



经济及社会理事会

Distr.: General
18 April 2016
Chinese
Original: English

供作决定

联合国儿童基金会

执行局

2016 年年度会议

2016 年 6 月 14 日至 17 日

临时议程* 项目 7

2015 年关于儿基会评价职能的年度报告

摘要

本报告记载了儿基会在 2015 年开展的评价活动。报告审查了儿基会评价职能的现况，并报告了订正评价政策的执行进展情况。报告审视了儿基会评价职能的管理情况，还审视了旨在促进联合国系统内和更大范围内评价一致性的努力。报告提供了有关绩效和成果以及用于评价的人力资源现况的最新资料。报告审查了经选定的评价的影响，说明儿基会各小组如何使用评价的情况。报告还载有供执行局审议的决定要点。

* E/ICEF/2016/5。



一. 引言

1. 联合国秘书长潘基文在谈及评价工作为 2015 年后议程服务的专题时指出，“无论是在哪个领域和在哪个级别进行的评价，都将在执行新的发展议程中发挥重要作用”。秘书长是在联合国评价小组 2015 年 3 月在纽约主办的启动评价活动周的会议上说这番话的。他接着说：“评价并非易事，也不受欢迎，但它至关重要。当前预算拮据的环境使评价比以往任何时候都更重要”。
2. 儿基会认识到评价在支持组织问责、学习和不断改善绩效方面的重要作用。儿基会还确认其订正评价政策(E/ICEF/2013/4)规定的如下重要职责：与各级发展伙伴就评价活动密切合作，并协助发展评价能力。评价有助于儿基会展示成果、促进问责并为更好的政策和方案提供证据，从而直接促进实现儿基会的使命、任务和优先事项。
3. 本报告记载了儿基会在 2015 年开展的评价活动。报告审视了儿基会评价职能的现况，并报告了订正评价政策的执行进展情况，同时提供了关于绩效和成果的最新资料。同前几年一样，本报告审视了对儿基会职能的管理情况，还审视了旨在促进联合国系统内和更大范围内评价一致性的努力。今年注重了记载经选定的评价的影响，以便说明儿基会各小组如何使用评价的情况。报告最后载有供执行局审议的决定要点。

二. 对评价职能的管理

4. 执行局坚持对儿基会评价职能进行监督，并维持它在儿基会的核心作用。
5. 2015 年，评价办公室编写并向执行局提交了 2014 年评价职能年度报告。高级管理层还编写了说明管理观点的配套文件。对此，执行局通过了一项正式决定(第 2015/10 号决定)，其中注意到取得的进展，并吁请儿基会进一步加强这项职能。
6. 评价办公室还向执行局提交了 4 份评价报告，每份报告都附有管理层的回应：
 - (a) 对儿基会应对菲律宾台风“海燕”的工作的实时评价；
 - (b) 对 2003-2012 年儿基会在教育和性别平等方面开展的上游工作的评价；
 - (c) 对公平结果监测系统的形成性评价；
 - (d) 对题为“作为社会保护措施元现金转移：2010-2014 年儿基会评价提供的证据”的综合评价。
7. 全球评价委员会在儿基会发挥着重要的协商职能。该委员会在 2015 年 6 月举行了会议，审议了关于加强评价职能和修订全球专题评价计划的建议。在审议

了后述的建议后，儿基会在各部门开展了一次协商进程，促进制定了 2016-2017 年最新全球专题评价计划(E/ICEF/2016/3)。该计划已于 2016 年 2 月提交执行局。

8. 外部审计咨询委员会将评价视为监督系统的重要部分，并监测评价职能的绩效。委员会在 2015 年提交执行局的 2014 年年度报告中对中央和实地评估稳定的质量表示满意，并建议管理层进一步改进它对评价的响应的及时性和质量。¹

9. 评价办公室正在继续促进理解和执行订正评价政策，并为此编写了两页摘要广为分发。²

三. 联合国系统内评价的一致性

A. 大会的决定：

10. 2015 年 9 月 24 日，大会通过了一项题为“改变我们的世界：2030 年可持续发展议程”的目标宏伟和富有远见的新发展议程。³ 必须通过系统地收集数据和由国家主导的评价，定期后续和审查该议程的实施工作。大会该决议呼吁在必要时提供援助，以加强国家评价能力。⁴

11. 提及加强国家系统的依据是大会在前一年通过的有史以来第一项关于评价的决议(第 69/237 号决议)。大会该决议邀请联合国各实体应会员国的请求并根据它们的国家政策和优先事项，支持它们为进一步加强其评价能力而作出的努力。

12. 2012 年通过的联合国系统发展方面业务活动四年度全面政策审查(第 67/226 号决议)也指出必须加强评价能力。四年度全面政策审查强调了评价的重要性和必须使用评价证据，并促请联合国发展系统各成员“加紧努力，协助方案国提高其监测和评价发展方面业务活动的国家评价能力”。

13. 2015 年，儿基会致力于支持各国发展国家评价能力，以便达到这些决议提出的要求。儿基会特别强调支持制定国家评价政策和框架，并支持进行国家主导的评价的国家能力。一种有效的办法是将发展评价能力与高度优先政策问题相挂钩。例如，菲律宾政府在 2015 年 7 月核可了一项国家评价政策框架，据以配合最近进行的规划和预算拟订工作改革。该框架是通过儿基会和其他伙伴提供的技术援助、培训和便利条件制定的。柬埔寨政府要求协助制定准则，据以建立国家监测和评价系统，以支持及时监测和评价国家战略发展计划，儿基会正在以类似的方式对此作出回应。

¹ www.unicef.org/about/execboard/files/UNICEF_AAC_2014_Report-16Apr2015.pdf。

² www.unicef.org/evaluation/files/Evaluation_Policy_Brief-2pager_Final.pdf。

³ 大会第 70/1 号决议。

⁴ 见第 74(g)和(h)段。

14. 儿基会还同其他伙伴合作，支持议员倡导评价工作并在其工作中使用评价(见第四节)。

B. 联合国评价小组

15. 联合国评价小组是由联合国系统约 46 个组织的评价单位组成的专业网络。联合国评价小组的目的是提高联合国系统各部门的评价一致性，并据此支持通过下述途径发展全系统的评价一致性：实施商定规范和标准；发展和交流工具及指导；进行成员间同行审议；不断加强外联活动，并与联合国系统内外的评价倡议结成伙伴关系。

16. 近年来，联合国评价小组在上文提到的大会决议方面发挥了宣传和协调作用，并成为经 2012 年四年度全面政策审查授权对联合国发展方面业务活动进行独立的全系统评价的临时协调机制的成员。

17. 因此，联合国评价小组在促进全系统评价一致性方面发挥着日益具有影响力的作用。儿基会为此继续投资于联合国评价小组的活动并往往主导这些活动。2015 年，评价办公室主任被任命为联合国评价小组副主席，儿基会因而进一步发挥更有力的领导作用，对联合国评价小组的伙伴关系议程负有特别的领导职责。

18. 儿基会在安排和主办 2015 年 3 月在纽约举办的联合国评价小组评价周方面发挥了重要作用。评价周的活动包括：上文所述的高级别会议，秘书长在该会议上谈到了评价在新发展议程中的作用；儿基会和全球环境基金在儿基会大楼共同主持的评价做法交流活动，与会者认为这次交流非常成功；以及 2015 年年度大会。

19. 联合国评价小组通过各种机构间工作组进行工作。2015 年，联合国评价小组首次设立了一个人道主义评价关注小组。该小组的目的是(a) 促进人道主义评价；(b) 确保充分考虑到联合国评价规范、规则和准则规定的有关标准；(c) 改进联合国实体所作人道主义评价的质量。儿基会是该小组的组成部分。第一次联合活动是对人道主义原则进行案头审查，其目的是更好地了解如何评价人道、公正、中立和独立等核心人道主义原则，并强调良好的做法、挑战和机遇。

20. 2015 年 11 月，联合国评价小组与另一个评价网络国际评价合作组织正式结成伙伴关系，使全球各地的国家评价协会相互联系。联合国实体、国际评价合作组织以及其他组织加在一起被称为“评价伙伴”(EvalPartners)，它们之间的非正式联盟关系在过去几年中倡导更多地使用评价和加强评价能力，从而确立了它本身拥有重大影响力的地位。更多的详情见第四节。

C. 机构间评价

21. 联合国实体联合进行的机构间评价为加强评价的一致性提供了重要渠道。这些评价采用的形式各异，儿基会在每次评价中发挥了积极的支助作用，包括提供技术和管理投入以及若干财政捐助。

22. 现已按照经大会在 2012 年四年度政策审查中授权建立的独立的全系统评价机制启动了两次联合评价。这一机制正在在包括联合国评价小组的若干行为体参与下检验一种旨在使对全系统问题的独立评价具有良好基础的办法。儿基会同其他组织一起提供了财政和技术支助。为试点评价选定了以下两个专题：

(a) 对《联合国发展援助框架》各次评价的元评价和汇总，以消除贫穷为特别重点。该评价即将在 2015 年年底完成。

(b) 评价联合国发展系统对提高各国统计分析和数据收集能力以支持实现千年发展目标和其他国际商定发展目标的贡献。评价的最初报告已在 2015 年年底完成。

23. 对“再次努力消除儿童饥饿和营养不良现象倡议”专题评估的评价提供了各种证据和备选办法，以供指导进一步的行动。⁵ 这次评价是由世界粮食计划署(粮食计划署)为牵头机构，随同联合国粮食及农业组织(粮农组织)、世界卫生组织(世卫组织)以及儿基会联合进行的。

24. 2015 年，机构间人道主义评价小组对联合国应对重大人道主义紧急情况的工作进行了两次联合评价。机构间人道主义评价小组由人道主义事务协调厅(人道协调厅)协调，在机构间常设委员会(机构间常委会)主持下开展工作。机构间人道主义评价小组致力于就与人道主义问题相关的评价事项采取协调一致的行动。在 2015 年年底几近完成的机构间人道主义评价分别涉及应对南苏丹冲突和中非共和国危机的工作。

25. 这些机构间人道主义评价为联合国系统提供了重要的调查结果。诸如南苏丹等旷日持久的紧急情况通常植根于长期冲突和欠发达的格局，因此，各种方案应该支持涉及一系列救济、恢复、发展和应对能力的努力；导向目标必须应对大量可能难以与之联系的人群的需求；而且为了增强可持续性，必须使国家和地方利益攸关方密切参与应对工作。

四. 评价的一致性：全球

A. 发展网络

26. 2012 年，儿基会和另一个组织共同创建了全球网络“评价伙伴”。该网络的目的是加强创新伙伴关系，据以提高民间社会组织的评价能力，对决策者产生影响，并倡导建立更强劲的国家评价系统。⁶ 另一个创建伙伴是国际评价合作组织，

⁵ www.wfp.org/content/faowfpunicefwfpwhodfatd-canada-joint-evaluation-renewed-effort-against-child-hunger-and-unde。

⁶ www.mymande.org/evalpartners。

这是一个使全球各地的国家评价协会相互联系的网络。许多机构已加入“评价伙伴”，从而促成实施了一系列广泛的评价能力发展举措，包括开展同行互动，以促进在各发展评价协会中开展南南合作和三角合作。2015 年进行的对“评价伙伴”活动的评价确认，该网络已确立了它本身在全球各地评价活动中拥有重要影响力的地位，同时确认了若干个需要关注的领域，以便“评价伙伴”予以进一步发展。⁷ 如上文所述，联合国评价小组和“评价伙伴”已经正式结成核心伙伴关系，共同主持这一伙伴关系工作的是“评价伙伴”主席和联合国评价小组主管伙伴关系的副主席，后者目前还担任儿基会评价办公室主任。

27. 2014 年，“评价伙伴”成员指定 2015 年为“国际评价年”，并在全球各地举办了 90 次活动，最终促成在尼泊尔举办了一次极其成功的评价周活动，使评价人员、学者、活动分子、官员、政府部长和议员聚集一堂参加讨论评价的会议。这些会议包括两年一度的“全球评价伙伴论坛”，尼泊尔总理在该论坛发表了讲话，此外还包括在尼泊尔由议会议长主持举行的具有里程碑意义的大会。

28. “全球评价伙伴论坛”的重要成果是最后敲定并核准了 2016-2020 年《全球评价议程》。该《议程》的重点是评价能力发展和专业化，它涉及评价人员个人和关注评价工作的机构，也涉及关注委托进行、授权开展或资助评价活动从而为评价创造有利环境的机构所在的范围广泛的领域。儿基会是整个论坛期间的主要捐助组织，并正在安排以“评价伙伴”的名义出版《全球评价议程》。

29. “全球论坛”还目睹在“评价伙伴”内启动了若干个小网络，其中包括“ValSDGs”，其目的是倡导对在《2030 年议程》框架内实现《可持续发展目标》的进展情况进行评价。“评价伙伴”的另一些小网络包括：“EvalGender+”，其重点是围绕公平和两性平等进行评价；“EvalYouth”，其目的是增强青年评价人员的技能和知识；“EvalIndigenous”，其目的是加强土著人民参与对他们有影响的发展进程的规划和评估工作。儿基会在“EvalSDGs”中发挥主导作用，并正在参与其他网络的工作。

30. 在相关的活动中，儿基会还与“评价伙伴”合作，使议员支持评价活动，并鼓励他们在工作中使用评价。在若干区域设立了一个议员论坛，并在 2015 年 11 月启动了“全球议员发展评价论坛”。

31. 儿基会正在继续支持发展和更新一个称为“My M&E”(“监测和评价”)的全球评价网络平台，该平台与“评价伙伴”有关联。儿基会已促成提供可通过“My M&E”获得的学习资源，特别是最后拟定并启动了一门以阿拉伯语讲授的电子学

⁷ [www.mymande.org/sites/default/files/files/EvalPartners-Evaluation-Executive-Summary-\(2015_02_03\).pdf](http://www.mymande.org/sites/default/files/files/EvalPartners-Evaluation-Executive-Summary-(2015_02_03).pdf)。

习课程，其主讲者为 33 位知名评价专家。⁸ 儿基会还在继续支助数目日增的有助于提高评价能力和一致性的全球、区域和国家评价会议和闭门会议。

人道主义网络

32. 在评价人道主义行动方面，儿基会在机构间人道主义评价小组中以及在启动联合国评价小组人道主义评价关注小组方面发挥了重要作用。此外，儿基会还在继续与“人道主义行动问责和绩效动态学习网络”合作。2015 年，该网络的主要成果之一是编写了《评价人道主义行动试点指南》的最后文本和开设了以《试点指南》为基础的配套在线培训课程。培训课程开设以来，已有 2 200 余人参加了培训。

33. 人道主义行动问责和绩效动态学习网络的《指南》还为由评价办公室在 2015 年启动的一项举措提供了基础，该举措的目的是建设儿基会国家办事处评价人道主义行动的能力。由于儿基会在国家一级进行的人道主义行动评价为数不多，这一举措采用了实际参与的办法，向致力于开展此种评价的国家办事处提供培训和支助。更多的详情见第四节。

五. 儿基会的评价职能：重要的绩效信息和分析

全球评价报告监督系统

34. 本节所述信息主要来自全球评价报告监督系统。该系统提供：关于国家、区域和全球评价报告的独立评估和评级；提交各办事处的关于报告质量的反馈意见；纳入管理部门信息表的绩效信息汇编。全球评价报告监督系统质量评级的依据，是外部审查员利用以联合国评价小组标准为基础的准则进行的审慎和有系统的评估。

35. 自 2010 年建成以来，全球评价报告监督系统已多次向执行局报告了为改进和加强该系统而进行调整的情况，最近的一份报告是在 2015 年提交的(见 I/ICEF/2015/10)。2015 年进行了一项全面的外部审查，审查认为全球评价报告监督系统十分重要，因为它有助于儿基会履行评价职能，有助于在区域和国家两级进行质量控制，也有助于儿基会的决策。通过全球评价报告监督系统的反馈，评价能力已有提高，评价报告质量也有所改善，因此不言而喻，评价进程的质量多年来不断得到改善。为了取得更好的成果，在决策时已利用提交管理部门和执行局的关于评价质量的报告来加强各种系统和进程。评估还认为，全球评价报告监督系统的实施费用合理，而且该系统拥有相当大的潜力，可据以进一步加强儿基会的知识管理工作。

⁸ www.mymande.org/elearning。

36. 评估再次确认了全球评价报告监督系统的重要性和价值，但仍然提出了若干改进意见。这些建议要求：采用较不复杂的评级工具；更新现有的指导文件；找出被不正确地标记为“研究”或“审查”的评价，并将它们纳入全球评价报告监督系统；制定一项迅速上传评价报告并及时编制审查和反馈意见的政策；改进传播全球评价报告监督系统成果的工作；进一步加强评价办公室与研究办公室间的协调，从而增强一致性；定期审查全球评价报告监督系统，确保它继续与它的目的相符。

37. 评价办公室认为该项审查既相关也很有用，因此它正在与其他办公室协商下编写一项管理层回应。日后将向执行局报告实施商定建议的行动。

绩效概述

38. 归为一套的 6 项关键绩效指标概述了 2015 年的绩效趋势，并辅之以前几年的信息。大多数年份提供了关于全球专题评价计划执行情况的第 7 项关键绩效指标，这些评价都是由评价办公室进行的。本报告未提及这项指标，因为在提交执行局 2016 年第一届常会的《2014-2017 年全球专题评价计划：供 2016-2017 年审查和更新》(E/ICEF/2016/3)中已经报告了进展情况。

39. 根据下节的分析得出了 4 个主要结论：

(a) 儿基会各办事处连续第三年报告业已完成的评价报告不到 100 份，考虑到对评价的财政投入和人力资源基础一直有所增加，这一数字似乎异常。必须作出进一步的审查和分析，了解是哪些因素提高或限制了评价的覆盖率和完成率；

(b) 评价质量继续了 2013 年开始出现的趋势，良好/优质评价的百分比越来越高。质量的提高是随同持续注重较高层面的成果(即在结果和影响层面的成果)的情况出现的；

(c) 评价做法某些方面的弱点继续引人关注，这涉及妥善考虑到道德操守问题。不过与上一年相比，这一指标有了明显改善，对以下两个相关优先事项的评级也是如此：将人权、两性平等和公平问题纳入评价，以及使利益攸关方以适合其能力和利益的方式参与评价工作；

(d) 现在可通过儿基会虚拟集成信息系统跟踪预算使用情况。数据表明，2015 年儿基会用于评价的经费占其方案资金的 0.65%，政策的要求是 1%。与 2011 年的 0.33%和 2014 年的 0.5%相比，此种增加值得欢迎。不过，评价支出水平仍然明显低于目标要求。

40. 必须持续重视下述要求：(a) 提高提交评价数据库的评价报告的评价覆盖率，并提交更多的报告；(b) 改进管理层回应的实施情况；(c) 达到评价支出占方案支出总额 1%的整体政策目标。鉴于儿基会高度分权，要处理这些问题，就必须在儿基会各级进一步加强评价的规划、预算拟订和实施工作。

关键绩效指标执行情况

指标 1：已经管理并提交全球评价数据库的评价数量⁹

41. 2014 年，儿基会各办事处完成并提交了 82 项评价，这明显少于 2013 年提交的 96 项。各区域提交的评价数量不同，最少的一个区域只提交了 3 项评价，东部非洲和南部非洲提交的评价多达 16 项。¹⁰

42. 执行局第 2014/10 号决定要求提供更多关于评价覆盖率的资料。全球而言，74%的国家办事处在过去三年至少完成并向全球评价数据库提交了一项评价，但显然有 35 个(26%)国家办事处未做到这一点。儿基会认识到在某些情况下开展评价有困难，但它必须重视缩小评价覆盖率的差距并履行它在这方面的责任。订正评价政策(E/ICEF/2013/14)明确规定了评价覆盖率的标准。¹¹

指标 2：主题分配

43. 2014 年主题分配细目显示，过去四年中，仅涉及单一部门的评价百分比差别很大(见表 1)。主要部门年年不同：2013 年是营养，2014 年是教育。鉴于所涉数目相对较小，这是正常的波动。

44. 有两个具体结果值得注意。第一，只有 6%的评价注重于人道主义方案拟订工作。这与人道主义行动支出水平很高(而且越来越高)的情况不相称。第二，所有经审查的评价中约有 9%特别注重于性别平等，此外还在部门评价中经常重视两性平等。这是很好的结果，并维持了可见于在 2012 年完成的评价中同样积极的结果。

⁹ 同其他指标一样，关键绩效指标中的数字是指在 2014 年完成并在有数据可查的最近一年即 2015 年评估的评价数据。评价通常在年底提交，本报告于 2016 年初编写，因此没有对 2015 年完成的评价的分析。

¹⁰ 在提交的 82 份报告中，有 13 份报告迟交。表 1 关键绩效指标中数据的依据是 69 份按时提交供全球评价报告监督系统审查的分析。

¹¹ 见第 49 段：评价工作将通常在以下情况下进行：(a) 在参照或推广方案之前……；(b) 在应对严重人道主义紧状况之时……；(c) 在长期未对方案实施情况进行评价之后……；(d) 用于每个方案构成部分结果的评价支出已达到 1 000 万美元。

表 1
2010–2014 年评价报告主题分配

主题 ^a	2010 年 基准百分比	2012 年 百分比	2014 年 百分比
针对具体部门的主题，按《2014-2017 年战略计划》	66	41	66
关键成果领域			
儿童生存和发展	25	26	23
保健			13
营养			0
水、环境卫生和个人卫生			10
教育	15	13	32
儿童保护	15	8	6
艾滋病毒/艾滋病	10	5	4
社会包容	1	1	1
覆盖若干部门的方案	34	43	16
贯穿各领域的主题	—	10	9
组织绩效/其他	—	6	9

资料来源：2011-2015 年全球评价报告监督系统的报告。

^a 部门名称现与儿基会《2014-2017 年战略计划》所用名称相符。鉴于新老名称很一致，这些数字可与 2014 年前的数字直接比较。

指标 3：已进行的评价的类型

45. 主要以产出层面为对象的评价所占的百分比如预期持续下降。产出层面的成果通常应该通过监测和审查进程予以评价。评价应该注重更高层面和更复杂的成果。然而，影响层面评价的比例在 2014 年急剧下降，这一数字表明必须进行监测，确保无论是通过评价或以其他方法使方案影响逐步得到充分证实。

46. 形成性评价的百分比，也就是为评估执行工作是否顺利开展而在方案周期初进行的评价的百分比，继续维持最早在 2012 年所见的较低水平(见表 2)。这可能表示，多层面和实时方案监测办法向管理人员提供的数据开始比传统方案监测办法提供的数据更为丰富，从而降低了对形成性评价和产出层面评价的需求。

表 2
已进行的评价的类型

评价类型	2009 年	2012 年	2014 年
按经审查的方案规划结果 ^a			
产出层面评价	33	18	15
成果层面评价	24	30	68
影响层面评价	43	52	18
按管理意愿 ^a			
形成性层面评价	45	24	29
总结性层面评价	55	76	71

^a 总结性评价通常也包含一些形成性部分，影响层面评价也可能涉及产出层面问题。本表只记录这些评价的最初目的。

指标 4：儿基会评价报告质量

47. 质量评级近年来呈现出积极的趋势，这一趋势在 2014 年仍在继续(见表 3)。2014 年的数据显示了迄今最高的满意/优质评级和最低的低劣评级。这是连续第三年的出色成果。全球评价报告监督系统评级人员注意到了 2014 年具体的积极成果：88%上传的评价报告附有职权范围，前几年仅为 57%。这使评级更为准确，并通过与考虑对类似主题进行评价的办事处分享工具，支持全球评价报告监督系统的知识管理职能。

表 3
2009–2014 年已完成的评价的质量

质量评级	2009 年	2012 年	2014 年
对行动置信(良好或优质)	36	62	74
对行动几近置信(几乎满意) ^a	34	30	23
低劣	30	8	3

^a 选择“对行动几近置信”是因为报告内容显然准确，对评价的管理也正确，但因为存在一些缺点，不能充分置信。问题往往是报告的说明不充分，或者虽然做法总体良好，但在一小部分工作中存在着缺陷。

48. 总体质量评级是许多具体评级的综合。其中两项评级，即关于利益攸关方的参与和人权/性别/平等的评级，是将机构承诺主流化的敏感指标(见表 4)。2014 年，这两项指标都有了重大改善。

49. 作为对将两性平等问题适当纳入评价的更细致分析的组成部分，儿基会目前正在采用《联合国全系统性别平等和增强妇女权能行动计划》准则，每年对评价

内容进行一次具体审查。儿基会评价报告超过四分之一达到(14%)或超过(13%)要求, 45%接近达到但尚未达到要求, 28%没有达到要求。这种情况虽不理想, 但大大高于 2013 年的基线, 该年的报告 5%达到要求, 60%接近达到要求, 35%未达到要求。

表 4

将某些机构重点纳入已完成评价的主流: 2010–2014 年

机构重点 ^a	达到“良好”或“优质”评级的报告百分比		
	2010 年	2012 年	2014 年
适当纳入人权、性别平等和公平问题	18	44	60
使所有利益攸关方以适合它们的能力和利益的方式参与	40	48	66
道德操守问题和考虑因素以及适当的道德操守保障	10	32	49

^a 根据联合国评价小组的评价质量标准, 选自每份报告中经过分析的内容。

50. 儿基会在先前的报告中承诺将通过各种举措, 更加重视在管理它诸如评价和研究等证据职能方面的道德操守问题。2014 年的筹编工作促成在 2015 年公布了关于在评价和研究以及相关职能中保障道德操守的指南。儿基会各办事处显然了解到关于道德操守的讨论以及加强监督的必然性。2014 年 49%的百分比虽然尚不令人满意, 但它大大高于 2010 和 2012 年的数字。这将继续得到重视。

指标 5: 评价的使用, 包括管理层回应

51. 使用是一项关键的评价原则。由于评价质量有所改进, 现在有很好的理由切实使用评价成果, 以此利用投资于评价的时间和资金。管理层的回应将评价建议变为商定行动。儿基会订正评价政策要求为每项评价制定一份管理层回应。2009 年启动了由一个数据库、指导、培训、监测等组成的评价管理层回应系统, 其目的是通过明确管理责任和对商定建议采取后续行动, 加强对评价的使用。这方面有两个关键阶段: (a) 初步回应, 详尽说明办事处对评价建议所作的一整套承诺, (b) 随后提供资料, 说明在 12 个月的回应实施期间实际采取的行动的最新情况。这方面的目标是 95%的承诺得到充分履行(因为认识到在有些情况下可能证实无法履行所有的承诺)。

52. 2014 年向数据库提交初步回应的比率从 2013 年的 92%降至 81%, 但仍然远远高于 2009 年 10%的基准。到 2015 年年底, 2014 年评价的商定行动(包括已经实施和实施中的行动)的实施率为 76%, 并有望在一两年后已经实施和实施中行动的比例达到前几年达到的 86%。但是, 与这一总体的积极趋势同时存在着另一种趋势。过去四年期间, 约有 31%的商定行动在完成评价后两年或更长时间内依然被列为“正在实施”。还有 14%的商定行动或是已被取消(2%)或是从来就没有

开始(12%)。因此加在一起,约有 45%的管理层回应承诺没有在规定期限内被标注为“已经完成”。这是不佳的绩效。

53. 为了了解出现这些数字的原因,评价办公室委托一个外部顾问小组进行一次审查。这次审查于 2015 年年底开始,目的是确定评价管理层回应系统是否已满足了儿基会的需求,特别是审查评价管理层回应系统、该系统的总体效益以及可能加以改进的领域。

54. 正在进行的审查得到的结果表明,对评价管理层回应系统的指导通常明确、适当,也符合儿基会订正评价政策。评价管理层回应进程的总体优势包括支持问责制、透明度和监测进程。不过,评价管理层回应跟踪系统似乎没有充分涵盖所有正在为回应评价建议而作出的行动。儿基会许多工作人员认为评价管理层回应有助于更好地规划和决策,但较少的工作人员认为评价管理层回应有助于加强知识交流。虽然儿基会的评价管理层回应系统与参照组织的评价管理层回应系统类似,但如能进一步明确各种作用和责任、增强支助、提供验证证据并提高伙伴的参与度,评价管理层回应的总体进程将会得到改善。审查还建议加强对遵守要求的情况进行监督。

55. 评价办公室认为,同对全球评价报告监督系统的审查一样,对评价管理层回应系统的审查既相关也很有用。评价办公室将在与其他的办公室协商后编写对这一审查的管理层回应,并在适当的时候向执行局报告实施商定建议的行动。

指标 6: 整个组织用于评价的支出

56. 儿基会每年向执行局报告整个组织用于评价的支出,直到 2011 年为止,因为儿基会财务管理系统在该年有所变更,致使数据变得不具可比性。但自 2014 年以来,在采用了虚拟集成信息系统的财务管理系统后,就又可以提交有关这一指标的报告。2015 年,虚拟集成信息系统的数据显示,儿基会方案支出总额为 45.05 亿美元,为儿基会评价活动支出或承付的总额为 2 920 万美元。2014 年的评价支出总额为 1 890 万美元。

57. 2015 年的数据表明,儿基会专用于评价的支出占其方案预算支出的 0.65%,比 2014 年的 0.5%提高了 30%。此外还表明,在落实将评价支出至少提高到占方案支出总额 1%的政策目标方面取得了显著进展。不过,对照政策目标,儿基会用于评价的支出不足。在 7 个区域中,有一个区域超过了 1%这一门槛值(拉丁美洲和加勒比达到了 1.2%),其他区域的最低支出百分比为 0.1%和 0.2%。儿基会总部的评价支出达到 2%,促成其提高的部分原因是从区域向总部调拨了用于评价重大紧急情况的经费。

六. 儿基会的评价职能：人力资源

58. 有效评价的先决条件是提供充分的人力和财政资源。本节对用于 2015 年评价的人力和财政资源的相关信息与至少两年的基准数据进行了比较，以了解这段期间的趋势。

59. 截至 2015 年 12 月，设有包括职称为评价人员的专业人员员额的国家办事处有 106 个，2013 年有 100 个，2008 年有 85 个。儿基会所有 7 个区域办事处和 4 个总部地点(纽约、日内瓦、哥本哈根和佛罗伦萨)都设有此种员额。表 5 表明 2008 至 2015 年期间评价专业人员数目的变化情况，其中明显地分列了第 3 级及以上级别人员。¹² 64% 以上的评价职位属于国际专业人员职类。

表 5

2008-2015 年儿基会评价专业人员数目增加情况

	2008 年 3 月	2013 年 12 月	2015 年 12 月	增加 2013-2015 年	增加 2008-2015 年
第 3 级或以上级别国际评价专业人员	50	95	118	+24%	+136%
第 3 级或以上级别国家评价专业人员	40	63	66	+5%	+65%
第 3 级或以上级别评价专业人员 (共计)	90	158	184	+16%	+104%
第 1 级或第 2 级评价专业人员	没有报告	90	95	+6%	无
外地办事处评价专业人员百分比	84	90	94	+4	+12
第 3 级或以上级别女性评价专业人员百分比	42	50	48	-4	+14
第 1 级和第 2 级女性评价专业人员百分比	没有报告	40	36	-10	没有报告

资料来源：人力资源司各年数据。

60. 过去两年期间，2008 年以来呈现的积极趋势有所加强。较高级别的员额增加了 24%，这有助于儿基会履行它提高评价技能的机构承诺和提供在这一领域的职业发展机会。性别均等的目标在较高级别员额中已几近达到，在第 1 级和第 2 级国际专业人员中已经达到。儿基会还一直在利用来自工业国家和方案国家的专业

¹² 第 3 级是期望拥有广泛评价能力的最低级别，但并未规定这些员额必须是该级别或以上级别。

人员的才能,并在国际专业队伍中实现了 50/50 的均衡(其中 65 人来自工业国家, 64 人来自方案国家)。140 名本国干事来自方案国家。¹³

61. 第 3 级或以上级别的人员配置代表了提供优质评价的能力。设有此类员额的办公室从 2013 年的 78 个增加到 2015 年的 84 个。另有 24 个办事处仅设有第 1 或第 2 级评价工作人员。大约 20% 的国家办事处没有专设评价工作人员。

62. 本两年期间的员额增加率(24%)远远超过设有这些员额的办事处数目的增加率(8%)。造成这种差异的主要原因是资源较充足的方案办事处一直在扩大它们的监测和评价人员队伍。三分之一的办事处设有 3 名或 3 名以上评价专业人员,它们包括从很大的办事处(如驻埃塞俄比亚和驻印度的办事处)到中等规模的办事处(如驻玻利维亚和驻菲律宾的办事处)不等。许多设在受紧急情况影响的方案国家(如利比里亚、南苏丹、阿拉伯叙利亚共和国和也门)的办事处设有 3 个或更多的监测和评价工作人员,在这些办事处可看出儿基会对提供人道主义援助的承诺。此外,儿基会整整四分之一的评价职位设在属于 D 类的艰苦工作地点(占监测和评价员额的 10%)以及属于 E 类的极端艰苦工作地点(占此类员额的 16%)。

63. 此外还通过集中管理的新的和新出现的人才计划(“新人才计划”)填补了两个第 3 级评价员额,使按“新人才计划”所设的监测和评价员额总数在 7 年内达到了 19 个。在 12 名已结束为期两年首次任用的青年专业人员中,有 7 人在一个方案办事处获得了全职职位,从而达到了过渡到由经常预算提供经费的职位这一目标。

64. 儿基会正在继续监测评价工作的水平。2015 年,负有评价职责的工作人员有 90% 以上还履行第二项职能(见表 6)。履行第三项工作职能的百分比略有提高,主要原因是大型办事处在办事处增设了第 1 和第 2 级监测和评价干事,负责支助方案监测和其他监测工作,包括公平结果监测系统的工作。

表 6

2013-2015 年涉及评价的专业工作职能组合情况^a

职称所述的 职能数目	在第 3 级或以上级别员额中 所见的组合情况	2013 年百分比 (数目)	2015 年百分比 (数目)	2013-2015 年的 变化
1	评价	9(15)	9(16)	0%
2	监测和评价	68(106)	65%(120)	-5%
	规划和评价			
	研究和评价			
3	规划、监测和评价	23(37)	26(48)	+13%
	社会政策、监测和评价			

资料来源: 人力资源司 2014 和 2016 年数据。

^a 为了与先前的报告保持一致,这些数据仅涉及第 3 级或以上级别的员额。

¹³ “工业国家”和“方案国家”是儿基会内部使用的国家分类正式术语。

65. 增强支助多证据职能的做法加强了作为整体的儿基会方案,但也会降低对评价的注意力。先前的定期调查已经确认,监测和评价干事用于评价的时间约占其工作时间的 15%。2015 年对两个区域此类干事的调查衡量了最近的时间分配状况。该调查确认监测和评价工作人员的时间和注意力是需求极大的紧缺资源,并确认这些区域的监测和评价工作人员用于评价的时间占其工作时间的 10%左右,这一比例低于先前的估计数。

表 7

监测和评价工作人员时间分配估计数

监测和评价干事的任务 请考虑到你在过去 6 个月的工作情况 分配你用于每项任务的时间的百分比	时间分配百分比		
	中东欧/独联体 ^a	拉加 ^b	平均数
情况监测(研究、调查、情况分析)	23	27	25
支助规划职能/规划股	16	16	16
分析和传播数据	14	13	13.5
儿基会方案监测(第 1 级/第 2 级)	11	14	12.5
支助加强国家统计系统的能力	12	c	—
评价	10	10	10
社会政策	10	11	10.5
其他	4	9	6.5

^a 中欧和东欧以及独立国家联合体。

^b 拉丁美洲和加勒比。

^c 拉丁美洲和加勒比的调查没有向答卷者提供这一选项。这一任务领域的工作可能反映在“其他”的数字中。

66. 工作人员培训和支助很重要。2013 年启动了一项内部电子学习方案,目的是提高儿基会国家办事处有效管理评价工作的实际能力。顺利完成课程的工作人员获颁了证书。共有 296 人完成了题为“在人道主义环境中进行评价”的电子学习课程,7 人完成了最近开设的题为“儿基会分权评价职能简介”的课程。在完成课程后,监测和评价专家获邀提交并实施一项旨在取得高级证书的知识应用计划。

67. 本报告已经指出,与儿基会人道主义方案拟订规模日益扩大的情况相比,对儿基会开展的人道主义工作、尤其是它对 1 级和 2 级紧急状况所作回应的评价不足。为此,儿基会启动了一项举措,以提高它制定、管理和使用对人道主义行动评价的能力。该举措包括为评价管理人员及其政府和民间社会伙伴举办一系列培训讲习班并提供技术指导,这些都是在评价周期的关键里程碑时刻举办和提供的。

由评价办公室领导的培训，与作为整个培训期间个案研究的实际评价的实施工作相挂钩。

68. 2015 年在 3 个区域举办了讲习班并启动了 12 项评价。这一举措将在 2016 年推广到另两个区域。现在的意向是，这项工作不仅要建设对人道主义行动进行评价的能力，还要通过完成正在进行的评价，在儿基会及其伙伴中增进和加强关于儿基会支持的人道主义行动的知识和学习。

69. 评价办公室正在继续调整其结构，以履行《战略计划》规定的职责。2014 年和 2015 年期间，领导职能有所增强，专题评价工作人员队伍有所加强。2015 年年底，评价办公室在以下领域设有第 4 或第 5 级评价专家：在紧急情况中设有两个此种工作人员职位，在保健/艾滋病毒/艾滋病/营养、教育、早期儿童教育、饮水、环境卫生和个人卫生(讲卫生运动)和机构效率等领域各设一个此种职位。第 3 级人员队伍目前包括 3 名临时助理专业人员和 1 名“新人才计划”专业人员。负责监督评价能力发展和知识管理的科由一个第 5 级职位和一个第 3 级职位组成，每个职位都由核心资金提供经费。除了评价专业队伍外，评价办公室还有 3 个一般事务职位和两个管理职位(1 名 P-5 级高级顾问和一名 D-2 级主任)。

七. 加强儿基会内的评价：有影响力的评价

A. 使用评价促进决策、学习和政策宣传

70. 儿基会正在全球和地方各级开展评价，以提供可靠的证据，促进在儿基会内以及在伙伴和利益攸关方中的决策工作，并开展有充分依据的宣传和政策改革。评价结果必须得到儿基会以及伙伴和受益者的接受和理解并附有明智的建议，同时促成作出以全面和及时的方式予以落实的管理层回应，否则，这种结果的价值有限。执行局已要求更多地注意这些问题。¹⁴

71. 使用评价证据是重要的考虑因素，不仅在报告评价结果的阶段如此，在从规划和制定评价活动到执行和报告的整个评价周期中也是如此。“评价影响”是一种思考方式，可据以思考评价可以在改善绩效和方案拟订以及切实倡导更好的政策方面产生的作用。在确定评价影响程度时应该考虑到的因素包括利益攸关方是否接受调查结果，这些结果有没有促成利益攸关方采取行动，以及它们有没有产生即时或延迟的直接和间接影响。另一些方面包括如何使用评价支持决策或改善规划和方案拟订及执行工作，包括改善方案绩效和更多地关注诸如公平、性别平等和社会包容等优先事项。本报告下一节根据儿基会区域办事处和国家办事处提供的最新实例讨论了评价影响。

¹⁴ 例如第 2013/13 号决定。

B. 对一些在国家和区域两级有影响力的评价的分析

72. 本报告本节重点介绍了一系列评价以哪些方式促进并在一定程度上形成了决策、学习和政策宣传，从而可被视为有影响力的评价。¹⁵

73. 若干评价对儿基会和(或)其伙伴组织采用的方案拟订办法产生了影响。例如，由于对关于儿童早期发展的试点工作进行了评价，儿基会驻约旦办事处打算在尚未提供儿童早期发展服务的城市地区逐步推广实施一种“儿童友好空间+儿童早期发展”模块。利益攸关方已经利用在中东和北非进行的对一项关于作为变化推动力的巴勒斯坦青少年的多国方案的评价，促进在青少年和青年参与领域的决策和改变工作方法。这项评价还加强了儿基会为约旦、黎巴嫩、阿拉伯叙利亚共和国和巴勒斯坦国境内青少年进行的方案拟订工作。对阿富汗“让我们学习”方案的评价表明，社区级的创新教育办法为失学儿童和少女提供了完成学业的机会。在孟加拉国，对儿基会在该国战略定位的评价正在促进即将进行的国家方案拟订工作，该方案将继续采用一种强有力的提供服务的办法，同时更加注重气候变化和减少灾害风险，并加强实施上游政策和宣传。

74. 在孟加拉国，评价强调了利用生命周期框架实现加强方案间跨部门协作。同样，在印度，对儿童权利保护和教育的评价表明，“趋同方针”(以强化方式使若干部门合力支持若干地区或群体)将有益于童工问题和儿童权利战略。各项建议被纳入了国家办事处目前正在制定的童工问题战略，并促进了儿基会在查谟和克什米尔的方案拟订工作，尤其是关于与政府的教育、司法和警察部门合作的方案拟定工作。在巴基斯坦，对儿基会的大规模环卫方案的评价促成将一项明确的变革理论纳入了逐步推广方案的工作，同时为城市和农村地区确定了独特、明确的变革之路。

75. 若干评价对中央和地方各级的政策产生了影响。在马来西亚，对马来西亚国家方案公平重点的评价促成在国家、区域和全球各级加强了与主要利益攸关方的伙伴关系和联系，国家方案对此有所反映。最后，关于中东和北非的多国评价的

¹⁵ 在阿富汗、孟加拉国、印度、约旦、黎巴嫩、马来西亚、巴基斯坦、刚果共和国、阿拉伯叙利亚共和国、越南和巴勒斯坦国进行了以下的评价：“让我们学习”形成性评价，2014年(阿富汗)；对儿基会在孟加拉国的战略定位的评价，2015年(孟加拉国)；对在古吉拉特邦、拉贾斯坦邦和马哈拉施特拉邦促进保护儿童权利和教育工作的评价，2015年(印度)；对儿基会对约旦境内叙利亚儿童紧急社会心理支助应对工作的评价，2015年(约旦)；对马来西亚国家方案公平重点的评价，2015年(马来西亚)；对儿基会在巴基斯坦的大规模环卫方案第一阶段(2013-2014年)的评价，2014年(巴基斯坦)；对改善刚果共和国境内土著人民生活质量方案的评价，2015年(刚果共和国)；对儿基会支助的越南教育和培训部 2006-2014 年以母语为本的双语教育的评价，2015年(越南)；对“巴勒斯坦青少年：推动积极变化的动力——营造促进和平与和解的环境”方案第三和第四阶段的评价，2015年(多国方案：约旦、黎巴嫩、阿拉伯叙利亚共和国和巴勒斯坦国)。可查阅 www.unicef.org/evaluation/index_90792.html。

结果和建议已经对联合国近东巴勒斯坦难民救济和工程处涉及约旦、黎巴嫩和阿拉伯叙利亚共和国境内青少年和青年的工作产生了影响。

76. 某些评价有助于指导方案推广工作。在约旦，儿基会已在继续着手推广非正规帐篷住区儿童友好空间方案，目的是惠及更多的儿童，特别是弱势儿童。在越南，在对以母语为本的双语教育方案进行评价后，省级主管机构一直在利用其本身的资源更广泛地实施双语教育方针并使之制度化。

77. 另一些评价强调了公平问题的重要性。在约旦，在非正规帐篷住区采用的“儿童友好空间”办法将惠及通常无法进入儿基会服务空间的弱势儿童。在阿富汗，儿基会注重于努力向最边缘化的儿童尤其是女童提供受教育的机会，以此实现教育公平。评价推出了实施一项更大型方案的构想，这种方案不但可有助于应对教育不公平现象，还可有助于应对社会和经济不公平现象。在刚果共和国，评价方案的目的是提供关于如何改善得不到充分服务的土著居民生活质量的意见。在马来西亚，评价促成制定了若干项注重公平的战略。这些战略的目的是根据利益攸关方为确认处理公平问题的重要倡导方而开展的调查，应对青少年和女童问题。评价的研究阶段同与重要的伙伴讨论国家方案的时机相符，这有助于提高公平议程的知名度、特别是在政府部委的知名度，从而支持有利于弱势群体、包括残疾儿童、移民儿童和无国籍儿童的举措。

78. 这些实例简要地表明了评价可以在各种情况下发挥的积极影响和促进作用。现在的挑战依然是如何更全面一致地衡量和监测这种影响。如果儿基会要最大程度地利用并得益于它的评价工作，它就必须应对这一挑战。目前正在进行的对评价管理层回应系统的审查，在某种程度上阐明了这些问题，但还必须进一步开展工作，记载并分析评价的作用和影响，特别是按照《2030年议程》在不断变化的发展格局下开展此种工作。评价拥有加强、注重和加快实现全球目标进展的潜力，但必须更清楚地了解哪些评价进程最为有用，并了解评价证据可如何以最佳方式促进全球进步以造福儿童。

八. 结论：挑战和前景

79. 本报告概述了正在儿基会评价职能各重要方面取得持续进展的情况，这种进展是通过全球评价报告监督系统的指标和其他信息衡量的。本报告还注意到已经加大力度，特别是通过联合国评价小组和富有活力的“评价伙伴”倡议，促进评价办法和做法的一致性。

80. 然而，本报告也强调了迅速变化的发展格局，并强调了如果要按照《2030年议程》取得迅速的进展，就必须应对的各种挑战。如果要实现秘书长所称的“无论是在哪个领域和在哪个级别进行的评价都将在执行新的发展议程中发挥重要作用”的情况，就必须实施重大的变革。这方面关键的重点转移，是转向国家主

导的评价和加强国家评价系统和能力。同时还必须找到适当的方式，在产生和使用数据、信息和知识的宽泛框架内最佳地使用评价证据。必须在全球范围内获得充分的专门知识和资源，据以切实利用知识，发挥全球各地下一代儿童的潜力，以求建成一个繁荣、和平的世界。

九. 决定草案

81. 执行局

注意到 2015 年关于儿基会评价职能的年度报告([E/ICEF/2016/11](#))。
