



经济及社会理事会

Distr.: General
3 April 2012
Chinese
Original: English

供采取行动

联合国儿童基金会

执行局

2012 年度会议

2012 年 6 月 5 日至 8 日

临时议程 * 项目 7

儿童基金会评价职能和主要评价工作的年度报告

摘要

本年度报告是按照评价政策（E/ICEF/2008/4）和执行局有关决定编写的。报告介绍了评价工作的全球发展情况以及儿童基金会评价职能的现状，并扼要说明了在加强权力下放的评价职能方面的进展。报告对暴力侵害儿童方面的方案编制情况进行了一次专题分析，概述了 2005 至 2010 年期间开展的 52 个国家级评价的元评价结论。

* E/ICEF/2012/9。



导言

1. 儿基金会的评价职能是按照执行局在 2008 年核准的评价政策来行使的。它通过促进本组织的透明度、学习和问责，加强了本组织的工作。评价有助于儿基金会为儿童和妇女提供更好的成果，因为在发布评价报告以提高透明度的同时，它还为本组织提供及时、有战略重点的客观信息，使其了解各项政策、方案和倡议的执行情况。儿基金会的评价办公室负责领导和支持整个组织的评价职能。此外，它还委托他人进行独立评价，并与联合国系统内的机构以及其他伙伴机构开展联合评价活动。

2. 已在评价政策提供的框架内取得了稳步和持续的进展。不过，近几年来，组织环境和更广泛的国际背景都出现了重大变化，预计在 2012 年还出现进一步变化。因此，儿基金会拟对其评价职能和评价政策进行一次全面审查，目的是更新评价政策，以便在 2013 年中提交执行局供其审议。

一. 评价中的全系统一致性问题

协调和简化

3. 近年来，儿基金会一直在加强关注发展合作中取得的成果，同时也在加强关注各种评价方法、做法和活动。联合国评价小组帮助整个联合国系统在共同评价规范、标准和能力方面确保一致性。儿基金会在这一领域提供了实际领导和支助，并且继续在 2011 年有效参与总部和国家一级的工作：在总部一级，通过各种区域倡议，而在国家一级，则是通过联合国国家工作队监测和评价工作组。

4. 儿基金会为联合国评价小组工作队提供领导或为联合国评价小组主要工作领域提供适度供资和技术投入。具体情况包括如下：

- (a) 人权和性别平等：完成并传播评价指导，包括一本评价员手册；
- (b) 影响评价：编写指导材料；
- (c) 联合国发展援助框架（联发援框架）评价：编写和传播指导；
- (d) 国家评价能力发展：采用一种联合国系统协作通用办法、联合能力建设活动和查明良好做法；
- (e) 联合国评价小组评价实践交流：分享儿基金会关于重视公正的评价以及关于权力下放的评价职能的评价指导。

5. 在儿基会的支助下，联合国评价小组为评价一体行动做法做出了贡献，包括从2008年的可评价性评估(通过为2009-2010年国家主导的评价提供质量支助)到通过担任联合国评价小组主席的方式参与现行一体行动独立评价的管理工作。

6. 事实证明，联合国评价小组是在坚持遵守国际标准的同时，在整个联合国系统内协调各种评价做法的一种有效手段。它支持了涉及各种系统性问题的复杂评价，现在正在制定联合国规范工作的评价指南。但是，联合国机构的评价能力参差不齐，将继续需要儿基会通过联合国评价小组，以提供工作队领导、技术咨询和选择性供资等方式提供支助。

联合评价

7. 联合评价包括联合国内部的机构间活动以及其他多合作伙伴倡议。儿基会在2011年开展的主要联合评价活动包括：评价联合国女童教育倡议；以及为联合国促进性别平等和增强妇女权能署(妇女署)启动的2012年联合性别方案联合评价的设计、管理和供资提出建议。

8. 在所开展的联合评价活动中，有很多涉及人道主义行动：(a)对中央应急基金的评价，该基金是儿基会应急行动的最大单一供资来源(由联合国人道主义事务协调厅领导)；(b)全球后勤组群联合评价(由世界粮食计划署领导)，应在2012年报告；和(c)关于应对海地和非洲之角紧急事态的机构间实时评价。

9. 随着发展合作逐步成为由国家领导和组织的多方利益攸关方伙伴关系，联合评价变得越来越重要。但是，考虑到各机构的任务授权不同且能力参差不齐，对此种评价工作进行协调可能极具挑战性。需要给予进一步关注，寻找快速而有效的工作方法，以便进行高质量的联合评价。

二. 儿基会的评价职能：业绩和结果

10. 2011年，一组三个主要业绩指标被纳入年度报告，以便介绍整个儿基会评价业务的概貌。今年，为了提供更加详细的资料，本报告介绍了六个指标。

指标 1：接受管理并提交全球评价数据库的评价数量

11. 2010年，儿基会管理了140个评价及700个研究和调查(平均每个办公室1个评价和5个研究/调查)。这与整个儿基会在2008-2009年的平均年度总数(分别为248和813个)形成对比。数量下降的原因不清楚，但可以反映出评价分类更加明确，而且评价办公室和区域办事处强调选择评价的战略主题并将有限的资源用在数量较少但质量更高的评价上。

12. 向全球评价数据库提交评价报告的年度比例继续增长，从2009年不到30%增长到2010年的71%和2011年的91%。该数据库是组织提交报告、对报告进行

归档整理和质量评估的一项重要工具。为了为管理人员提供一种便利的管理和监督工具，儿基会开发了一个评价跟踪仪表盘，以便纳入 VISION（虚拟集成信息系统）的业绩管理模块。

指标 2：主题分配

13. 在 2010 年开展并接受全球评价报告监督系统¹ 审查的评价分布于以下中期战略计划重点领域：25%调查儿童幸存和发展问题；15%调查教育和性别问题；10%调查艾滋病毒/艾滋病问题；16%调查儿童保护问题；1%调查政策宣传和伙伴关系问题；33%调查各种交叉专题和组织绩效问题。这表明涵盖范围已遍及所有战略重点领域。

指标 3：所开展评价的类型

14. 评价对各级成果的有效性进行评估。在 2010 年开展的评价中，被评估的最高级别的方案成果如下：影响（33%）、结果（26%）和产出（33%）。为了支持较高级别的成果，儿基会期待今后的报告能够显示重视产出的评价有所减少。

15. 与其他联合国机构共同管理的评价比例从 2009 年的 1%增加到 2010 年的 6%。同样在 2010 年，与非联合国组织共同管理的评价比例为 26%（而 2009 年为 17%）；另外，国家主导的评价占 6%；儿基会主导的评价占 43%；有 19%的评价未明确报告管理安排。

指标 4：儿基会评价的质量

16. 2011 年，全球评价报告监督系统的外聘审查人员对各级儿基会办事处上一年度开展的 89 个评价进行了质量审查。总的来说，在 2010 年的报告中，按照儿基会评价标准，有 40%被评为满意（2009 年为 36%），还有 30%被评为基本满意但有些工作需要改进。但是，还有 30%的质量被评为不满意；这是 2009 年所完成评价中不满意所占比例（15%）的两倍。虽然认为这只是一年的反常，但评价办公室将会对这个问题进行监测，并作进一步的分析。

指标 5：评价的使用情况，包括管理方反应

17. 评价政策要求，每个评价都应该填写管理方反应。通过宣传和监督，已在遵守有关编写管理方反应以及向跟踪系统上传已填写反应的政策方面取得进步。上传率从 2009 年的 12%提高到 2010 年的 47%和 2011 年的 65%（针对上一年度完成的评价报告）。为了进一步改善遵守情况，已将适当的业绩指标纳入 VISION 的评价跟踪仪表盘；正在对相关指导进行修订，以便考虑到近期在利用该系统方面取得的经验；并且还在准备一个电子学习模块。

¹ 生成指标 2-4 所用数据的外部审查的 2011 年全球评价报告监督系统报告可参见：
http://www.unicef.org/evaldatabase/index_60935.html。

指标 6：企业级评价

18. 企业级评价由评价办公室管理，已被纳入综合监测和评价框架。2010-2011 年综合监测和评价框架为开展 15 个企业级评价提供了承付款。截止 2012 年 4 月，有 10 个评价已经完成（67%）；3 个正在执行当中（20%）；2 个（13%）被取消（原因是缺少数据和经费）。综合监测和评价框架的评价结果概要以及已经完成的评价可参见儿基会网站的评价网页（www.unicef.org/evaluation）。

19. 2011 年，完成了史上第一次对儿基会幼儿发展工作的全球评价，提出了将这一做法纳入主流的各种建议。去年，还完成了关于教育在建设和平中的作用的研究调查，确定了针对儿基会和更广泛的联合国系统的主要教训和建议，并且通报了一项新的重大教育和建设和平方案。在对编制儿基会方案时采用的注重终生技能、营养和人权的做法进行评价方面取得了进展。在体制专题方面，编写了一份综合报告，总结了前几年所开展的若干国家级方案绩效评估的经验；并且对儿基会亚太共享服务中心的教训进行了审查和记录整理。儿基会与荷兰政府共同管理了一项关于莫桑比克方案所产生影响的评价，此方案的目的是将能够获取安全饮用水和安全卫生条件的人口数量增加 100 万。到目前为止，在这个地区开展的高质量影响评价数量很少，而这一评价就是该地区一系列全球评价之一。

20. 还对儿基会在人道主义应急方面的业务情况进行了评价，特别是对本组织应对 2010 年 1 月海地地震时的业务情况进行了一次独立审查。这是一次内容详实的评估，为制定企业应急启动程序做出了贡献，并且还涉及重大紧急事件的简化标准操作程序。²

21. 在儿基会内部进行广泛咨询之后，已对 2012-2013 两年期的综合监测和评价框架进行了修订。已向执行局提供 2012-2013 年综合监测和评价框架，并且已将其张贴到儿基会网站的评价网页中。

三. 儿基会的评价职能：人力和财政资源

22. 每两年向执行局提供一次关于可用于评价的人力和财政资源的信息。这样有利于对两年期的预计轮换情况进行跟踪，并且报告长期趋势，而不是报告比较多变的年度起伏情况。

人力资源

23. 截至 2011 年 12 月，有 95 个国家办事处设有职称中含“评价”二字的各级别专业员额，与此相比，2010 年有 89 个，2008 年有 85 个。另外，全部 7 个区

² 已在第 9 段提到对人道主义活动的联合评价。

域办事处和两个总部也都设有此种员额。表 1 说明了 2008 年 3 月至 2011 年 12 月期间第 3 及以上职类专业评价人员的数量变化。³

表 1：2008-2011 年儿基会专业评价人员的人数增长情况

	2008 年 3 月	2010 年 3 月	2011 年 12 月	2010-2011 年的 增幅
国际专业评价人员*	50	66	73	+11%
国内专业评价人员**	40	37	41	+11%
专业评价人员（合计）	90	103	114	+11%
外地办事处专业评价人员所占比例	84%	84%	89%	+6%
女性专业评价人员所占比例	42%	50%	44%	-12%

资料来源：人力资源司各年度数据。

* P-3 及以上职类。 ** 本国专业干事 C 级及以上职类。

24. 2010 年以来，大部分变化都是积极的。较高级别的员额增长了 11%，这一增长正在帮助儿基会实现其做出的有关增加评价专项技能的共同承诺。已在国际专业职类增设岗位的事实正在为职业发展提供更多机会。但是，增加的全部评价干部均由男性候选人填补；这意味着 2010 年实现的各占 50%的比例出现了下滑。

25. 设有第 3 或以上职等员额的办事处数量在 2010 年从 53 个增加到 61 个。此种员额的数量是衡量本组织致力于从评价中获得高质量证据的一种指标。不过，儿基会的大多数国家办事处都没有第 3 或以上职等的评价工作人员。因此，它们缺少高质量评价所需的技术支持。很多办事处都由担任监测和评价“联络人”的行业专家负责监测和评价。

26. 儿基会首次能够报告主要工作人员投入评价工作的时间和做出的努力。如表 2 所示，在负有评价职责的工作人员当中，有 90%以上履行的是第二职能，有 28%履行的是第三工作职能。

27. 从逻辑上讲，监测与评价的结合加入了相关职能。增加规划以及社会政策等更间接职责的目的是试图混合不同的技能组合，这么做当然会减少可用于监测和评价的时间。虽然此种结合是可以理解的，但考虑到办事处的人员配备限制，它们并不是理想的做法。

³ 第 3 职等是确保系统能力的预期最低级别。

表 2：2011 年涉及评价的专业工作职能组合

职称中的 职能 数量	在第 3 或以上职等中注意到的组合	总数	比例
1	评价	10	9%
2	监测和评价 规划和评价 研究和评价	72	63%
3	规划、监测和评价 社会政策、监测和评价	32	28%

资料来源：人力资源司数据，2012 年。

28. 重要的相关证据由 2011 年对负有监测和评价职责的外地办事处级工作人员业绩情况进行的一次调查（以下称“权力下放的监测和评价职能调查”）提供。每个办事处负责监测和评价职能的领导都回答了很多与其工作有关的问题。如下文表 3 所示，当被要求计算其时间分配情况时，评价所占比例最低。从区域、办事处规模或职称来看，没有什么重大差别。

表 3：2011 年监测和评价工作人员用于监测和评价职能的时间比例

职类名称	任务范围（选择）	被分配的时间比例	工作人员对人力资源可用情况的满意度
形势监测	国家方案文件；联发援框架；方案指标	29%	75%
规划和绩效监测	部门监测；统计系统；住户调查	24%	77%
研究	调查；研究；形势分析	17%	75%
其他	为高级工作人员提供支助；捐助方报告；财务监测	16%	不详
评价	设计职责范围；监督咨询人员；评价报告；管理方反应	14%	70%

资料来源：《2011 年儿基会权力下放的监测和评价职能调查报告》，第 19-22 页。

29. 虽然绝大多数工作人员对人力资源的可用情况表示满意，但还是有相当数量的少数人表示不满意，特别是对评价任务，指出有一种过度紧张的感觉。

财务资源

30. 自 2006 年以来，关于评价职能的年度报告就已包括了关于评价支出的数据，这些数据依据的是管理人员分配给各项交易的编码。编码指南一直到 2010 年都没有变化，但在 2011 年进行了重大修改。修改的结果是允许各办事处在“评价”

项下对更加广泛的活动进行编码，结果导致 2011 年数据与前几年的数据没有可比性。就本报告而言，2010 年的数字将被用于说明前几年的可比数据。

表4：2006-2010年各国家和地区办事处的评价支出趋势

	(单位：千美元)			
项目	2006 年	2009 年	2010 年	2009-2010 年的趋势
评价支出	7 970	9 560	11 122	+16%
比较器：儿基会方案总支出	2 120 000	2 940 000	3 354 000	+14%
评价支出在所有方案支出中所占比例	0.38%	0.33%	0.33%	未变
相关支出：分析、研究和调查	27 900	46 310	47 715	+3%
相关支出：数据、数据库、调查和统计	24 700	35 230	40 060	+14%
研究/社会数据/评价方面的总支出在所有方案支出中所占比例	2.84%	3.10%	2.95%	-0.5%

资料来源：根据财务和行政管理司提供的数据计算，2011年。

31. 表 4 中所述趋势与前几年一致。0.33%的评价支出占比远低于方案支出的 1%。⁴ 虽然随着总体方案预算的增长，评价支出也在不断增长，但增长的趋势没有什么变化。

32. 在供资类型或供资来源方面，传统上，外地办事处一级的评价供资模式一直是整个儿基会供资模式的真实反映（见表 5）。

表 5：2006-2010 年各国家和地区办事处用于评价的供资来源

供资来源	(单位：千美元)					
	按来源分列的支出，年度平均值			按来源分列的百分比		
	2006-2007 年	2008-2009 年	2010 年	2006-2007 年	2008-2009 年	2010 年
经常来源	3 500	3 350	4 810	39%	36%	43%
其他来源 – 经	3 950	4 800	5 360	44%	51%	48%

⁴ 受常务副秘书长办公室委托于近期在联合国内部进行的一项评价职能审查指出：“虽然没有正式的预算比例标准，但一项广泛使用的指南是评价支出占总预算的 1%……在联合国系统中，大部分评价单位的支出都不到 1%。”《全系统独立评价机制：大会第 64/289 号决议委托的全面审查现有的联合国系统发展业务活动全系统评价体制框架——报告草案》，由 Angela Bester 和 Charles Lusthaus 编写（2012 年 1 月），第 38 页。

常资金						
其他来源 – 应急资金	1 600	1 200	950	18%	13%	9%
评价资源总额	9 050	9 350	11 120	100%	100%	100%

资料来源：财务和行政管理司各年度数据

33. 正如已经指出的那样，2011 年的管理数据与前几年的数据没有可比性。将更广泛的活动列入涉及评价的类别，导致支出数额达到 3 010 万美元，约为近几年平均数额的 3 倍。⁵ 由执行VISION导致的变化可能会引起进一步的调整。达到一个新的基准和稳定的数据要花一些时间。儿基会预期会在 2014 年报告中提供更加明确的信息。

34. 仅就评价办公室而言，与 2008-2009 年相比，2010-2011 年期间的支出略有增加（见表 6）。在本两年期内，主要的外部供资伙伴是荷兰、西班牙和瑞士。就 2012-2013 年而言，评价办公室的经常资源不会有任何增长。其他现有资源有所减少，从 2010-2011 两年期开始时的 320 万美元下降到目前的 176 万美元，表明资源动员方面已出现困难。在内部，评价委员会正在对评价方面的资源分配进行监测。

表 6：评价办公室：可用供资

	(单位：千美元)			
	2010 年 ^a	2011 年 ^a	2012 年	2013 年
支助预算	2 183	2 175	2 023 ^b	2 150 ^b
经常资源	787	1 383	1 100 ^b	1 100 ^b
其他资源 – 经常资金	818	1 075	827 ^c	
其他资源 – 应急资金	1 209	603	937 ^c	

资料来源：财务和行政管理司各年度数据

a/ 支出款项。 b/ 分配款项。 c/ 手上现有款项。 注意：2011年的数字为暂时数字。

⁵ 按供资来源，分列如下：经常资源占 34.7%；其他资源—经常资金占 52.9%；和其他资源—应急资金占 12.3%。

四. 加强外地的评价工作

概况

35. 权力下放的监测和评价职能调查发现，权力下放的评价职能行使得相对比较顺利，并且与儿基会评价政策保持一致。该调查还发现，权力下放的结构提供了一个框架，儿基会国家、区域和总部各级办事处可用其来支持组织学习、提高绩效、战略决策、成果报告和注重证据的政策宣传。调查还认为，儿基会为支助国家监测和评价系统做出了重要努力，但是，调查也注意到一些重要领域需要关注，以便使权力下放的评价职能实现一致的质量。下文将对具体的优缺点进行讨论。

治理和领导

36. 调查发现，权力下放的评价职能的治理有优点，但也有一些领域需要进一步改进。调查发现，遵守以下基准的情况最好：

- (a) 监测和评价工作人员在制定职责范围方面提供援助（98%的国家办事处）；
- (b) 监测和评价工作人员在选择评价团队方面提供援助（91%）；
- (c) 高级管理层核准职责范围（88%）；和
- (d) 高级管理层核准对评价团队的选择（89%）。

有两个领域需要改进：

- (a) 区域办事处参与评价报告草案的质量审查（86%）；和
- (b) 在国家办事处内部，高级管理层对管理方反应的准备情况进行监督（76%）。

37. 为了便利对权力下放的评价职能进行监督，评价办公室开发了一个仪表板，用以对照关键绩效指标来说明每个区域的进展情况，这些关键绩效指标来自各种内部数据库。这个仪表板将在 2012 年被纳入 VISION 的绩效管理模块。中东和北非区域办事处已经启用一种类似的实时连续桌面信息检索引擎，以便于进行国家一级的监督。

38. 区域评价委员会由国家代表和区域办事处高级工作人员组成，就区域评价优先事项以及如何进一步遵守评价政策提出咨询意见。这有助于各国家办事处制定战略，以便根据评价政策规范调整评价职能治理（中非共和国和尼日尔），使评价成为其年度管理计划中的一个正式优先事项（乌兹别克斯坦）以及开发工作流程跟踪系统（厄立特里亚）。

加强评价的战略规划

39. 儿基会继续将综合监测和评价计划作为一种决策工具，用以支助对国家一级评价活动的战略规划。约有 86% 的儿基会办事处指出，由高级管理层领导的会议对综合监测和评价计划进行了讨论；有 55% 的办事处报告说，各国政府和包括其他联合国机构在内的合作伙伴参与了综合监测和评价计划的规划阶段。埃塞俄比亚和坦桑尼亚根据其政府规划日程表对监测和评价计划进行了调整。受到非洲之角紧急事态严重影响的埃塞俄比亚、肯尼亚和索马里在东部和南部非洲区域办事处（东南非区域办事处）的支助下，制定了紧急事态综合监测和评价计划。

40. 在区域管理队的领导下，美洲和加勒比区域办事处（美加办事处）对国家综合监测和评价计划进行了一次区域分析，以便查明若干国家正在规划评价的专题领域，目的是加强协调与配合，以便收集关于区域优先事项为儿童做些什么工作的证据。东亚和太平洋区域办事处、中东和北非区域办事处以及西非和中部非洲区域办事处都制定了区域专题评价的计划。中欧和东欧以及独立国家联合体区域将财务规划的比例固定，将 5% 的专题供资专门用于评价活动。

促进和支助质量评价

41. 外聘审查人员对 2011 年 4 月之前提交的所有评价报告的质量进行了评估（这些报告涉及 2010 年开展的各种评价）。⁶ 进行了一次全球分析，并与各区域及国家办事处讨论了各项结论和后续行动。个别反馈报告已发送到儿基会各办事处，并且提出了有关如何提高今后评价质量的实用建议。突出了 35 个质量最好的评价，并且与所有相关方案部门进行了分享。

42. 各区域办事处为提高国家办事处管理的评价质量做出了重大努力，特别是通过各种质量保证系统对监测和评价计划、职责范围和评价报告进行审查。但是，并非所有国家办事处都充分利用了这一区域支助。为了应对这一挑战，中东和北非区域办事处以及中欧和东欧以及独立国家联合体区域办事处规定 2011 年开展的所有评价都必须接受区域质量保证系统的审查。区域监测和评价干事正在越来越多地被邀请加入国家一级的战略评价指导委员会。西非和中部非洲区域办事处为确保各项国家方案的可评价性做出了重大努力：贝宁、喀麦隆、马里、几内亚比绍、尼日尔和刚果共和国都开展了可评价性评估。在机构间一级，美加办事处重新设立了一个联合国评价能力发展问题工作队。

加强评价使用和管理方反应

43. 评价办公室和各区域办事处加强了对编写和提交管理方反应的宣传和技术支助。亚太共享服务中心对良好做法进行记录整理，并且进行了传播，中东和北非区域办事处就如何克服瓶颈问题编写了指南。

⁶ 关于全球评价报告监督系统的描述，见 2011 年执行局评价报告第 22-23 段。

44. 权力下放的监测和评价职能调查广泛报告了在利用各种评价、研究和调查方面的良好做法，特别是将其结论用于 (a) 报告结果 (92%的国家办事处)；(b) 组织学习 (84%)；和(c) 循证政策宣传 (82%)。调查发现，需要在传播评价结论方面做出更大努力，包括通过编写传播材料 (73%)。

加强内部评价能力

45. 各级办事处继续为提高内部评价能力做出重大努力。其中包括知识管理活动，如网上实践群和网络研讨会 (“webinars”)、服务台职能、电子公报和时事通讯以及常规网络会议、考察访问和培训研讨会等。各区域办事处对中东和北非区域办事处以及亚太共享服务中心评价时事通讯等加强全球倡议的各种具体努力进行区域管理。权力下放的监测和评价职能调查表明，监测和评价工作人员普遍相信其在管理评价方面的技能。但是，他们对建设和平或人道主义局势的评价管理不太自信。

46. 调查指出，虽然儿基会监测和评价工作人员充分利用了知识管理机制，但有关监测和评价的系统培训尚未充分形成制度。为了解决这一瓶颈问题，评价办公室正在与人力资源司和各区域办事处合作，共同开发一个在线自学归纳方案，涉及对国家一级评价职能的治理，并且为管理和使用各种评价提供核心技能，方案重点关注人道主义方面。

47. 评价办公室正在带头加强儿基会各办事处及合作伙伴在评价注重公正的干预措施方面的能力。编写并散发了一份关于关注公正的评价的设计和管理手册，并且带有培训；与妇女署合作设立了一个电子资源中心；并且与几个组织合作发起了一个系列网络研讨会。

48. 值得注意的是，权力下放的监测和评价职能调查表明，很多监测和评价干事的任期都相对较短：20%在儿基会内部的监测和评价经验不到 2 年；49%拥有 2 至 5 年的经验；31%拥有 5 年以上的经验。小办事处和中等收入国家办事处的工作人员的经验最少。这些数字表明，在聘用具有专业经验的监测和评价人员方面存在问题。

五. 国家评价能力建设

联合国评价小组和国家评价能力

49. 2011 年，联合国评价小组核可了关于联合国评价小组成员在国家评价能力建设方面的作用的概念文件，其起草者是一个由儿基会和联合国开发计划署共同领导的工作队，但目前该工作队由儿基会和国际劳工组织 (劳工组织) 共同主持。该文件的目标包括促进知识共享、加强国家主导的评价的质量以及为全球和区域评价协会提供支助。

全球学习

50. 儿基会与几个组织合作，巩固了其作为国家主导的监测和评价系统的知识中心的地位，扩大并更新了一个Web 2.0 互动平台：*MyM&E*。这是一个公开的机构间网站（www.mymande.org），便于全世界共享国家主导的监测和评价系统方面的知识。该网站的成员数量在 2011 年增加到 2 839 个；自 2010 年以来，访问者的人数翻了一番，访问者来自 45 个以上的国家，人数达到 42 700 人；网页下载量也在一年内增加了 50%，达到 111 000 次。

51. 与妇女署和联合国评价小组合作，启动了一个关于公正、人权和性别平等的电子评价资源中心，为世界各地的评价者提供了免费和便利地获取最新评价办法和方法的途径。

52. 作为 *MyM&E* 网站的重要组成部分，安排了一个关于关注公正的评价的系列现场网络研讨会。在 2011 年举办了 7 次网络研讨会，特色是请来一些具有国际名望的主旨发言者，吸引了来自世界各地的 1 000 多位参与者。这一系列网络研讨会的共同发起者还包括联合国评价小组、开发署、劳工组织、妇女署、克莱尔蒙特大学和洛克菲勒基金会以及一些全球专业评价协会。

加强国家主导的监测和评价系统

53. 除了支助公共部门评价能力之外，儿基会还越来越多地在评价方面与大学和研究中心等国家英才中心合作。美加办事处为绘制拉丁美洲和加勒比地区的英才中心提供便利；儿基会国家办事处目前正在与各中心接洽，以便为各种战略评价提供支助。在塞内加尔，世界银行将非洲高级管理研究中心确定为非洲法语国家评价和结果问题区域机构学习中心，儿基会正在为该中心提供支助，以便提高其作为一个培训机构的能力。在津巴布韦，儿基会帮助建立了业务研究和评价协作中心，以便为方案规划提供高质量的证据。

54. 在儿基会的支助下，摩洛哥评价协会成为国家宪法修订委员会的一员。根据新的宪法，政府承认有义务对公共政策和方案进行评价。目前，儿基会正在支助摩洛哥评价协会倡导和支持建立一种国家评价制度，以便对包括适当预算分配在内的公共政策进行评价。同样，在加纳，儿基会为国家发展规划委员会提供支助，以便对已经分配并用于支助监测和评价系统的资源进行一次基准评估。这次评估导致财政部向所有部委下达指示，要求它们向监测和评价系统分配充足的资源。在吉尔吉斯斯坦，儿基会为在卫生和教育部以及总理办公室设立并运作监测和评价单位提供了支助。

六. 概要：应对儿基会面临的评价挑战

55. 前几节介绍了儿基会评价制度的各个方面。本节将概述儿基会面临的主要评价挑战，并对目前为应对这些挑战及加强儿基会有效利用评价结果所采取的行动进行总结。

56. 儿基会的评价政策为本组织的评价职能提供了一个健全的框架。在执行这项政策方面已经取得稳步进展。已经为该职能的监测和支助管理设立了一项有效的监督制度。另外，在 2012 年，将把一些选定的关键绩效指标纳入 VISION 系统，以便于管理方进一步关注政策的遵守和执行问题。

57. 加强监督已经帮助儿基会兑现其主要政策承诺。例如，向中央评价数据库提交评价报告的遵守率现已超过 90%。但是，更好的管理信息也突出了需要进一步改进的领域，特别是在评价质量方面。在这里，主要决定因素是评价职责范围的质量；拥有充足技能和知识的监测和评价工作人员的管理和监督；咨询人员的质量；提供充足的预算；以及持续的管理方关注，包括从评价专题的战略选择一直到确保切实有效的管理方反应和后续行动。

58. 权力下放的监测和评价职能调查表明，国家办事处正在密切关注职责范围的制定和评价团队的选择情况，同时在管理方反应方面也需要做出更大的努力。虽然遵守这一要求的比例在过去三年里增加了两倍，但在 2010 年完成的评价中，只有三分之二准备了正式的管理方反应（按照 2011 年报告的数据）。目前，评价办公室正在更新关于管理方反应的指导意见，并将确保广泛传播，同时，各区域办事处也将对外地办事处的执行情况进行监督。

59. 虽然权力下放的监测和评价职能调查表明在利用评价材料方面有一些良好做法，但也还有做出更大努力的空间。评价办公室正在开始与区域监测和评价工作人员合作，以扩大传播材料和渠道的范围，目的是加强对评价证据的理解和采用。

60. 正在通过一种新的电子学习方案、一系列流行网络研讨会和常见的电子公报，以及通过 My M&E 网络平台、在线评价实践群和评价办公室服务台，为培养工作人员的技能 and 知识提供帮助。

61. 最后，将在 2012 年进行一次全面的评价职能审查，以便查找制约因素并查明进一步改进和完善该制度的机会，同时为更新评价政策以便满足儿基会不断变化的需求提供合理依据。

七. 专题分析：评价和预防暴力侵害儿童

引言

62. 2011 年的年度报告中载有一个专题分析章节，旨在向执行局通报儿基会在重要方案规划领域内的评价情况。专题分析可以说明(a) 在对评价结论进行综合的基础上得出的评价证据；和(b)儿基会如何发展该专题内的知识和数据，包括需要缩小的差距。

63. 2012 年报告的专题是预防暴力侵害儿童。“暴力”是一个宽泛的问题，与《儿童权利公约》（1989 年）第 19 条及相关条款所定义的特征有重叠。因此，这一分析不限于直接的身体暴力，还涉及在儿童遭受伤害时的儿童保护问题，包括各种形式的剥削。

64. 在儿童保护领域，一项重要成果是加强了“数据的监测、研究、评价和利用”。之所以将预防暴力侵害儿童选作本次报告的主题，是因为这是一个具有挑战性的方案领域，而且所有合作伙伴都认识到加强可用于提供证据的各种指标、工具和方法非常重要。

信息库

65. 2005-2010 年期间完成的评价以及利益攸关方磋商成为本次报告的信息库。最近，儿基会开展了一次题为“防止儿童遭受暴力：评价结论综述”⁷ 的元评价，对主要由各外地办事处委托的各种评价进行了一次审查（只有一个全球评价可用）。从专题方面来讲，有三个或更多评价被发现涉及以下领域：有害的传统做法；在家照看儿童；武装冲突之后的恢复和过渡；地雷行动；自然灾害；童工；贩运儿童及儿童移民；虐待及剥削儿童；街头儿童以及校园暴力。拟在 2012 年对紧急事态中的儿童保护情况进行一次全球评价，这次评价将为紧急事态的应对提供参考，而在此方面，本次报告所介绍的情况不够。

66. 评价质量审查的等级评定情况见下文表 7。

表 7：2005-2010 年期间 52 次儿童保护评价的质量评定情况

总体评定	%
不满意	27
满意	48
很好	23
优秀	2
合计	100

资料来源：外聘审查人员运用儿基会标准。

⁷ 可参见 www.unicef.org/evaluation。

67. 评级范围表明其结果差于被调查年份儿基会的平均评定得分。评价质量中的缺点与方案设计中的不足是联系在一起的：在有些个案中，试图评估成效因方案成果框架中的弱点（指标不好）以及监测和评价安排疲软（设计简单、基准未测定、由疲软的信息管理系统提供的数据不充分、职责混淆以及能力加强工作无效）而受到阻碍。需要更多地关注解决成本效率和成本成效问题。评价需要得到儿基会办事处的始终如一的支持，避免工作范围中途出现变化。

68. 也有很多优点。评价重点关注相关性和适当性、有效性、效率和可持续性为核心评价标准。有些评价也对公正问题进行了考虑。所使用方法多种多样，但通常都是恰当的；在有些个案中，所用方法包括量的调查，运用对照组对影响和因果关系进行评估。例如，在津巴布韦，对儿童保护系统进行的一次分析在不同方案和对照领域内调查了 1 300 名儿童和成人。虽然紧急事态后的状况具有挑战性，但更容易获得供资，在紧急事态后这一背景下，评价的质量高于平均值。

关于有效方案规划的结论和教训

69. 本次分析重点关注了以下四个专题领域：

- (a) 达到高级别的方案相关性和有效性；
- (b) 形成全面的儿童保护制度；
- (c) 适应具体的社会背景；
- (d) 儿童安全防护。

70. 这一结论是根据其在关于暴力侵害儿童的方案规划战略中的中心地位以及充分证据的可用性得出来的。

(a) 相关性和整体成效

71. 在所审查的 52 份报告中，有 43 份明确讨论了方案的相关性，其中有 23 份（占 53%）被评估为完全与其背景相关、19 份部分相关，1 份不相关。相关性较低的情况最常出现在方案目标和干预措施与国家优先事项、法律政策框架或国际义务不一致时。

72. 关于整体方案成效，有 13% 被评估为无效，67% 被评估为部分有效，被评为很有效或完全有效的为 19%。

73. 虽然认识到应该在相关性和有效性方面加以改进，但结果还是被认为符合预期：

- (a) 儿基会和合作伙伴必须在现有制度内工作，即便是在这些制度会限制评价有效性时。方案是否会在随时间流逝而变得更加有效是一个关键问题。例如，

社会福利方案面临结构性挑战，如提名机制不健全、服务质量不高、缺少技术熟练的社区发展工作者、安全威胁和预算紧张。

(b) 为防止暴力侵害儿童而进行的方案规划所拥有的经过确认的证据基础有限。因此，该部门一直在继续创新，创新为各种成功带来了可能。在这些个案中学到的经验教训以及如何调整方案规划具有极其重要的意义。

(c) 已经制定的战略不多，因此，现有的衡量相关性的国家优先事项和框架可能发展不足。将注重权利的成果作为目标对于倡导政策改革很有效，适合于衡量相关性，对照它可以了解方案规划是否进展顺利。

(b) 儿童保护制度

74. 儿童保护方案往往强调建立儿童保护制度。作为主要组成部分，这包括法律、政策和标准；服务提供机制；人力资源；合作伙伴之间的协调与协作；传播和宣传；以及用以向决策者反馈信息的监测和评价。

75. 在很多涉及暴力侵害儿童的方案中，审查发现了一种有效的制度做法。在所审查的评价中，审查人员发现有 9 个采取良好做法，加强了司法部门：

(a) 建立执法部门和司法机构，使儿童案件的管理和后续行动效率得以提高，至关重要是增加了预算拨款；

(b) 建立区域机构提高了贩运者和剥削者的抓捕和定罪数量；

(c) 设立接近社区的机构增强了这些社区宣传司法、质疑有害社会规范和挺身与有害做法的行为者进行斗争的能力；

(d) 制定案例管理和儿童友好的程序加强了对受害者、证人及犯罪者的保护，包括将很多儿童犯罪者从监狱转移出来；

(e) 指示关注预审、审判和审后各个阶段，确保关怀的连续性。

76. 毫无疑问，这种制度做法的价值在评价中被确定下来。但在很多情况下，评价人员都注意到了一个或多个与这些制度的设计或执行有关的问题。他们发现以下战略是有效的补救手段：

(a) 确保获得支持这种制度的政治承诺，包括查明弱势或边缘化儿童；确定最低限度的一揽子服务计划；进行游说，以便将其纳入法律框架和政策；进行成本分析；以及要求充足预算；

(b) 将教育和卫生等各部门的儿童保护关切结合起来。这涉及脆弱性的共同原因，并且增加了整个生命周期内的行为者数量。评价发现，综合性的跨部门方案更为成功。但是，如果不可能采取跨部门做法，仍然有理由采取范围较窄但经过充分设计的干预措施；

(c) 规划“连续关怀”，包括预防、早期干预、临时关怀和恢复。如有关打击贩运儿童方案一样，注意到相比立即援助/临时支助，有一种忽视预防和早期干预的趋势；

(d) 将较长期的社会改革战略与满足社区迫切需求的短期实际“进入点”结合起来。例如，除了提高认识之外，儿童伤害预防方案还应该将提高认识与改变环境结合起来（在池塘周围设立围栏；道路照明）；

(e) 鼓励私营部门与公共部门、民间社会以及社区行为者合作。个体商号的参与往往会涉及通过培训和实习修改劳动惯例或对社会保护方案的支持，但这很少作为较大社会动员战略的一部分；

(f) 在建立伙伴关系时给予应有注意和适当监督，特别是在涉及弱势组织或新组织时。如果这方面做得不好，评价人员会怀疑没有进行机构评估或职责不明确；

(g) 在儿基会深度参与期之后执行谨慎的退出战略。退出战略往往不是从一开始就制定，没有财务可持续性计划，沟通不畅，管理不善。

(c) 背景

77. 不管是建立全面的制度，还是执行目标范围较窄的方案，针对暴力侵害儿童行为的倡议都必须对社会背景有敏感的认识。虽然一些社区规范和惯例对结束暴力侵害儿童行为有帮助，但其他的一些则可能有害。

78. 关于暴力侵害儿童方案的某些方面显示了对背景的极大敏感性。对滋生或打击暴力行为的信仰和社会组织进行调查有助于确定其与贫穷等根本结构性问题的联系。当儿童参与被社区视为鼓励积极成熟时，它显示了强大的结果。管理有方的参与有助于儿童增长知识和技能，并且有助于减少校园暴力。

79. 不过，评价人员也注意到有多种不完全适应的情况威胁到方案的成功。它们包括：

(a) 源自其他环境并且强行在社区推广的“再循环”干预措施及概念，而不是扎根于本地惯例和现有方案。例如，有很多社区成员抵制注重权利的思想，因为这种思想与本地环境不相关且缺乏可参与性。

(b) 未能加强家庭和父母支持。一些方案在向儿童提供信息或技能培训时未考虑到这一人际背景。

(c) 低估儿童参与的机会。有一种趋势是让儿童以一种不现实的方式指导各种方案，或者是根本不考虑儿童的意见，想当然地认为他们不知道其最大利益。有很多参与活动是为独立团体设计的，而不是作为社会结构的一部分。包括疾病儿童或来自边远或移民社区的儿童在内的社会边缘儿童的参与程度最低。

(d) 误调性别角色。虽然性别公正战略成为很多方案的主要组成部分，但执行情况往往不好。常见问题包括未能在设计阶段开展深入的性别分析、缺乏性别分类数据、未能消除女孩参与的阻碍，包括家务和结婚的社会压力，以及不注意男性和男孩问题的定义，纳入此类定义本会对男孩和女童都有所帮助。

(d) 儿童的安全防护

80. 儿基会认识到，在支助合作伙伴解决暴力侵害儿童问题时，所涉及的组织和人员应该在儿童安全防护方面高度协调一致，因为除了其他关切之外，这些方案还涉及保密问题、容易遭受危险和敏感的家庭做法。很少有评价明确提到儿童安全防护做法，但它们确实得出了重要结论。注意到在制定儿童安全防护政策方面有一些良好做法，包括行为守则、儿童保护能力培训、指定儿童安全防护协调人以及制定提名规程和“吹哨”政策。有些方案将这些政策扩大到涵盖游客和记者（老挝人民民主共和国）或研究人员和评价人员（哥伦比亚）。

81. 评价发现，合作伙伴没有制定儿童安全防护政策或这些政策执行不力时，会出现很多道德问题，包括在确定如何对待儿童时未能尊重儿童的意见；未能保护儿童的尊严；产生无法实现的预期；以及最重要的是以伤害的方式冷落儿童。将儿童置于危险中的具体做法包括将儿童送回虐待环境而不进行后续监测；要求儿童指认有可能实施了犯罪行为同龄人；要求儿童向成人报告暴力行为而成人却未采取任何行动；将培训机构设在儿童必须在不安全的环境下走很长一段路程的地方；以及违反保密原则。偶尔也有成人执行者遭遇风险的情况。

82. 研究这些问题是如何产生的并非本次审查的范围。这项工作若干项评价的一部分，这就引出了一个问题，即在儿基会提供援助的情况下，儿童安全防护政策和制度是否充分。审查这些问题是下文所介绍的一项建议的主题。

监测和方法发展倡议

83. 有关暴力侵害儿童方案规划要想取得成功，必须将监测数据和研究结论与评价证据相结合。所有监测和评价活动都取决于是否拥有有效的工具，如果没有很多经过实践证明的方法，这些工具的开发对于较新的方案领域极其重要。近几年，儿基会增加了其在整个儿童保护重点领域内对监测和评价以及方法研究的投入。

84. 要想通过多指标类集调查等住户调查进行监测，必须有效应对侵犯行为的隐蔽性和敏感性。尽管如此，还是取得了极大的进步。2000年以来，已经为多指标类集调查增加了六个模块：切割女性生殖器官；惩罚儿童；童工；出生登记；早婚或童婚；以及对待家庭暴力的态度。出版第一套具有可比跨国数据的出版物的基础是拥有来自最高质量的“黄金标准”监测的数据。关于儿童保护的总体情况以及儿童纪律的报告已经出版，并且已计划在2012-2013年期间出版关于切割女性生殖器官、童婚和脆弱性的报告。

85. 监测方面的进步尚未涉及家庭以外的弱势群体，包括生活在街头的儿童以及监狱中或机构照看的儿童。关于校园暴力、针对残疾人的暴力、移民当中的暴力或因帮派活动所致暴力的研究代表性也不充分。特别麻烦的是，缺少关于机构（校园、监狱、家庭护理的替代机构）内暴力的可靠数据，因为它是在国家以及收集这些数据的提供者的权力范围之内的事情。

86. 为支持开发新的数据收集工具和方法，进行了大量投资。与包括联合国维持和平行动部、人权事务高级专员办事处、开发署、世界卫生组织、美国疾病控制中心及各种学术机构在内的各种合作伙伴一起，共同制定了关于青少年司法、正规护理、法治以及武装暴力的全球指标。儿基会与负责暴力侵害儿童问题的秘书长特别代表建立了特别密切的关系，与其开展广泛合作，以期实现加强数据、研究和知识管理的共同目标。

87. 虽然采取全盘性做法比较罕见，也比较困难，但东部和南部非洲区域将在2012-2013年执行一项五国业务研究工作，以期加强国家儿童保护监测和评价制度。

评价的使用情况

88. 在总部一级，有关预防暴力侵害儿童方案的评价被用于大规模吸取教训的活动。在2008年和2011年组织的两次元评价活动中，儿童保护是唯一的内容。儿基会还为多个机构系统审查提供评价，最后形成一份报告——《我们在社区保护儿童活动中中学到什么？对有关社区儿童保护机制的证据进行一次机构间审查》（2010年），并导致肯尼亚和塞拉利昂目前也在开展一项机构间行动研究。

89. 在国家一级，普遍遵守了管理方反应要求，运用评价成果的方式则多种多样。在包括埃及、冈比亚、塞内加尔和塞尔维亚在内的很多国家，都有国家利用的有力证据。

90. 西非和中非区域最近对有关贩运儿童的方案进行了评价，结果导致对有关保护贩运受害者的区域指南进行了修改，并且发起了一项关于“移徙中的儿童”的多机构倡议，该倡议涉及移徙和贩运问题。

建议

91. 虽然是一个相对较新且成长较快的部门，但涉及暴力侵害儿童问题的方案规划正在对评价进行大力投入，包括从指标和工具到审查外地成果。已经形成一个重要的证据库，实践者也都渴望使用它。注意力被适当地重点放在相关性、战略有效性和成果衡量方面。所有这些良好做法都应该继续。

92. 对于本报告中指出的一些主要差距，建议儿基会优先采取四项行动：

(a) 加强监测和评价工作人员以及儿童保护工作人员的规划、监测和评价能力。儿童保护部门应该与方案指导、人力资源以及监测和评价工作人员进行协商，以确定可以为从事儿童保护及监测和评价领域的工作人员提供何种额外的支助、培训或技术指导。使现有工作人员进一步意识到且能够执行其提供良好证据的职责，这是可以采取的最关键的行动。

(b) 审查制度做法。应该对《2008 年儿童保护战略》中拟议的增强社会变革和制度部分的核心战略进行审查，并利用新的知识和外地证据作进一步的阐述。关键问题是确保纳入所有已知重要的方面，提供有关如何适应特殊情况的指导，并且采取一种联合办法将各种制度与社会变革结合起来。

(c) 开展全球一级的评价。明显缺少对预防暴力侵害儿童相关方案规划的全球一级评价。方案规划的范围要求进行共同评价。同样，主要评价应该在相关时涉及暴力问题。

(d) 审查儿童安全防护政策。儿基会应该审查其儿童安全防护政策，以确保对儿基会支助的各种方案伙伴及政策的运行环境而言，政策是充分的。此外，还应该审查各项政策的执行情况，包括培训、纳入伙伴关系协议之中以及监测和审查的充分程度。

八. 决定草案

93. 儿基会建议执行局核可以下决定草案：

执行局

1. 注意到关于儿基会评价职能和主要评价工作的年度报告（E/ICEF/2012/13）；
2. 重申儿基会评价职能发挥的中心作用及儿基会评价政策（E/ICEF/2008/4）所列各项原则的重要性；
3. 欣见报告中所述的表明持续地在权力下放层次加强评价职能的证据，并鼓励儿基会加快进度；
4. 注意到跟踪评价制度有效性的关键业绩指标以及关于人力资源和财政资源的行政数据；
5. 注意到为增强国家伙伴的技能、能力和制度以及加强南南学习而做出的积极努力；
6. 欣见年度报告按专题介绍了近期在儿童保护和预防暴力侵害儿童行为以及在管理方反应方面的评价成果；

7. 请儿基会：

(a) 报告为确保在拟订关键政策、战略和方案时系统地考虑和利用相关评价成果而采取的措施；和

(b) 报告管理方对预防暴力侵害儿童相关方案规划专题审查的执行的执行情况。
