订或更新。比起 2006 年年底订有评价政策的 7 个方案来说，这是大幅增加(IED-2006-006, 第 59 段)，但是没有具体评价政策(不包括方案规划、预算编制、监测和评价方法条例)的秘书处方案占一半以上。¹⁵

54. 为确保评价职能的独立性，联合国评价小组建议成立独立评价单位。有 10 个协调人报告他们的方案有一个独立的评价单位或办事处，专门负责评价工作。¹⁶ 虽然联合国评价小组建议评价管理人员直接向所属部门的负责人报告，但只有一

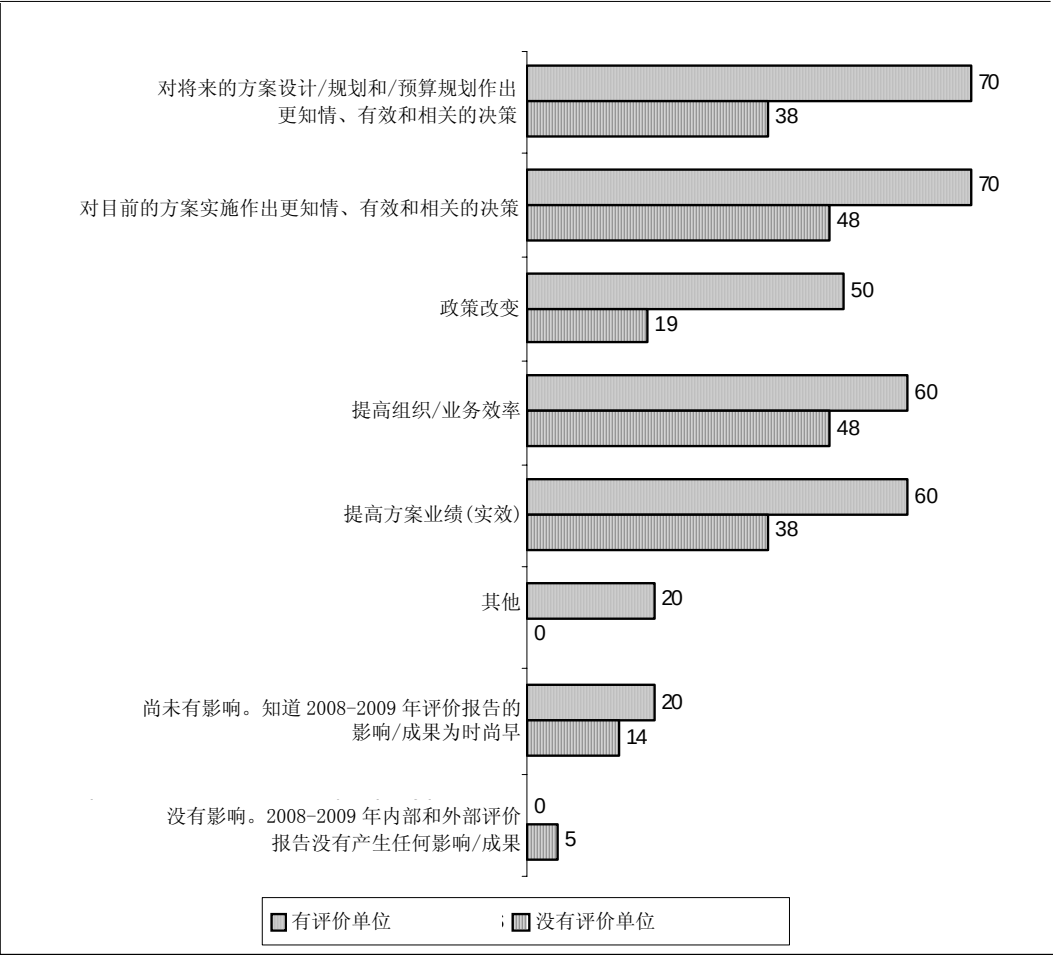
¹⁵ 截至 2009 年 12 月没有评价政策的 16 个方案为：经济和社会事务部、大会和会议管理部、管理部、新闻部、安全和安保部、非洲经济委员会、欧洲经济委员会(2010 年制订了评价政策)、秘书长办公厅、最不发达国家、内陆发展中国家和小岛屿发展中国家高级代表办事处、法律事务厅、外层空间事务办公室、非洲问题特别顾问办公室、联合国贸易和发展会议、联合国人类住区规划署、联合国日内瓦办事处和联合国维也纳办事处。

¹⁶ 大会和会议管理部、维持和平行动部/外勤支助部、新闻部、欧洲经济委员会、国际贸易中心、人道主义事务协调厅、联合国环境规划署、联合国难民民事务高级专员办事处、联合国人类住区规划署和内部监督事务厅。

半的独立评价单位有这种汇报结构。在没有独立评价单位的部门，评价工作由业务单位负责，通常由部门首长办公室协调。

55. 图六显示各协调人对来自 2008-2009 年各方案评价报告的成果说明以及对是否有独立评价单位的说明。在设有独立评价单位的方案，评价报告比较有可能产生有利的影响。在政策变化方面，这方面的分别特别引人注目，50%有评价单位的方案报告有这方面的影响，没有评价单位的方案只有 19%有政策变化方面的影响。

图六
评价报告的影响，按方案和是否有独立评价单位开列(百分比)

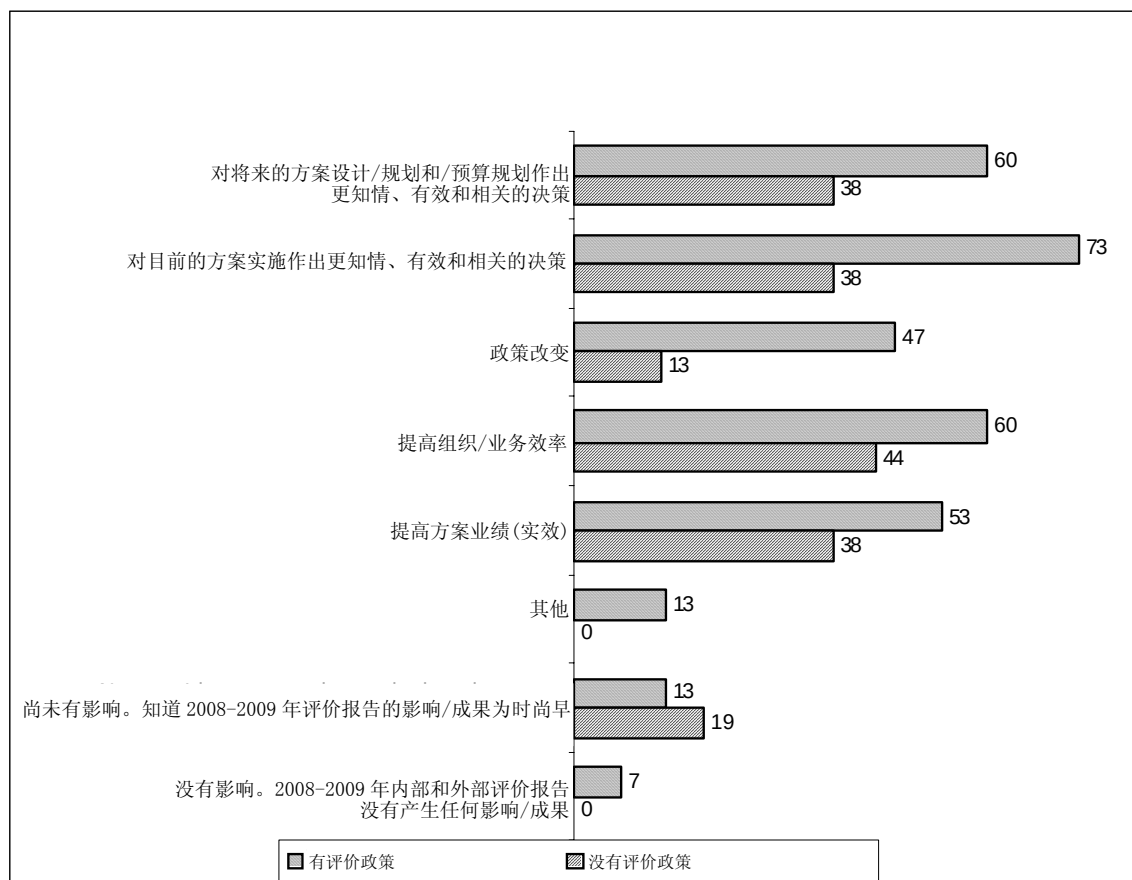


来源：监督厅方案协调人调查。

56. 如果将 2008-2009 年有或无评价政策的方案的评价报告所产生后果作一比较, 也可以观察到类似的效果(图七)。不过, 应对这些结果作慎重的解释。许多因素影响到评价报告是否产生后果。有无评价政策和/或评价单位可积极影响评价在方案中的用处, 并与评价过程和程序有联系(图八)。

图七

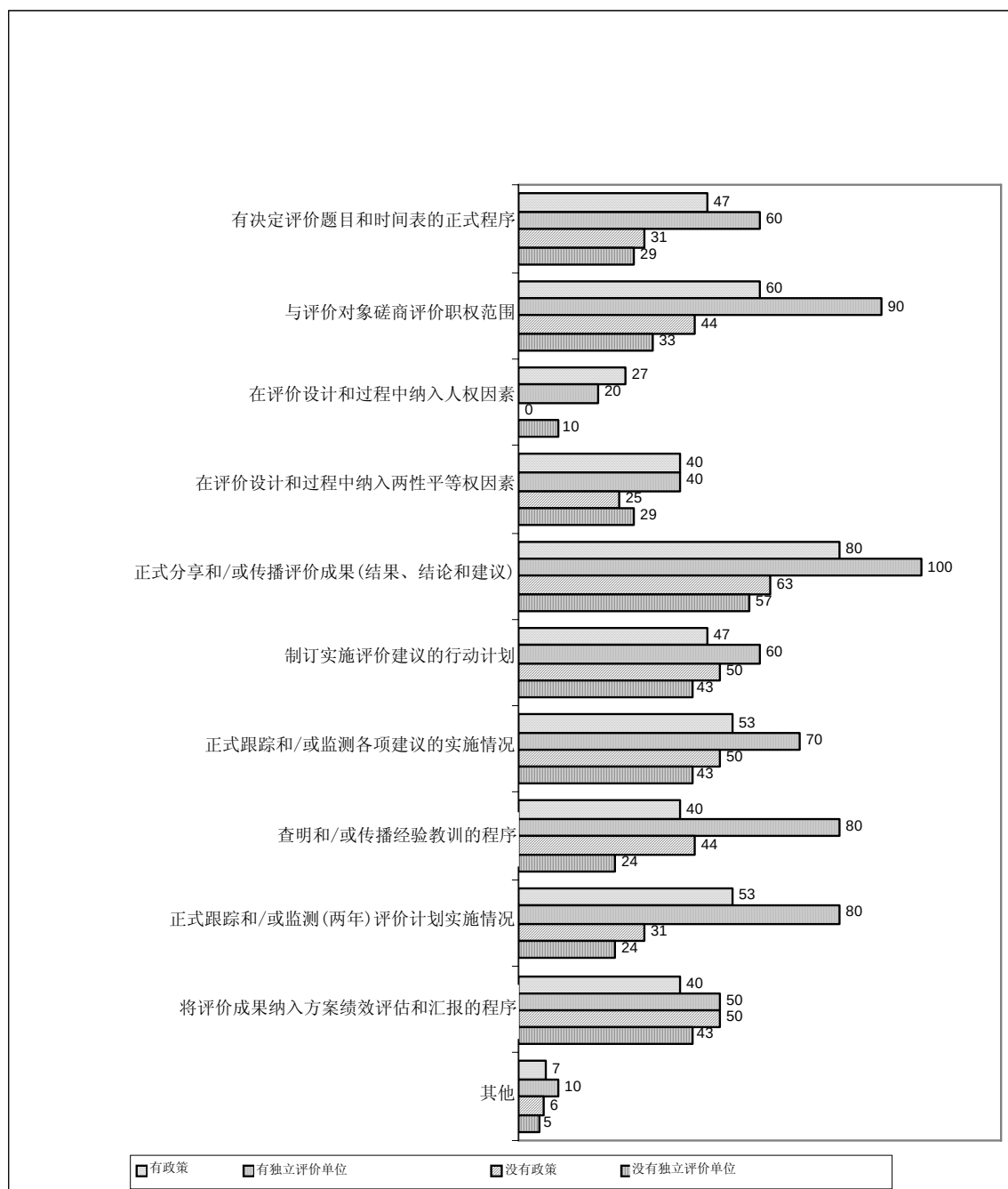
评价报告的影响, 按方案和是否有评价政策开列(百分比)



来源: 监督厅方案协调人调查。

图八

有和无评价政策和评价单位的方案设定的评价过程和程序(百分比)



来源：监督厅方案协调人调查。

57. 最后，有评价政策会积极影响总的评价质量，而有独立评价单位却没有这种影响。在质量评估中，所有评价报告的平均分为 2.62(介于“好”与“尚可”之间)，而有独立评价单位的方案的平均分为 2.55，有评价政策的方案的平均分为 2.33(优=1，非常差=5)。虽然第一个结果并不显著，第二个则十分显著。也就是说，凡是有评价政策的方案，其评价报告的质量会好很多。

评价尚未成为一个完全获接受的管理职能

58. 当问及在 2008-2009 年管理和/或进行评价的三个最有益的经验时，约有一半的协调人都提到与评价本身有关的因素(如它提供了一个“契机，开展一个参与性的、集体吸取经验教训的活动，了解实际工作的具体情况和如何做更好”，或者说它“侧重于收集有关统计数据，以供管理当局检查和评价”)。其他协调人提到与方案活动更密切相关的专题(例如这个过程被认为“信息含量大，有助于顺利完成次级方案的活动/产出”、“改进方案”和“横向一致”)。关于管理问题，如“创造一个透明和问责环境”，“给予和接受关于方案管理的真正回馈”和“介绍自我评价结果，让下级管理人员直接接触到高级主管，讨论和解决结构性问题”等等，都是大家提到的进行评价的其他有益处的方面。

59. 从协调人调查得出的对评价职能的看法一般是正面的，而监督厅对方案级监测和评价的检查结果缓和了这种正面的看法，因为它的目标群体较大。例如，在访问贸发会议工作人员时，管理层对监测和评价的承诺得到的评价不一。工作人员感到管理层只承诺必要时进行监测和评价，一般来说，监测和评价没有被高级方案工作人员看作是管理文化的组成部分。¹⁷ 监督厅对其他部门的监测和评价进行的检查也发现类似的结果。¹⁸ 监督厅经常发现管理人员不知道自己的监测和评价责任。¹⁹ 在调查中，关于评价是否可以有助于评估方案业绩、制订未来的方案、并吸取经验，以及关于监测和评价工具能否有效协助达成部门目标，具有管理责任的工作人员的评分比非管理人员的低。²⁰ 工作人员认为，监测和评价往往没有获得有效管理，因此工作人员视之为对他们的工作价值不大的行政活动。²¹ 检查的结论指出，需要得到所有各级管理人员作出坚定的承诺，才能全面实施综合自我评价议程。

¹⁷ 见以下部门的监测和评价审查报告：联合国贸易和发展会议(IED-10-001，第 28 和 42 段)。

¹⁸ 见以下部门的监测和评价审查报告：非洲经济委员会(IED-09-005，第 26 段)和管理部(IED-10-002，第 23 段)。

¹⁹ 见以下部门的监测和评价审查报告：安全和安保部(IED-draft，第 40 段)。

²⁰ 根据维持和平部/外勤支助部(IED-10-00)，附件六和图 6。

²¹ 同上，第 52 段。

F. 内部监督事务厅的评价活动和工作计划

60. 为了确立评价优先次序，监督厅利用系统化、战略风险规划方法，其中采用 2007 年推出的 12 个代用风险指标。前一个两年期报告 (A/64/63) 介绍了这个办法的详细内容。

61. 在 2008-2009 年，监督厅完成了下列评价 (产生 11 份报告)：

- (a) 深入评价政治事务；
- (b) 深入评价人力资源管理厅；
- (c) 人权高专办方案评价；
- (d) 最不发达国家、内陆发展中国家和小岛屿发展中国家高级代表办事处和非洲问题特别顾问办公室方案评价；
- (e) 经验教训专题评价；
- (f) 秘书处环保工作专题评价；
- (g) 联合国协调机构专题评价；
- (h) 联合国科特迪瓦行动 (联科行动) 方案评价；
- (i) 维持和平行动中解除武装、复员和重返社会的专题评价；
- (j) 建设和平基金特别评价；
- (k) 大会和会议管理部全球统筹管理措施特别评价。

62. 2010-2011 年，监督厅完成了 26 份评价报告，涵盖以下 9 个专题：

- (a) 联合国近东巴勒斯坦难民救济和工程处方案评价；
- (b) 管理部方案评价 (完成了 2 份摘要和 6 个办公室报告——副秘书长办公室、中央支助事务厅、方案规划、预算和财务厅、联合国日内瓦办事处、联合国维也纳办事处和联合国内罗毕办事处)；
- (c) 经济和社会事务部方案评价 (1 份摘要和 10 个次级方案报告)；
- (d) 联合国秘书处性别观点主流化专题评价；
- (e) 联合国秘书处应对气候变化工商伙伴关系专题评价；
- (f) 联合国利比里亚特派团 (联利特派团) 方案评价；
- (g) 联合国海地稳定特派团 (联海稳定团) 方案评价；
- (h) 联合国苏丹特派团 (联苏特派团) 方案评价；

(i) 维持和平行动部和外勤支助部与区域组织合作专题评价。

63. 大会在第 64/229 号决议核可了方案和协调委员会载在其第四十九届会议工作报告(A/64/16, 第二章, 二.B)关于评价的结论和建议。委员会选定经济和社会事务部方案评价, 供其在 2011 年第五十一届会议审议, 以及选定维持和平行动部和外勤支助部与区域组织合作的专题评价, 供有关政府间机构审议(同上, 第 41 段)。

64. 假设可以继续维持现有的资源水平, 监督厅将继续对所有秘书处方案评价实行 11-13 年方案评价周期。下表列出了下一个十二年周期评价各部门的计划。

表 1

内部监督事务厅方案评价 12 年周期

方案 (按检查和评价司风险评价排列)	完成或计划进行评价年份 (斜体为计划进行的评价)
1. 维持和平行动部	进行中, 有支助账户
2. 外勤支助部	进行中, 有支助账户
3. 人权事务高级专员办事处	2009
4. 联合国环境规划署	2013
5. 联合国难民事务高级专员	2012
6. 管理部(包括联合国日内瓦办事处、联合国内罗毕办事处和联合国维也纳办事处)	2010
7. 人道主义事务协调厅	2012
8. 联合国毒品和犯罪问题办公室	2013
9. 非洲经济委员会	2014
10. 联合国近东巴勒斯坦难民救济和工程处	2010
11. 政治事务部	2008
12. 联合国人类住区规划署	2014
13. 联合国贸易和发展会议	2014
14. 经济和社会事务部	2011
15. 大会和会议管理部	2015
16. 安全和安保部	2015
17. 裁军事务厅	2015
18. 秘书长办公厅	2016
19. 亚洲及太平洋经济社会委员会	2016
20. 新闻部	2017
21. 国际贸易中心	2017

方案 (按检查和评价司风险评价排列)	完成或计划进行评价年份 (斜体为计划进行的评价)
22. 拉丁美洲和加勒比经济委员会	2018
23. 非洲问题特别顾问办公室/非洲发展新伙伴关系	2009
24. 最不发达国家、内陆发展中国家和小岛屿发展中国家高级代表办事处	2009
25. 外层空间事务办公室	2017
26. 欧洲经济委员会	2018
27. 法律事务厅	2019
28. 西亚经济社会委员会	2019

65. 到本两年期结束时，监督厅将完成以下方案评价：政治事务部、最不发达国家、内陆发展中国家和小岛屿发展中国家高级代表办事处、非洲问题特别顾问办公室、联合国近东巴勒斯坦难民救济和工程处、人权事务高级专员办事处、管理部和经济和社会事务部。2012-2013 年，监督厅计划完成难民署、人道主义事务协调厅、联合国环境规划署和联合国毒品和犯罪问题办公室的方案评价。为了实现其对所有秘书处方案的 11-13 年评价周期的目标，监督厅将努力在接下来的六年完成其余 13 个评价。此外，为了跟进经方案和协调委员会核可的建议，监督厅将继续执行已获授权的三年度审查方案。

66. 关于评价维和行动(由支助帐户供资)，在 2008-2009 年期间，监督厅首次进行“任务整体”评价，2009 年以联科行动开始。2010 年，检查和评价司完成了对联合国海地稳定特派团(联海稳定团)和联合国利比里亚特派团(联利特派团)的评价，并在 2011 年完成了对联合国苏丹特派团(联苏特派团)的评价。2012 年，监督厅计划对联合国组织刚果民主共和国稳定特派团(联刚稳定团)和联合国东帝汶综合特派团(联东综合团)进行评价，或重新评价联合国海地稳定特派团，以更好地了解灾害对特派团的影响，以及如何可以减轻其影响。鉴于目前维持和平特派团的数量(14)，根据现有的资源，需要另外四年时间才能评价其余的特派团。

67. 2009 年，检查和评价司发起了一系列监测和评价的检查，以检查和核实方案业绩报告的数据和审评方案自我评价的质量，作为监测的一个手段，并借此通报秘书长和会员国有关方案汇编成果的可靠性以及各方案的学习与提高能力。

68. 自 2009 年以来，监督厅平均一年进行三次监测和评价。由于进行这些检查的方法和能力现已得到发展，监督厅计划争取实现更高的检查效率，努力做到每年平均进行五到六次检查。监督厅打算到 2014 年完成对所有方案的监测和评价检查，实际上是一个六年周期，比四年周期这个目标长一些(见表 2)。

表 2
内部监督事务厅监测和评价检查

方案 (按检查和评价司风险评价排列)	完成或计划进行评价年份 (斜体为计划进行的评价)
1. 维持和平行动部/外勤支助部	2010
2. 人权事务高级专员办事处	2011
3. 联合国环境规划署	2011
4. 联合国难民事务高级专员	2013
5. 管理部(包括联合国日内瓦办事处、联合国内罗毕办事处和联合国维也纳办事处)	2010
6. 人道主义事务协调厅	2011
7. 联合国毒品和犯罪问题办公室	2009
8. 非洲经济委员会	2009
9. 联合国近东巴勒斯坦难民救济和工程处	2011
10. 政治事务部	2012
11. 联合国人类住区规划署	2012
12. 联合国贸易和发展会议	2010
13. 经济和社会事务部	2009
14. 大会和会议管理部	2012
15. 安全和安保部	2011
16. 裁军事务厅	2012
17. 秘书长办公厅	2012
18. 亚洲及太平洋经济社会委员会	2013
19. 新闻部	2013
20. 国际贸易中心	2013
21. 拉丁美洲和加勒比经济委员会	2013
22. 非洲问题特别顾问办公室/非洲发展新伙伴关系	2014
23. 最不发达国家、内陆发展中国家和小岛屿发展中国家高级代表办事处	2014
24. 外层空间事务办公室	2014
25. 欧洲经济委员会	2014
26. 法律事务厅	2014
27. 西亚经济社会委员会	2014

69. 除了这个现已标准化的监测和评价检查方案，监督厅还针对已查明的风险开展专案检查。2009 年，监督厅对经济和社会事务部进行了一次人力资源和管理做法的检查。

选择供方案和协调委员会审议的评价专题

70. 方案和协调委员会不妨考虑想在 2013 年第五十三届会议审议监督厅 2012-2013 年工作计划中哪些评价，并请监督厅进行目前不在工作计划内的任何其他评价。2012-2013 年将对难民署、人道协调厅、环境署和禁毒办进行方案评价。除了这些方案评价外，方案和协调委员会不妨考虑计划对联合国秘书处内部司法系统进行的专题评价，这是通过监督厅的战略风险评估选择的专题。

跟进方案和协调委员会建议的行动

71. 在第六十四届会议上，方案和协调委员会建议监督厅继续改进和完善进行评价的方法，以便提出更多有意义的推论和结论(A/64/16，第 43 段)。监督厅通过以下五项行动回应这项建议：

(a) 向监督厅工作人员提供了实质性的评价方法培训：关于焦点小组、访谈技巧、质量数据分析、以及统计分析软件的训练讲习班；一系列涵盖各种专题包括影响评价的研讨会；关于评价设计和方案逻辑模式和理论的讲座。一些工作人员也参加了评价员研究所提供的评价培训，该研究所致力于提高评价方面的专业技能；

(b) 对维持和平行动整体进行评价，对联合国科特迪瓦行动、联合国利比里亚特派团、联合国海地稳定特派团、联合国苏丹特派团进行了方案评价；

(c) 开发和完善了评价方案适当时可用的当地人口调查组成部分，在评价联科行动、联利特派团、联合国海地稳定特派团、人权高专办和联合国近东巴勒斯坦难民救济时曾加以利用；

(d) 2008 年，监督厅要求对它的评价工作的独立性、可信度和用处进行一次外部同行评价。同行审议在联合国评价小组的主持下并与发展援助委员会发展评价网络合作开展，结果指出监督厅的独立性很强，总的来说对其工作的可信度也很满意。除其他外，这项审查建议监督厅更广泛地让各方分享评价成果，以提高评价的效用，特别是为了学习的目的；

(e) 2009 年，监督厅通过一个利益攸关方调查和一个由一名外部顾问进行的独立审查，对 2008-2009 两年期完成的监督厅报告进行了质量评价(见第 6 段，脚注 4)。专家发现，监督厅 75% 的评价报告和 43% 的检查报告的质量“好”。没有任何一份经审查的报告被评为低于“尚可”的质量。建议和结果两节被认为是监督厅报告中最强的部分，而方法、附件和结论数节则有改善空间。

72. 委员会又建议监督厅提请各秘书处注意，必须按照现有细则和条例，对各自方案进行影响评估和深入评价(A/64/16, 第 44 段)。监督厅采取了下列行动以回应这个建议：

(a) 将这项建议通知所有秘书处的协调人；

(b) 参加了联合国评价小组影响评价工作组，制定了联合国影响评价概念说明。监督厅与秘书处所有评价协调人分享这一结果。此外，监督厅将通过工作组参与制订一份专门为联合国系统量身定做的影响评价指引文件。

(c) 讨论影响评价及其用途，作为向大会和会议管理部、西亚经社会，经济和社会事务部和政治事务部等方案自我评价指引的一部分。

四. 结论

73. 总的来说，评价在秘书处的作用和地位高低参半。尽管有更多的方案采用了评价政策，制订了评价过程和程序，而且评价已在整个秘书处更均匀地扩散，可是 2008-2009 年观察到的资源增加没有在 2010-2011 年持续上升。以本组织的规模和复杂性来说，监测和评价的可用资源数额远远低于一般推荐下限(总预算百分之一)。

74. 来自评价的信息用于各种重要的目的，不过评价还没有成为一个完全被接受的管理功能。最常被方案协调人提及的是“改善目前的方案”，其次是“为未来的方案和/或预算规划学习”，以及“向更高级的管理层和政府间机构汇报”。几乎有一半的方案汇报使用评价信息于决策，各协调人提供了许多具体例子，说明它们如何使用这些信息来改进业务。

75. 管理人员对评价的支持比一般工作人员的支持度低，或许因此评价往往不是管理人员正式责任的一部分。2010 年高级主管同秘书长的契约决定删除“经常对所有方案和次级方案进行有效的自我评价”这个预期成果和“实施自我评价计划”这个目标，可能会起反作用。

76. 根据联合国评价小组，如果评价职能独立于其他管理功能，则可促进独立和公正的评价过程；评价负责人应直接向有关组织的理事机构或组织的领导人汇报。²² 各方案的情况往往并非如此。此外，许多方案还没有评价政策或独立单位。

77. 评价报告的平均质量令人满意，而且高质量的报告可以在秘书处不同的部门发现，但是在各部门之间甚或之内，报告的质量差别很大。监督厅提出一个知识管理倡议，以拉近在审查过程中发现的一些明显的能力差距。不过，检查已存在报告的质量不能解决有关评价有否处理最相关专题的问题。

²² 联合国评价小组，《联合国系统评价标准》，第 4 页(2005 年)。

78. 为了建立秘书处的评价能力，必须有训练有素的高质量的评价人员，协助各方案和次级方案查明有关评价专题，并保证自我评价提供的结果可导致改善倡议的实施和/或改进环绕倡议的变革理论。

五. 加强评价报告质量的行动

79. 为支持提高评价报告的质量，监督厅将：

- (a) 与各部门分享关于报告样品的详细质量评价；
- (b) 向各部门提供一个典型评价报告的“简易”模板，各方案可在自愿基础上使用。例如，该模板可以包括标准目录和核对表，列出每一章节的主要重点；
- (c) 拟定一个高质量的评价报告作为例子，说明各种高品质的单元。这些单元可包括：
 - (一) 说明各报告单元的特点；
 - (二) 方法一节的强项；
 - (三) 措辞良好的评价结果、结论和建议；
 - (四) 如何提出证据的例子；
 - (五) 如何纳入两性平等和人权方面的因素的例子。

六. 供方案和协调委员会审议的问题

80. 按照大会第 59/275 号决议，监督厅还提出以下问题，供方案和协调委员会审议。大会在该决议核可了委员会的要求，请监督厅提出“对委员会或其他适当政府间机构提供的政府间指导和后续行动将起到有益作用的问题”（A/59/16，第 383 段）。

81. 《方案规划、预算编制、监测和评价方法细则》第 107.2 条规定：“在编写自我评价报告方面，中央评价股将给予方法上的支助”；又第 107.3 条规定：“自我评价须由方案管理人员按照中央评价股所定的标准进行。该股负责质量标准、方法、评价资料的改写和转移以及专案研究。”

82. 大会第 64/259 号决议部分核可了秘书长提出执行成果管理框架(A/64/640)的建议，监督厅后来指出，无论监督厅或管理部都没有将此功能纳入其 2012-2013 年的战略框架，任何部门的 2012-2013 年战略框架也不再有任何业绩指标反映自我评价能力的发展。

83. 需要澄清方案规划、预算编制、监测和评价有关建设秘书处内部自我评价能力的规定如何在 2012-2013 年得以执行。在这方面，监督厅指出，将在大会第十六届会议审议秘书长关于执行关于问责制的大会第 64/259 号决议时进一步讨论自我评价问题。方案和协调委员会对这个问题的审议将有助于这个讨论。

主管内部监督事务副秘书长

卡尔曼·拉普安特(签名)

2011 年 3 月 28 日

附件一

两年度研究包括的部门/办事处名单和 2008-2009 年评价报告的数目

部门/办事处	报告数目
1. 经济和社会事务部	2
2. 大会和会议管理部	1
3. 管理部	1
4. 政治事务部	2
5. 新闻部	10
6. 维持和平行动部和外勤支助部	7
7. 安全和安保部	0
8. 非洲经济委员会	20
9. 欧洲经济委员会	13
10. 拉丁美洲和加勒比经济委员会	7
11. 秘书长办公厅	0
12. 亚洲及太平洋经济社会委员会	11
13. 西亚经济社会委员会	4
14. 贸发会议/世贸组织国际贸易中心	9
15. 人道主义事务协调厅	5
16. 裁军事务厅	0
17. 人权事务高级专员办事处	1
18. 最不发达国家、内陆发展中国家和小岛屿发展中国家高级代表办事处	0
19. 法律事务厅	4
20. 外层空间事务办公室	1
21. 非洲问题特别顾问办公室	1
22. 联合国贸易和发展会议	5
23. 联合国环境规划署	8
24. 联合国人类住区规划署	5
25. 联合国难民事务高级专员办事处	11
26. 联合国毒品和犯罪问题办公室	16
27. 联合国近东巴勒斯坦难民救济和工程处	9
28. 联合国日内瓦办事处	2
29. 联合国内罗毕办事处	0
30. 联合国维也纳办事处	0
共计	155
31. 内部监督事务厅	11

附件二

质量评价成果

编号	评价标准	评分 ^a					评分 总数	平均
		1	2	3	4	5		
执行摘要								
1	执行摘要载有报告的关键内容，特别是评价的专题和目标、方法，包括限制因素、主要结果、结论和建议	6	15	11	2	11	45	2.93
引言								
2	报告指出何时进行评价(评价期间) ^b	29				16	45	
3	报告指出何人进行评价 ^c	38				7	45	
引言整体								
4	报告清楚列明评价专题	17	19	5	2	2	45	1.96
5	报告清楚列明评价的宗旨和目标	8	27	9	0	1	45	2.09
6	报告清楚列明评价覆盖和不覆盖的范围(评价范围)	7	18	11	7	2	45	2.53
7	报告清楚列明关键的评价问题和标准	7	19	10	7	2	45	2.51
方法整体								
8	总的来说，评价方法令人信服地解释评价如何得出结果	9	18	13	3	2	45	2.36
9	评价方法说明数据来源、数据收集和分析方法	7	20	13	3	2	45	2.40
10	评价方法应对了各种方法挑战和/或限制因素，以及评价结果的可靠性和真实性问题	10	9	8	13	5	45	2.87
背景								
11.	报告提供足够的背景资料，评价专题的来龙去脉清楚明白(不致太琐碎)	12	15	10	5	3	45	2.38
结果整体								
12.	评价结果清楚涉及评价专题和目标	18	19	5	3	0	45	1.84

注：

如何读表

- 表中的总分都是将评价表中的 45 个个别评分相加的结果，例如“方法整体”：8 份报告的方法部分被评为优(1)，19 份报告的方法部分被评为好(2)，等等
- “平均”(右边最后一列)是唯一的不是将个别评分相加的结果，它是一个算术平均数，计算方法如下：所有计分部分乘以它们的评分再除以评分总数，例如“执行摘要”： $(6 \times 1) + (15 \times 2) + (11 \times 3) + (2 \times 4) + (11 \times 5) / 45 = 2.93$

^a 分数表：1=优；2=良；3=尚可；4=差；5=很差。

^b 29 份报告说明何时进行评价，16 份没有说明。

^c 38 份报告说明何人进行评价，7 份没有说明。

编号	评价标准	评分 ^a					评分 总数	平均
		1	2	3	4	5		
13.	评价结果有充足佐证	16	13	9	7	0	45	2.16
14.	评价结果清楚列明(避免含糊不清)	10	13	12	10	0	45	2.49
15.	评价结果不含评价人员主观判断(客观、根据证据)	9	21	13	2	0	45	2.18
16.	评价结果解释了事情为什么和如何发生和不发生 (成功或困难理由, 促成和限制因素)	8	17	11	9	0	45	2.47
	结论整体	6	9	13	10	7	45	3.07
17.	结论回答了评价的各大问题并聚焦于重大问题	7	12	7	12	7	45	3.00
18.	结论以结果为根据(逻辑联系)	10	13	8	7	7	45	3.29
19.	结论为结果增值(避免重复结果)	4	7	13	14	7	45	3.29
20.	结论反映评价人员的意见和判断(应当评论好坏)	10	14	9	5	7	45	2.67
	建议整体	7	12	15	7	4	45	2.76
21.	建议清楚和直接与结果和建议有关	12	14	10	7	2	45	2.40
22.	建议限于可处理的若干关键建议(避免“罗列-大堆” 应该怎样的建议)	7	8	19	7	4	45	2.84
23.	建议实事求是/可以操作	8	17	14	2	4	45	2.49
24.	报告清楚指明谁应当实施建议	12	7	11	10	5	45	2.76
	格式整体	9	16	13	7	0	45	2.40
25.	报告易读易懂(避免用艰深复杂的字和不作解释的 简称)	10	21	10	4	0	45	2.18
26.	报告一气呵成, 符合逻辑	13	15	10	7	0	45	2.24
27.	报告用相关图表说明要点和重要	10	9	13	6	7	45	2.80
	报告总分	7	12	17	9	0	45	2.62

附件三

秘书处各实体对报告草稿的评论^a

管理事务部

第 44 段

内部监督事务厅在这一段提到 2010-2011 年认列给秘书处监测和评价活动的资源共达 4 990 万美元。该段又提到 2010-2011 年增加了 520 万美元。方案规划、预算和账户厅希望提请监督厅注意秘书长关于 2010-2011 年拟议方案预算的报告的导言(A/64/6(Introduction))。其中第 61 段指出,认列给监测和评价的资源共为 4 880 万美元,比 2008-2009 两年期概算增加了大约 400 万美元。监督厅似可注意到秘书长报告的这个数字。

第 64 段

该段原文为:

“考虑到其所得资源,监督厅提议对所有秘书处方案实行 10 年方案评价周期,但须核可其在 2012-2013 年拟议预算提出的小幅资源增长。”

方案规划、预算和账户厅指出,关于方案规划、预算和账户厅同监督厅在 2011 年 2 月 7 日方案主管审查会议讨论 2012-2013 年拟议预算时曾提到将方案评价周期缩短到 10 年,但经常预算没有提出增加资源以支持缩短周期,因为大多数计划接受评价的方案(即联合国难民事务高级专员办事处、人道主义事务协调厅、联合国环境规划署、联合国毒品和犯罪问题办公室和非洲经济委员会)多半是由预算外资源资助的。因此,按照大会第 50/7 号决议,监督厅打算从各个具体方案的预算外资源谋求进行评价的经费。在这方面,监督厅似可澄清上一段提到的资源供资方面的问题。

监督厅或可在提到秘书处方案主管进行的自我评价时在“评价”二字前加上“自我”两字,以免与监督厅负责的中央评价混淆(见第三.E 节,秘书处方案一级评价做法和能力概况;以及第五节,提高评价报告质量的行动)。例如,在第 43 段之前的小标题,应在“评价”之前加“自我”两字;在第 47 段之前的标题,应在“实施”前加“自我评价”四字;在第 49 段之后的斜体小标题,应在“方

^a 内部监督事务厅于此全文照录收自管理事务部、西亚经济社会委员会、人道主义事务协调厅、裁军事务厅、人权事务高级专员办事处、联合国环境规划署、联合国毒品和犯罪问题办公室、联合国日内瓦办事处和联合国近东巴勒斯坦难民救济和工程处对关于加强评价作用以及评价结论在方案设计、交付和政策指示中的应用的双年度报告草稿的评论。列入这种资料符合大会第 263/64 号决议的规定,是独立审计咨询委员会首先建议的。总的来说,提出评论的实体都同意监督厅的评价结果和结论。对监督厅报告草案的评论都已酌情纳入最后报告。因此,收自秘书处各实体的评论所提的段数与最后报告的段数有几处不一致。

案一级”后加上“自我评价”四字；在第 78 段最后一句，应在“评价”二字前加“自我”两字，以免混淆。

第 85 段的建议案文：

“在这方面，监督厅注意到将在大会第六十六届会议主要部分审议秘书长关于执行有关问责制的大会第 64/259 号决议的报告时进一步讨论自我评价问题。”

西亚经济社会委员会

西亚经济社会委员会希望提出下列问题供你考虑：

(a) 在图四，西亚经社会预算建议的评价资源被列为 4.43%，我们认为很可能是一个错误。根据贵我办公厅之间的讨论，我们希望正式要求加以审查，因为根据我们的计算，这一比例与 2008-2009 两年期差不多，约为 0.69%；

(b) 关于第 54 段，注意到有 10 个部门设有独立评价单位或独立的评价功能，最好在注脚中全部列出这些实体。

我们期待着阅读最后报告。

人道主义事务协调厅

感谢你让大家分享关于加强评价作用以及在方案设计、交付和政策指示中应用评价结果的报告草稿。人道协调厅希望澄清几个事实：

在第 6 页，第 10 段，报告指出人道协调厅仅在 2008 年和 2009 年进行内部评价。原因可能是大家对何谓内部这个术语的理解不同。不过人道协调厅认为，所有在 2008 年和 2009 年进行的评价(包括机构间实时评价和中央应急基金的评价)都是外部评价——因为它们都是对外授权和由外部(而不是内部)顾问进行的。

第 10 页，第 27 段，报告指出，人道协调厅对应对气旋纳吉斯的评价发现“分组方式使协调效率低，所提供的服务有时是多余的。”这是不正确的，评价报告没有这样说。根据关于气旋纳吉斯的机构间实时评价的主要结果，^b 在中央一级的协调被认为是相当不错，但在各分组与它们在实地对口单位联系方面、在外联受惠人方面、以及在分组间规划和协调方面，是有弱点。

因此，我们强烈建议将“分组方式使协调效率低，所提供的服务有时是多余的”这一句改为“在中央一级的协调相当不错，但在各分组与它们在实地对口单位联系方面、在外联受惠人方面、以及在分组间规划和协调方面，是有弱点。”

^b 随函附上的评价报告已由监督厅删除，可到以下网址查阅：http://ochanct.unocha.org/p/Documents/Final%20report%20of%20the%20Inter_Agency%20Real-Time%20Evaluation%20of%20the%20Humanitarian%20Response%20to%20the%20Pakistan%20Floods.pdf.

在第 16 页，报告指出，2010-2011 年人道主义事务协调厅总方案资源分给评价的资源是 0.41%(图四)。我们不知道这个数字是如何计算出来的(也许这只是反映了人道协调厅经常预算分配给评价的一部分?)。然而，2010 年人道协调厅花了大约 1%资源用来进行评价，主要是通过预算外捐款。因此，报告中的数据错误地显示人道协调厅花在评价的支出大大低于 2008-2009 年期间的经费。虽然该报告指出这些数字是根据对 2010-2011 年预算专册的分析算出来的，人道协调厅希望监督厅进一步澄清这些数字是如何确定的。

裁军事务部

裁军事务部认为，由于本部的自我评价报告未列入报告草稿附件一(因为它们不属于所覆盖的时间范围)，那么在该报告第 14 页提到自我评价活动的结果是不通的。因此，为了一致起见，裁军事务部宁愿不要提到这一点。

人权事务高级专员办事处

人权事务高级专员办事处(人权高专办)极感兴趣地阅读了监督厅关于加强评价作用以及在方案设计、交付和政策指示中应用评价结果的报告草稿。关于反馈要求，谨请注意人权高专办现阶段没有任何特别意见。我们期待着收到最后报告。

联合国环境规划署

文件的用意是综述联合国秘书处评价活动的关键方面，它很有可能成为评价职能的工具，也是向大会汇报的一种手段。报告易读易懂，但需要考虑到以下各点，才能改善质量，提高向预定目标受众提供反馈的效率：

(a) 方法一节只提供了关于“提供评价报告关键结果、结论和建议的摘要”这个公开目的的有限资料。没有提战略背景。用什么参数/标准界定哪些来自经审议报告的一般结果、结论和建议值得注意？需要清楚说明用何种指引选择与战略背景有关的结果摘要；

(b) 接着上一点，结果摘要的标题没有清楚说明理由。在秘书处一级，每个专题的选定评价提出的具体结果的一般相关性同样不清楚。读者无法清楚知悉选定结果对秘书处工作的影响；

(c) 第三节，第 13 段指出大多数评价都在部门或方案以下一级进行。这是令人惊讶的，因为根据预设，评价资源往往已安排给项目活动。这种模式也在预料之中，因为项目绝对数目(必然)比方案/部门数目多很多。关键要点是，有没有将足够资源/努力分配来进行具有战略重要性的评价(这往往是高一级的活动)；

(d) 第 39 段指出，许多评价 (55%) 没有就政策指示作出建议。不过，报告没有探讨所出现的政策指示建议与进行评价的类别或层级有何联系。人们可以随便预期，项目一级的评价作出这种建议的频率较低，因此项目评价的相对频率可解释这种模式；

(e) 从方法角度看，我们不知道图二显示的 71 个项目的模式能否反映所有 155 份评价报告的情况，又或评价类别的分布会不会因选定标准而改变。图二如能列出所有 155 个项目，提供更多的资料，就更加有用；

(f) 没有说明计算列于第三.E 节图四的监测和评价预算数字的根据。这些数字包括什么，不包括什么，都不清楚，这样就使数据的解释成为问题。这些数据大概来自调查答复。如果答复者从评价职能的角度汇报，监测预算就可能不包括在内，又如果答复者从监测和评价职能的角度汇报，汇报的预算可能就涵盖两套活动。所用的方法需要清楚说明，将来的研究应确保适用共同的标准；

(g) 第 8(a) 段业已表明，我们期待收到对经过监督厅审评的环境署评价报告质量的回馈意见。

经济合作与发展组织发展援助委员会最近完成的题为“发展机构评价”的比较研究介绍了评价职能是如何管理和提供资源的。这份研究报告是一个有用的范本，可供监督厅以后编制这份报告时参考。

联合国毒品和犯罪问题办公室

一般评论

报告提出了非常令人感兴趣的结果和结论，特别是关于接受评价为一项管理职能这一点。

报告如果有一个关于建议和行动的段落，比第五节所载的一段更为详细，将大有帮助。虽然报告第五节列出了具体的行动，它集中关注的仅仅是加强评价报告的质量，而不是如何使评价被接受为一种管理职能。

目前尚不清楚实际审查了各部/厅/方案多少评价报告。附件一最好多列一栏，标明 2008-2009 年报告中获审查的数目。这将澄清关于本组织八个优先领域的结果、结论和建议的第三.B 节。在目前的报告中，似乎联合国毒品和犯罪问题办公室只有一项评价获审查——这可以作为一个例子，但不能反映禁毒办整套评价工作。

具体评论

关于采取行动提高评价报告的质量的第五节，谨请注意考虑下列各点：

(a) 提高术语的共同理解：背景资料和证据、证据和结果、结果和结论之间都有混淆；

(b) 应设法解决如何将两性平等和人权纳入评价报告主流的问题。

因此，联合国毒品和犯罪问题办公室建议第 81 段 (b) 提到的评价报告模板列入定义以及将两性平等和人权主流化的指引。

第 42 段列出了所有方案，联合国毒品和犯罪问题办公室除外。

第 75 段提到，“监测和评价的[……]资源数额远远低于一般推荐下限”。最好说明这个下限是如何定义的。这些信息将有助于建立一个机构评价文化，联合国毒品和犯罪问题办公室致力于创建这种文化。

联合国日内瓦办事处

我代表联合国日内瓦办事处总干事感谢让我们分享报告草稿并征求我们的意见。我们认为，报告草稿清楚描述了整个秘书处的评价活动，精确点出了进一步加强评价作用必须解决的关键问题。我们特别欢迎给方案和协调委员会提出的问题。

联合国近东巴勒斯坦难民救济和工程处

联合国近东巴勒斯坦难民救济和工程处审查了报告草稿，谨表扬监督厅全面和清楚地分析了联合国系统的评价活动。近东救济工程处同意并赞赏报告草稿。