



联合国儿童基金会

执行局

2022 年年会

2022 年 6 月 14 至 17 日

临时议程*项目 7

2021 年儿基会评价职能年度报告

摘要

本报告概述了 2021 年儿基会评价职能的情况，报告概述了儿基会在 2019 冠状病毒病疫情暴发以来所面临的充满挑战的业务背景，以及为继续满足儿基会及其利益攸关方的学习和问责需求而采取的适应措施。报告同时概述了 2021 年评价职能的绩效，并分析了 2018-2021 年四年期的成就和经验教训。

报告最后对评价职能和所汲取的经验教训进行了前瞻性思考，特别是对于评价职能如何确保在下一个四年期符合目的，以支持联合国可持续发展目标“行动十年”计划而言，这些经验教训意味着什么。

第六节载列供执行局审议的决定要点。

*E/ICEF/2022/9.

说明：本文件全文由儿基会负责处理。



一. 概述

1. 2021 年，儿基会继续面临 2019 冠状病毒病疫情和实现可持续发展目标的进展持续减弱的双重挑战，评价职能也是如此。尽管面临众多重大挑战，但评价职能部门还是适应了变化，仍然坚如磐石。2019 冠状病毒病疫情所带来的限制催生了一系列创造性的解决方案，包括更多地使用实时评价、实时评估和远程数据收集模式。此外，增加评价支出有助于在实现儿基会评价政策（E/ICEF/2018/14）中概述的承诺方面取得显著成果。提交的评价数量增加了，地理覆盖面也扩大了，而且评价职能使用了更加多样化的产品，使各办事处在充满挑战的情况下仍然能够开展评价工作。
2. 2021 年的评价质量也保持在较高水准。同时，也发现了一些需要继续改进的方面，包括那些对确保评价得到利用最关键的方面：执行摘要和经验教训部分。此外，尽管为应对 2019 冠状病毒病疫情而提供的成套评价解决方案有所扩大，但非传统的评价工作（如影响评价、可评价性评估以及联合评价和机构间评价）所占比例仍然相对较小，政府参与评价的情况也是如此——鉴于联合国可持续发展目标“行动十年”计划的紧迫性，所有这些领域都将至关重要。最后，虽然对管理层回应这一要求的遵守率很高，但在某些情况下，管理层回应的及时性有所下降，而且高遵守率是否意味着会对相应的评价建议采取一贯有意义的行动，仍是一个未知数。
3. 在此背景下，现在比以往任何时候都更加重要的是，评价职能要完全适应儿基会的问责和学习需求，提供证据以加强儿基会的应对工作，并帮助儿基会和合作伙伴加快为各地所有儿童取得进展。因此，本报告的关键要点提供了重要的经验教训，将为 2022-2025 四年期的前进方向提供参考。首先，有必要更加重视结果和影响级别的评价，帮助儿基会和合作伙伴衡量儿基会儿童权利工作的最终成果和实现可持续发展目标的进展情况。同时，评价职能必须开展更迅速、更轻便的工作，以帮助最大限度地提高儿基会及其合作伙伴在第一时间取得成功的可能性。其次，不能把稳定的评价质量看作理所当然；如果要保持高质量，就需要保持警惕并制定创造性的解决方案。第三，必须更加关注整个组织针对评价的后续行动，包括管理层作出及时、有力的回应并且相关回应得到积极落实和监测。最后，儿基会如果要履行根据执行局在这方面的决定所作的承诺，如果要继续在实现为评价职能设定的业绩目标方面取得进展，如果要成功实现为评价职能设定的战略方向，就需要持续关注评价职能是否得到充分的资源。
4. 在确保评价职能能够发展、改进和应对前所未有的挑战，以便支持“行动十年”方面，这些措施和其他措施都至关重要。将于 2022 年委托对评价进行独立的外部同行审查，进一步探索这些领域，并在此过程中为 2023 年修订儿基会评价政策提供信息。

二. 引言

5. 2021年伊始，形势就极具挑战性。随着大部分可持续发展目标无法如期实现，世界仍在应对2019冠状病毒病疫情，联合国可持续发展目标“行动十年”计划的紧迫性凸显出来，儿童权利受到的威胁程度为几代人所未见。由于疫情暴发，来之不易的进展被中断，在某些情况下甚至遭到削弱。与此同时，旷日持久的武装冲突、粮食不安全、贫困程度的上升以及气候变化引起的自然灾害，继续对实现儿童权利和加速实现可持续发展目标的进程构成重大威胁。在此背景下，现在比以往任何时候都更加重要的是，评价职能的工作要完全适应儿基会的问责和学习需求，提供证据以加强儿基会的应对工作，并帮助儿基会和合作伙伴加速实现世界各地所有儿童的权利。

6. 尽管疫情在2021年带来了许多挑战，但儿基会还是取得了一些重要成就。其中之一是评价工作的绝对支出增加，与2018年相比，2021年的支出增加了1500万美元。中东和北非地区以及南亚地区的支出增长尤其显著，无论是从绝对支出还是从占方案支出的百分比来看，这些地区的投资都有所增加。2021年，评价的地理覆盖率也有所提高，特别是在覆盖率目标历来很难实现的中东和北非地区。

7. 也有迹象表明，评价的质量有所提高。就评价方法和结果而言，全球评价报告监督系统的质量分数稳步提高，而建议的质量仍然很高。这些进展提高了评价的可信度和效用，这表现在提交评价管理层回应的遵守率提高，2021年达到近100%。重要的是，一系列极具影响力的评价——即根据其建议采取关键行动的评价，以有意义的方式为儿基会的工作提供了信息。

8. 本着联合国改革议程的精神，并根据联合国系统发展方面业务活动四年度全面政策审查（大会第75/233号决议），与联合国其他机构联合管理的评价数量有所增加。这一趋势包括在分散的层面上进行联合评价的努力。

9. 在许多情况下，这些成就是2018-2021四年期形成的积极趋势的顶点，是随着评价职能的不断成熟和发展而建立的坚实基础。评价职能现已准备好加强对进一步提高评价质量的重视，与儿基会管理层一道提高管理层回应的质量和及时性，并利用评价证据为决策提供信息。

10. 对关键业绩数据的审查表明，在这些领域还需要做更多的工作。有必要在迄今取得的进展的基础上，应对剩余的挑战，并在2022-2025四年期打造新一代的评价职能。因此，2018-2021年四年期的结束标志着一个关键时刻，可以借此总结过去四年的成就和经验教训，并思考评价办公室和整个评价职能如何能够改善评价证据的质量、及时性和使用，以支持联合国可持续发展目标“行动十年”计划。

11. 本报告审查了关于2021年评价职能状况的数据，既对照既定的关键业绩指标，又以四年期的前三年为基准，审查了其业绩。报告简要分析了影响评价职能的各种力量，以及评价职能如何应对挑战和把握机遇。报告还概述了关键业绩指标的进展情况，并分析总结了2022-2025四年期的主要成果。整个文件突出强调了对儿基会

2021 年工作有特别影响的评价实例。这些“2021 年有影响力的评价”展示了被认为高度融入了性别平等视角，并且为儿童带来了积极和切实变化的高质量工作。

三. 儿基会的业务背景：充满挑战和创造性解决方案的一年

12. 在 2019 冠状病毒病疫情持续肆虐期间，仍然需要在 2021 年调整传统评价方法，以期应对前所未有的挑战。为了将工作转向与 2019 冠状病毒病有关的评价，已计划的几项评价被搁置，以期就儿基会如何应对这场疫情、如何更好地防范未来的突发公共卫生事件总结经验教训。例如，对乍得 2019 冠状病毒病疫情应对工作进行了实时评价，为通过分发粮食和现金转移帮助最弱势人群的干预措施提供了信息，对儿基会在加蓬的 2019 冠状病毒病疫情应对工作进行了实时评价，为该国制定新的国家合作方案的提供了参考。

13. 同时，许多没有明确侧重于 2019 冠状病毒病疫情应对措施的评价也纳入了与 2019 冠状病毒病疫情相关的调查内容。例如，对支持少女的阿富汗妇女领导倡议的总结性评价范围进行了更新，纳入了一项分析，即如何提供关键用品，以帮助受 2019 冠状病毒病影响的少女做好重返学校的准备。对儿基会在蒙古的地理定位方法的评价也扩大了范围，以评估疫情对儿童的影响。

14. 在促使评价职能部门创造性地思考和采取行动方面，这场疫情还促成了一系列评价工作，包括 2021 年的实时评估和社区快速评估。根据以往在人道主义评价方面的经验，这些以及其他针对疫情相关限制而推出的创造性解决方案表明，不进行现场数据收集也可以进行评价。

15. 虽然事实证明这些调整对满足儿基会在充满挑战的时期的学习和问责需求至关重要，但它们也存在较大的局限性。在某些情况下，与关键利益攸关方和权利持有人的接触机会减少，影响了最难接触到的群体和面临风险的儿童的均衡代表性。缺乏实地接触意味着评价人员与他们工作环境的互动有限，而且不得不依赖不太理想的数据集。评价职能部门尽一切努力考虑相关的风险和限制，并调整抽样战略，以减少缺点和意外影响及偏差。然而，在许多情况下，远程数据收集是应对当前挑战的一种变通方法。虽然这种方法对满足儿基会的评价需求至关重要，但它不能长期替代更传统、更可靠的利用证据的方法。

方框 1

2021 年有影响力的评价：将性别问题纳入儿基会 2019 年冠状病毒病疫情应对工作中

2019 年冠状病毒病疫情又增加了一层障碍，有可能使妇女和女童进一步与信息、服务和就业隔离。务必了解儿基会 2019 年冠状病毒病疫情应对措施在多大程度上适应和考虑了妇女和女孩所面临的独特挑战。采用参与式、共同设计和共同解释的方法，确保了儿基会国家办事处和执行伙伴在应对过程中能够获得“对儿基会南亚 2019 年冠状病毒病疫情应对工作中纳入性别平等问题及其有效性的实时评价”的结果，使他们能够进行中期调整 and 适应。

四. 评价职能业绩概览：正在取得的进展和关键经验教训

16. 对过去一年和 2018-2021 四年期整体关键业绩指标相关数据的分析表明，评价职能在艰巨的挑战面前仍然坚如磐石。所取得的进展在很大程度上归功于四年期间所做的大量投资。增加评价支出后，在履行儿基会订正评价政策（[E/ICEF/2018/14](#)）中概述的承诺方面有明显改善。

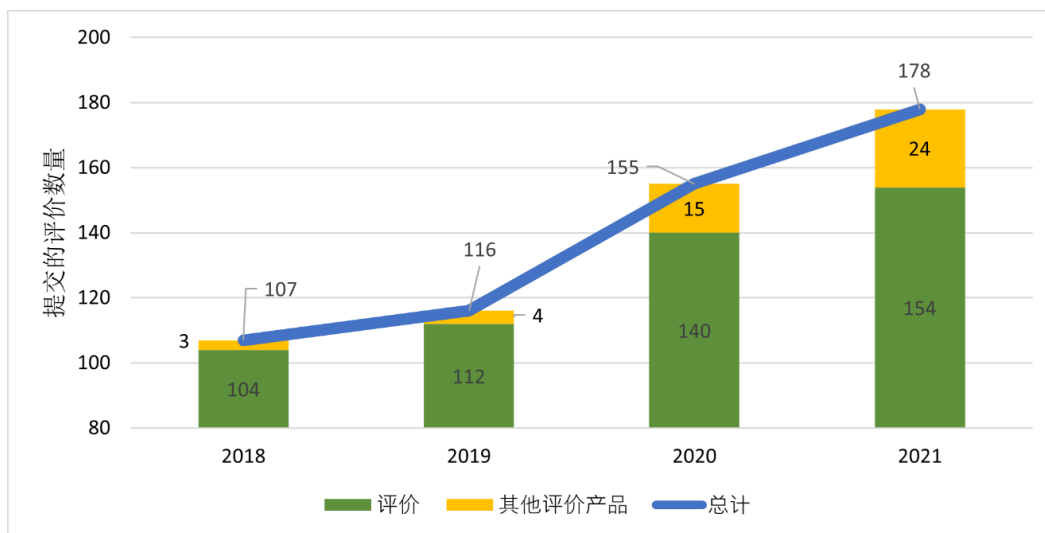
17. 本节分析了在评价提交、地理覆盖范围、质量和使用方面的表现，并概述了为加强评价职能所做的投资。

A. 提交的评价数量和地理覆盖范围

18. 儿基会订正评价政策中规定的覆盖范围规范旨在确保儿基会对与《2022-2025 年儿基会战略计划》的目标领域有关的方案、战略和政策进行足够数量的评价，以便为政策、实践和决策提供信息。2021 年，提交的评价数量和地理覆盖范围都有所增加。在前几年势头的基础上，儿基会的评价职能部门提交了有史以来数量最多的评价报告。

19. 儿基会的评价工作在儿基会的三个层面（总部、区域和国家层面）开展，其中大部分在国家层面开展。在 2021 年提交的 178 个评价产品中，有 14 个在总部层面开展，164 个在分散层面开展。如图 1 所示，2021 年产生的评价产品比 2020 年多 23 个。在整个四年期间，评价产品的数量增加了 67%。

图 1
2018-2021 年提交的评价数量



资料来源：证据信息系统集成（EISI）。

20. 在整个 2021 年，2019 冠状病毒病疫情以及越来越多的 3 级紧急情况，继续强调了生成证据的重要性，这些证据不仅足以决策提供信息，而且可以快速、高效和具有成本效益地生成。评价职能以一系列创造性的解决方案回应了这一需求，包括更多地使用实时评价和实时评估。

21. 2021 年进行了 6 次可评价性评估。这些评估力求通过衡量方案的准备程度来加强方案的规划和设计，以便在日后进行有意义的评价。此外，还进行了两次审查，旨在就教育管理信息和监测系统的方案执行情况提供及时反馈。这些审查的结果对于推广方案至关重要。2021 年提交了 14 项实时评估，而 2020 年提交了 5 项；所有这些评估都与 2019 冠状病毒病疫情有关，有助于为国家和区域 2019 冠状病毒病疫情应对措施提供证据基础。

22. 这套扩大的评价产品是确保管理人员有一系列工具来帮助他们产生证据以改进方案编制的核心。然而，随着灵活性的增加，需要更加明确如何在不牺牲评价的独立性、质量和可信度的情况下进行适应性调整。今后，需要更明确地区分独立评价工作和其他类型的工作。还需要有更明确的指导，说明哪些类型的工作有资格获得评价政策中指定用于评价的 1% 的资金，并计入执行局第 2018/10 号决定规定的 1% 支出基准。

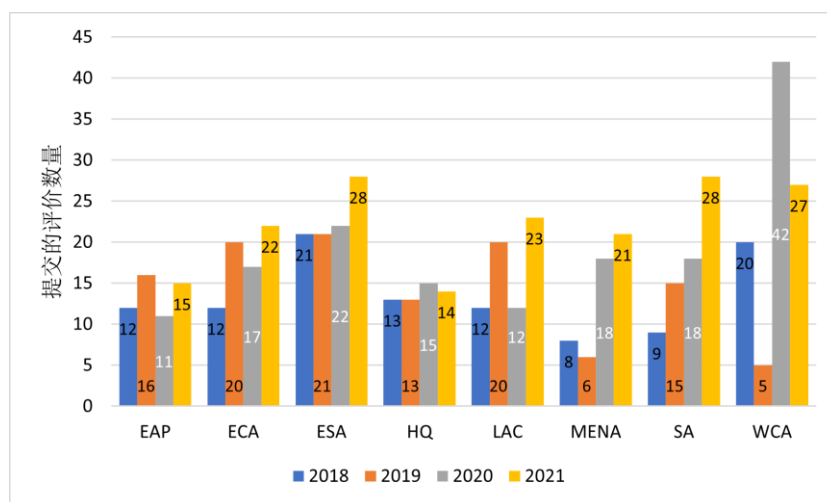
23. 如表 1 和图 2 所示，2021 年，评价职能的所有层面都出现了这两个收入趋势——产出增加和转向更加多样化的成套评价产品。除总部以及西非和中非地区外，所有地区提交的评价数量都超过了往年。东部和南部非洲地区以及南亚地区提交的评价数量最多，而总部提交的评价数量最少。

表 1
2021 年提交的评价和其他评价产品的数量，按地点分列

地点	提交的产品				总计
	评价	可评价性评估	审查	实时评估	
EAP	12	1	0	2	15
ECA	19	1	0	2	22
ESA	23	0	0	5	28
HQ	7	2	2	3	14
LAC	22	0	0	1	23
MENA	20	0	0	1	21
SA	25	1	2	0	28
WCA	26	1	0	0	27
总计	154	6	4	14	178

图例：EAP：东亚和太平洋地区；ECA：欧洲和中亚地区；ESA：东部和南部非洲地区；HQ：总部；LAC：拉丁美洲和加勒比地区；MENA：中东和北非地区；SA：南亚地区；WCA：西部和中部非洲地区。
资料来源：EISI。

图 2
2018-2021 年提交的评价产品数量，按地点分列

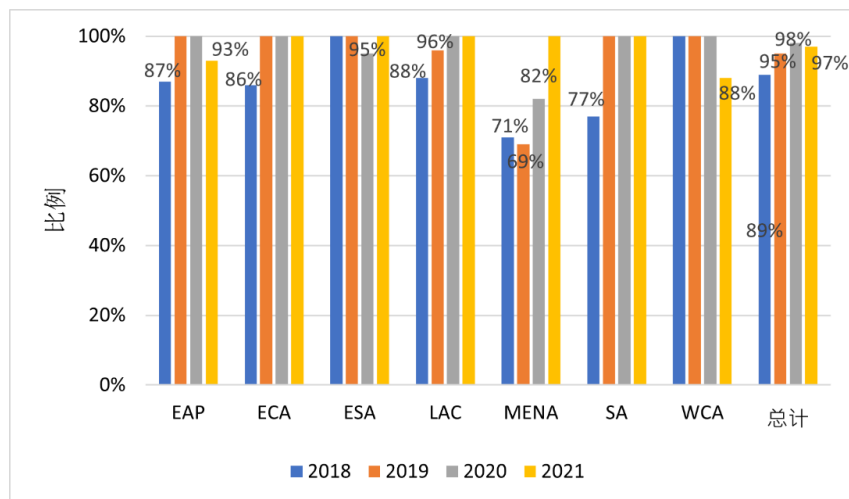


图例：EAP：东亚和太平洋地区；ECA：欧洲和中亚地区；ESA：东部和南部非洲地区；HQ：总部；LAC：拉丁美洲和加勒比地区；MENA：中东和北非地区；SA：南亚地区；WCA：西部和中部非洲地区。
资料来源：EISI。

24. 如图 3 所示，尽管暴发了 2019 冠状病毒病疫情，但从 2020 年到 2021 年，各区域的总体地理覆盖率（指已开展评价的国家和区域办事处）仅略有下降（从 98% 降至 97%），这归功于使用了远程技术和其他变通方法。此外，多国评价（指各办事处汇集财政和人力资源开展评价）使评价工作得以在资源有限的情况下进行。7 个

地区中有 5 个地区达到了 100% 的覆盖率，东部和南部非洲地区以及中东和北非地区的情况也有所改善。东亚和太平洋地区以及西部和中部非洲地区的覆盖率有所下降，因为前者有一个国家，后者有三个国家自 2019 年以来没有进行评价。据设想，覆盖率低的国家办事处将在 2022 年进行评价。

图 3
各区域的评价覆盖率，2018-2021 年*



* 条形图上方如果没有数据标签，就意味着实现了 100% 的地理覆盖率。

图例：EAP：东亚和太平洋地区；ECA：欧洲和中亚地区；ESA：东部和南部非洲地区；HQ：总部；LAC：拉丁美洲和加勒比地区；MENA：中东和北非地区；SA：南亚地区；WCA：西部和中部非洲地区。

资料来源：EISI。

B. 评价对可持续发展目标的影响和集体贡献

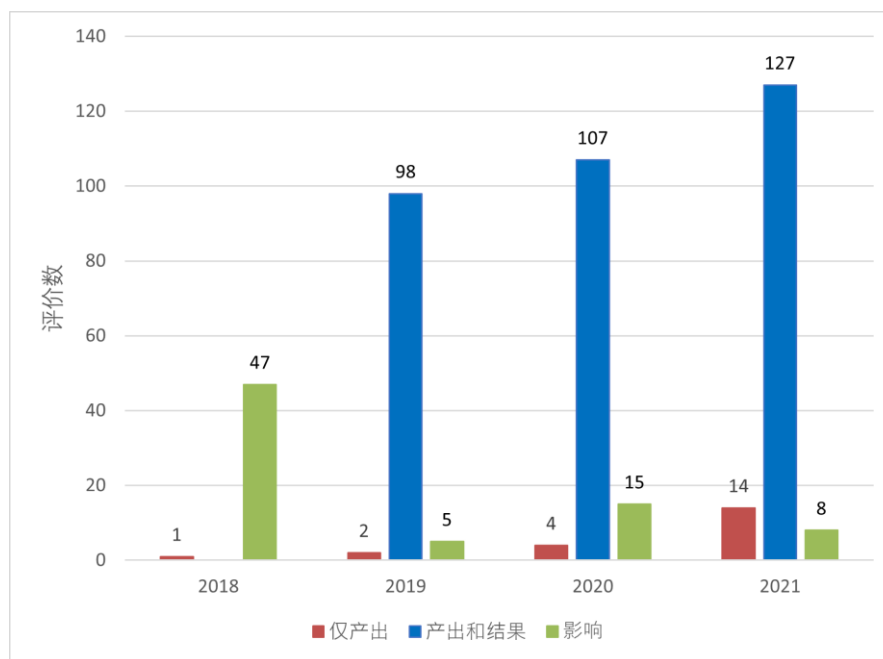
25. 虽然这场疫情突出了加快证据生成的必要性，但也凸显了跟踪可持续发展目标进展的必要性。这方面一个令人担忧的趋势是，自 2018 年以来，影响和结果级别的评价数量急剧下降，当时近一半的评价包括影响级别的成果。其中部分原因是目前使用的“影响”定义与前几年相比更为严格，即只有涉及对照组的实验设计才被列为影响。此外，由于以前的全球评价报告监督系统模板没有对结合了产出和结果的评价进行分类的选项，因此在评估两个级别的成果时，似乎都选择了“结果”。

26. 展望未来，将根据衡量组织和方案效果的成果链的层级来界定影响，而不是根据所使用的具体方法——也就是说，根据经济合作与发展组织发展援助委员会对影响的定义：“干预措施已经产生或预计将产生重大积极或消极、预期或非预期的更高层次效果的程度”。¹通过强调“更高层次的效果”——即通过经历特定干预的人的生活或经历中出现的物质、系统和/或持久变化来衡量的效果——将影响与结果区分开来，后者侧重于需要首先发生的中介或工具性变化（例如，颁布的政策变化和其他规范性指导、应用的技术援助、使用的服务），以促进实现最终期望的影响。

¹经济合作与发展组织。《采用更好的准则实现更优质的评估：评估准则的概念界定和应用原则（修订版）》（巴黎，发展援助委员会发展评估网络，2019 年 12 月）。

27. 同时，在已评估的 154 项评价中，有 8 项为影响一级，132 项为结果和产出一级，只有 14 项为产出一级。对产出和结果的关注是 2019 年以来人们注意到的一个趋势。这表明传统的项目监测可能存在差距，因此需要开展着眼于产出和结果的形成性评价。预计将在结果和影响级别开展更多的评价，使儿基会能够更好地确定其干预措施的效果。

图 4
2018-2021 年，按级别分列的评价重点



资料来源：GEROS。

28. 与此同时，2021 年产生的评价很好地结合了回顾过去的总结性评价和展望未来的形成性评价。超过一半的评价（53%）既是总结性又是形成性评价，而大约三分之一（34%）纯粹是形成性评价。12 项评价是纯粹的总结性评价，还有一项元评价。这种混合目的表明，儿基会的问责和学习需求得到了同等程度的满足。

表 2
2021 年开展的评价，按类型分列

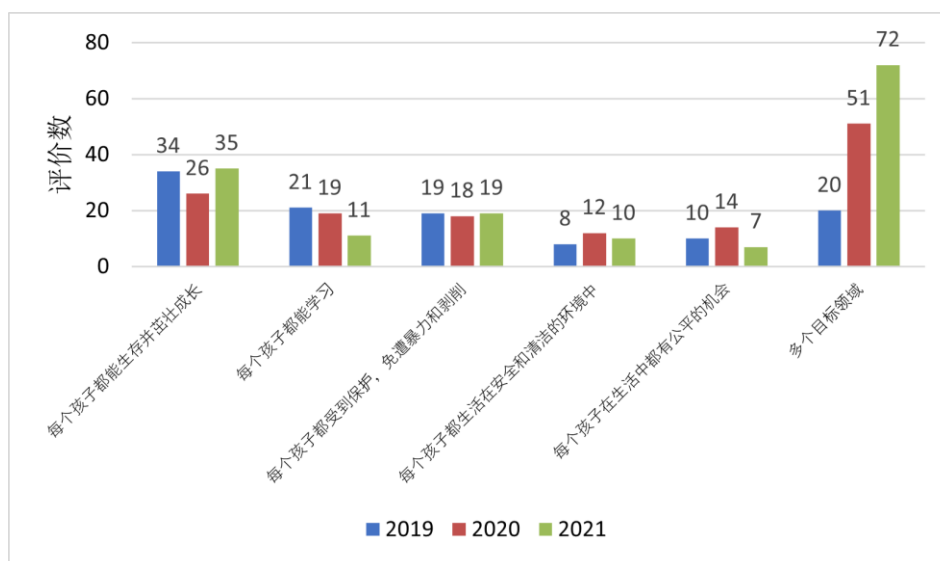
评价类型	评价数	占有所有评价的比例
形成性	53	34
总结性	19	12
总结性和形成性	81	53
元评价	1	1

资料来源：GEROS。

1. 目标领域分布

29. 自 2019 年以来，随着区域和国家办事处转向开展为整个国家方案提供信息的更具战略性的评价，而不是评价单一的干预措施，涉及儿基会《2018-2021 年战略计划》的多个目标领域的评价数量逐年增加。

图 5
2019-2021 年评价的专题覆盖面



注：2018 年的数字不详，因为对专题覆盖面的系统跟踪是在 2019 年首次推出的。
资料来源：GEROS。

30. 2022-2025 年全球评价计划（E/ICEF/2022/3）旨在确保广泛覆盖所有五个目标领域以及人道主义行动和跨领域主题，因此预计目标领域覆盖面上的平衡将继续得到改善。

2. 联合评价和机构间评价

31. 儿基会各级评价工作人员继续积极参与旨在加强整个联合国系统评价工作的联合国评价小组及其工作组。其中包括为指导文件建言献策，并分享经验以促进学习。

32. 除了这种规范层面的总体合作之外，自 2018 年以来，与一个或多个联合国机构联合管理的评价也有所增加。在 2021 年与其他联合国机构联合管理的 10 项评价中，有两项为全球层面：联合国系统和开发银行实现可持续发展目标 6 的工作的机构间评价综合，以及对人口基金-儿基会消除女性生殖器切割联合方案：加速变革-第三阶段-2018-2021 年的联合评价。

33. 特别值得注意的是在科特迪瓦、加纳、尼日利亚、摩尔多瓦共和国、南苏丹、斯里兰卡和津巴布韦等国家，在分散层面为开展联合评价所付出的努力。例如，在摩尔多瓦共和国开展的评价是一项全系统评价，评估联合国系统对国家发展优先事

项的集体反应。这类积极趋势需要在“行动十年”中加速，以支持收复失地的努力，并使实现可持续发展目标的进展走上正轨。

34. 迄今为止，发布管理层对联合评价和机构间评价的联合回应的做法各不相同。诸如此类的评价并不总是引发联合管理层回应，即使有，也往往不够规范，后续行动也薄弱无力。然而，确实存在强有力的联合管理层回应的正面例子，特别是与机构间人道主义评价有关的例子，可以从中挖掘出最佳做法。迄今为止获得的经验表明，在建立了正式的联合/机构间问责机制（如机构间常设委员会和《人人享有健康生活和福祉全球行动计划》）的地方，联合管理层回应最为有力。鉴于对联合行动和机构间行动的高度重视，并根据四年度全面政策审查，这类行动在 2022 年及以后将非常重要。

35. 在国家评价能力发展方面取得的进展较少。在全球评价报告监督系统评估的评价中，只有 4 项由东道国领导，11 项由东道国共同管理。虽然比前几年略有增加，但与评价政策中的承诺相差甚远，今后还需要进一步努力。

36. 评价职能继续支持各国评价自身实现可持续发展目标的进展情况。2021 年 3 月和 5 月，儿基会与联合国非洲经济委员会和非洲英语国家学习、评价和成果中心合作，连续第三年为政府和儿基会工作人员举办了关于将评价纳入非洲国家自愿审查的培训。2021 年 5 月 18 日至 20 日，儿基会与芬兰外交部、国际环境与发展研究所、国际发展评价协会、德国发展评价研究所和 EVALSDGs 一起，为非洲和拉丁美洲的政府官员提供了关于将评价与国家优先事项联系起来和实现可持续发展目标的培训。

37. 评价职能还继续支持众多国家（包括巴西、柬埔寨、斐济、几内亚、印度、科索沃²、马达加斯加和乌干达）开展公共政策评价、评价管理和国家主导的评价。

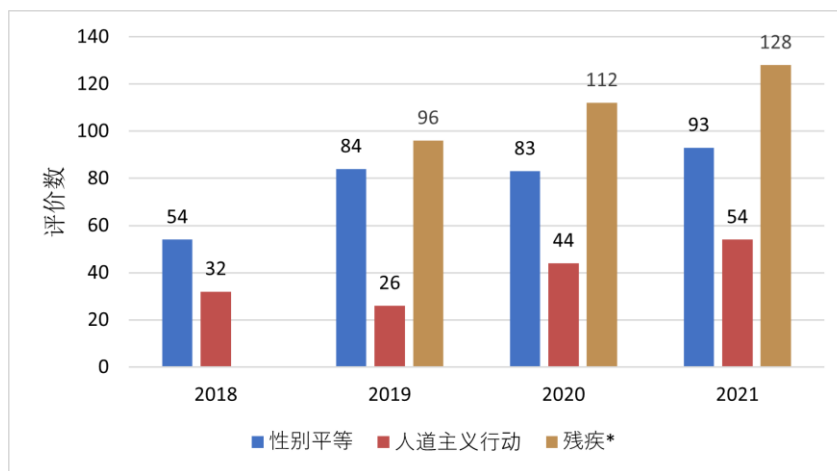
3. 评价对残疾、性别平等问题和人道主义行动的覆盖

38. 评价对残疾、性别平等问题和人道主义行动的覆盖是评价职能的关键重点领域。随着各办事处将残疾问题纳入所有政策和方案的主流，在评价中涵盖残疾问题的评价数量继续增加。评价小组正在具体考虑将残疾问题纳入评价设计、结论和建议中。为确保不让任何一个人掉队，评价人员与残疾人和/或其家人进行焦点小组讨论，并按残疾状况对数据进行分类，以便在评价报告中充分反映残疾人的需求。评价办公室正在进行一项关于将残疾问题纳入评价的综合工作，因此将为评价职能制定政策指导。此外，正在将关于儿基会包容残疾儿童这一工作的一项跨部门评价列入范围。

39. 与 2020 年相比，2021 年涵盖性别平等问题和人道主义行动的评价数量有所增加。93 项评价纳入了性别平等问题，而 2020 年为 83 项。54 项评价将人道主义行动作为贯穿各领域的主题，而 2020 年为 44 项。在人道主义评价中，评价职能仍然致力于对受影响人口负责。确保问责制的方式是让受影响人群了解他们的权利和权益以及评价小组的预期行为标准，并让他们作为主要利益攸关方参与建议的制定。

² 所有提及科索沃的地方都应在联合国安全理事会第 1244（1999）号决议的范围内理解。

图 6
2018-2021 年涵盖残疾、性别平等问题和人道主义行动的评价数量

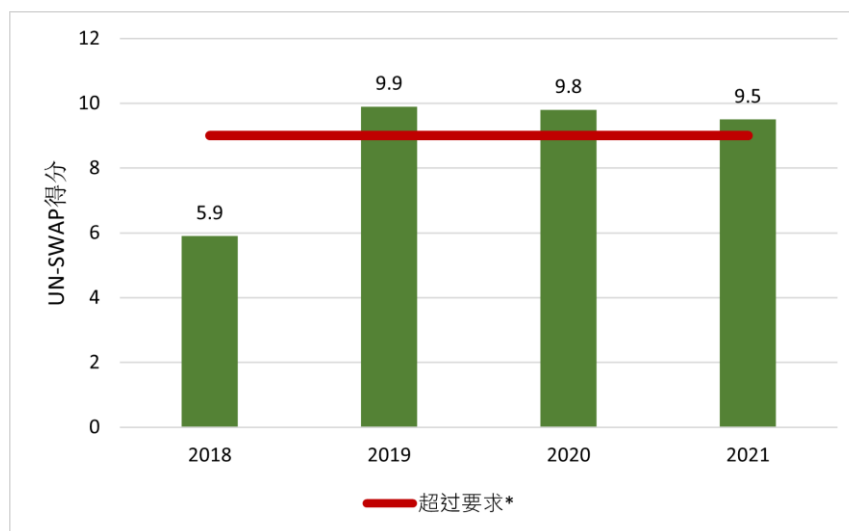


* 2019 年开始在全球评价报告监督系统中跟踪对残疾问题的覆盖情况。

资料来源: GEROS。

40. 在全球评价报告监督系统中进行的分析表明，在将性别平等和增强妇女权能纳入评价分析范围（评价框架中的评价目标、标准、问题/指标）方面取得了持续改进。在使用促进性别平等的方法、方法和工具以及数据分析方面，以及在评价结果、结论和建议中反映性别平等分析方面，取得的进展较少。儿基会在《联合国全系统性别平等和增强妇女权能行动计划》（UN-SWAP）中的表现仍然很好。如图 7 所示，在考虑了 2019 年对儿基会《2018-2021 年性别平等行动计划》的评价后，尽管个别评价的总体评估较低，但 2021 年评价组合的总体表现“超过要求”。

图 7
《2018-2021 年联合国全系统性别平等和增强妇女权能行动计划》下的评价业绩



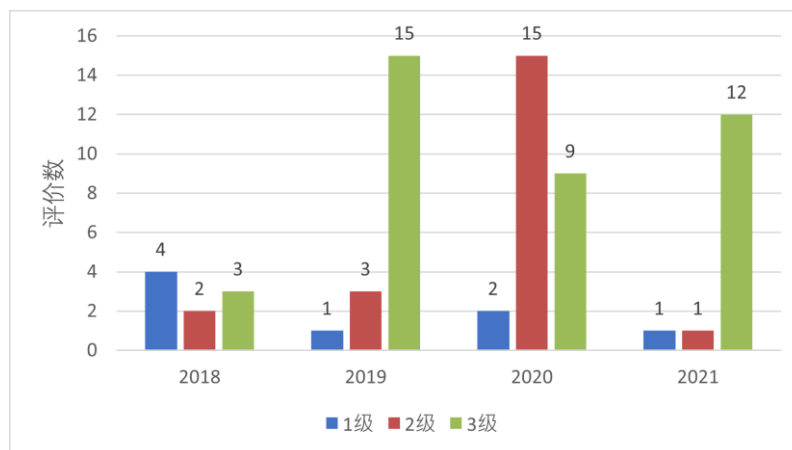
* 一个实体要被评为“超过要求”，评价报告的 UN-SWAP 得分必须达到 9.01 或以上。

资料来源: GEROS。

41. 评价职能在 2021 年仍然侧重于对已宣布的紧急情况的评价。如图 8 所示，2021 年进行了 14 项人道主义行动评价。这包括在目前被指定为三级紧急情况的国家进行的 8 项评价（即 4 项在阿富汗，1 项在南苏丹，3 项在也门）和 4 项侧重于 2019 冠状病毒病疫情的评价。此外，还分别对索马里和莫桑比克德尔加多角省的一级和二级紧急情况进行了评价。在这些评价之外，还对 2019 冠状病毒病疫情的应对工作进行了 12 项实时评估（未包括在图 8 中，因为它们没有被预先定义为评价）。

图 8

2018-2021 年进行的人道主义评价数量，按紧急情况级别分列

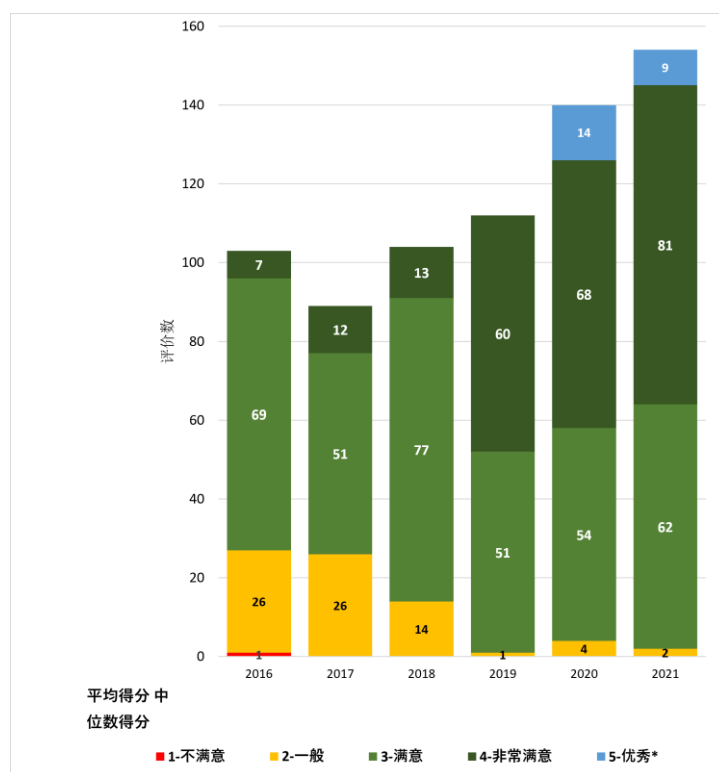


资料来源：EISI。

C. 评价质量

42. 2021 年的总体评价质量仍然很高，虽然 2020 年略有下降，但 2021 年的评价业绩恢复到疫情之前的水平。在由一家外部公司使用全球评价报告监督系统独立评估的 154 项评价中，9 项（6%）被评为“优秀”，81 项（53%）被评为“非常满意”，62 项（40%）为“满意”。只有 2 项评价（1%）被评为“一般”。没有任何评价得到“不满意”的评级，这一积极趋势自 2016 年以来一直持续至今。

图 9
2016-2021 年总体评价质量评级



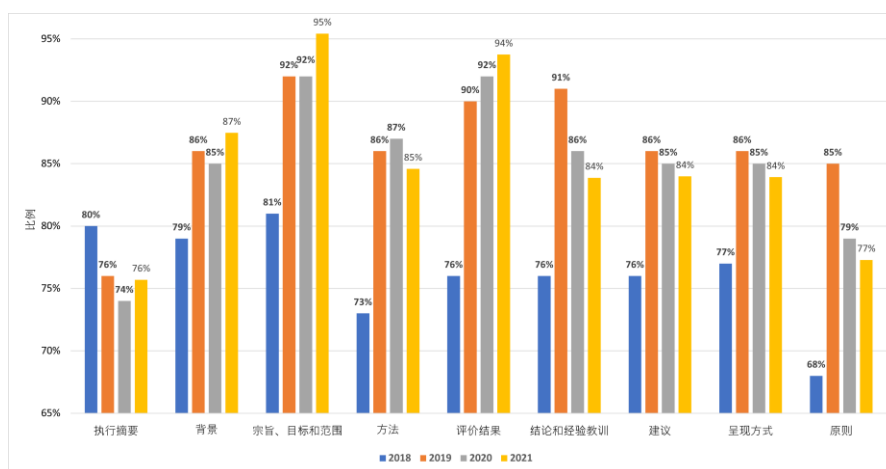
*“优秀”类别是在全球评价报告监督系统审查后于 2020 年周期推出的。

资料来源：GEROS。

43. 对评价报告的各个要素进行更细致的分析后，可以发现情况喜忧参半。2021 年一些评价的经验教训、执行摘要和评价原则部分的表现不如报告的其他方面。在这方面值得注意的是，在全球评价报告监督系统中占 5% 权重的经验教训部分，现已被认为是一项强制性内容。这一变动是在 2021 年推出的，以强调精心整理的经验教训对组织学习的重要性。此外，评价质量也受到执行摘要的影响，这些摘要要么太长，要么缺乏足够的细节而无法被视为独立的文件。此外，一些报告没有明确指出如何将残疾问题纳入评价；这一因素是最近纳入全球评价报告监督系统的一个维度，导致评价原则下的得分较低。

44. 在许多情况下，由于无法进行现场取证，评价的质量和可信度受损。例如，全球评价报告监督系统关于 2021 年评价组合的数据显示，与前两年相比，评价在方法上没有那么有力。

图 10
2018-2021 年评价质量评级平均得分，按报告章节分列



注：每个报告章节的质量得分从 0 到 100 不等（其中 0 对应的是每年包含的所有报告中可能的最低平均质量评级，100 对应的是完美平均质量评级）。

资料来源：GEROS。

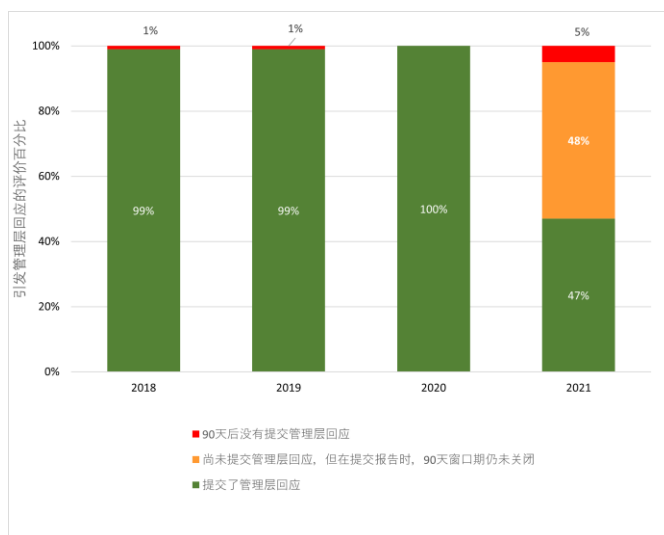
D. 确保组织从评价中汲取经验教训

1. 评价管理层回应的提交和执行

45. 管理层回应是确保评价建议得到落实的一个重要平台。儿基会的评价政策要求在证据信息系统集成（EISI）数据库中公布评价后的 60 天内提交管理层回应。虽然由于 2019 冠状病毒病疫情暴发，这一时间框架在 2020 年和 2021 年延长至 90 天，但在 2022 年又恢复至 60 天。

46. 总体而言，儿基会各办事处都按时提交管理层回应。在 2018 年至 2020 年期间进行的评价中，多达 99% 的评价有管理层回应。在 2021 年周期，管理层已经为 73 项（47%）评价提交了回应。在 81 项没有管理层回应的评价中，只有 7 项已经过了 90 天的窗口期，现已逾期。

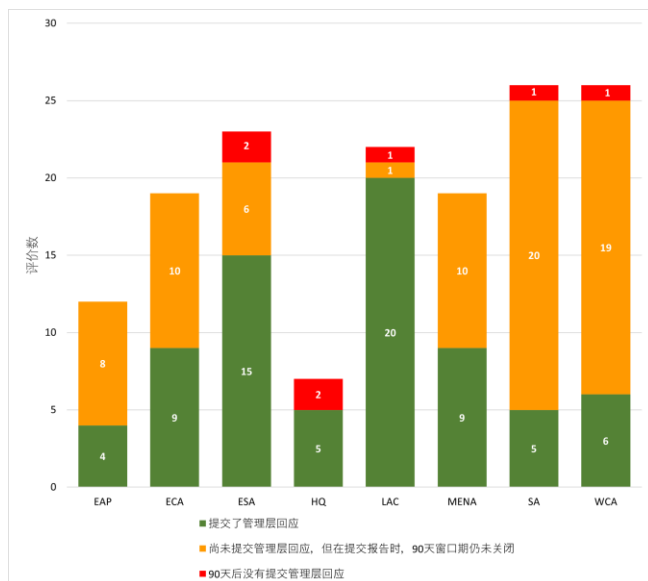
图 11
2018-2021 年遵守管理层回应要求的及时性*



* 由于 2019 冠状病毒病疫情暴发, 2020 年和 2021 年提交管理层回应的时限从 60 天延长到 90 天。
资料来源: EISI。

47. 在全球一级逾期的管理层回应过多, 而 7 个区域中有 3 个区域 (东亚和太平洋、欧洲和中亚, 以及中东和北非) 没有逾期的管理层回应。

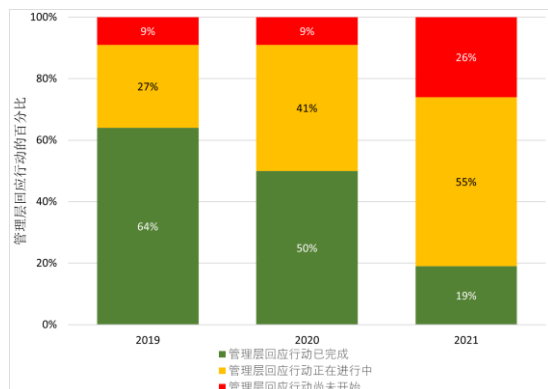
图 12
2021 年遵守管理层回应要求的及时性, 按地点分列



图例: EAP: 东亚和太平洋地区; ECA: 欧洲和中亚地区; ESA: 东部和南部非洲地区; HQ: 总部; LAC: 拉丁美洲和加勒比地区; MENA: 中东和北非地区; SA: 南亚地区; WCA: 西部和中部非洲地区。
资料来源: EISI。

48. 近年来，管理层回应中承诺采取的执行的执行率显著增加。在整个儿基会，针对 2019 年完成的评价的管理层回应行动的执行率目前为 91%（64%已完成，27%正在进行），只有 9%的行动尚未开始。对于 2020 年进行的评价，执行率为 91%（50%已完成，41%正在进行）。对于 2021 年进行的评价，执行率为 91%（50%已完成，41%正在进行）。

图 13
2019-2021 年管理层回应中确定的行动的执行状态



资料来源：EISI。

方框 2

2021 年有影响力的评价：将人道主义和发展方案编制联系起来

《对儿基会将人道主义与发展方案编制联系起来这一工作的形成性评价》为儿基会《2022-2025 年战略计划》和 2020 年四年度全面政策审查的措辞提供了信息，既在内部的儿基会规划、方案编制和报告方面，又在外部通过相关的联合国全系统机制强调“加强协作、一致性和互补性”以及“对建设和平和维持和平的贡献”。

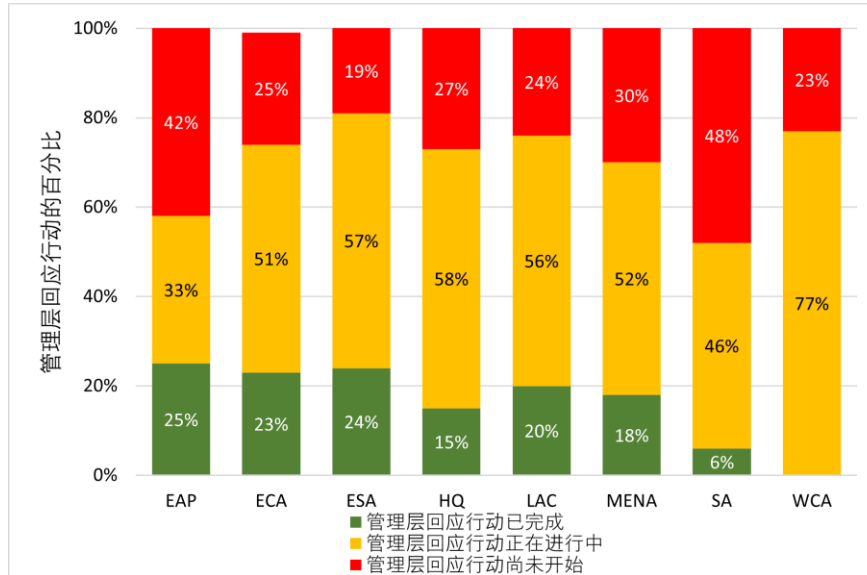
儿基会 2017-2021 年秘鲁国家合作方案的国家方案评价为 2022-2026 年国家合作方案的设计和实施提供了信息，有助于在国家和国家以下各级制定方案战略和确定实施重点。方案设计利用从 2019 冠状病毒病疫情中汲取的经验教训，明确阐述了将应急管理纳入国家方案。此外，评价还用于向秘鲁政府、秘鲁社会、国家合作伙伴、合作机构和国际捐助者说明资源的使用情况。

1998 年的《儿童法》对保护加纳儿童的权利至关重要，它明确规定了社会工作者在处理抚养费案件中的作用。然而，这导致社会工作者处理的案件数量过多。

《对儿童保护方案的形成性和总结性评价》概述说，应改用家事法庭系统。因此，修订了儿童保护和家庭福利部门间标准作业程序，并且儿童抚养案件现在被视为“低风险”，这意味着它们不应由社会工作者管理。由于这次评价，还提出了对《儿童法》的额外修订。

49. 从区域来看，对于 2019 年至 2021 年期间完成的评价，管理层回应行动的执行速度略有不同。然而，执行工作受到了 2019 冠状病毒病疫情的影响。随着疫情相关限制的取消，预计相关行动将在 2022 年重新获得势头。

图 14
2021 年管理层回应中确定的行动的执行状态，按地点分列



图例：EAP：东亚和太平洋地区；ECA：欧洲和中亚地区；ESA：东部和南部非洲地区；HQ：总部；LAC：拉丁美洲和加勒比地区；MENA：中东和北非地区；SA：南亚地区；WCA：西部和中部非洲地区。
资料来源：EISI。

方框 3
2021 年有影响力的评价：改善方案设计

儿基会 2017-2021 年秘鲁国家合作方案的评价为 2022-2026 年国家合作方案的设计和实施提供了信息，有助于在国家和国家以下各级制定方案战略和确定实施重点。方案设计利用从 2019 冠状病毒病疫情中汲取的经验教训，明确阐述了将应急管理纳入国家方案。此外，评价还用于向秘鲁政府、秘鲁社会、国家合作伙伴、合作机构和国际捐助者说明资源的使用情况。

2. 知识管理和评价结果的传播

50. 如果决策者不容易获得最高质量的评价，也不容易转化为广大利益攸关方的实际学习机会，那么即使是最高质量的评价也没有多大用处。此外，高度遵守管理层回应流程并不一定意味着管理层回应是有意义的，或有针对性地处理了相应的建议。展望未来，评价职能和管理层需要新思维，以确保评价工作在“行动十年”中为儿童实现最大的积极影响。

51. 确保管理层回应的质量和执行是儿基会管理层的责任。然而，评价职能可以做很多事情，以确保管理层回应的遵守率不仅高，而且有意义。如前一节所述，自2019年以来，对确保评价接受度最重要的两个领域——执行摘要和经验教训部分的质量有所下降。评价办公室将努力加强这些领域以及管理层回应本身的模板和指导。从2022年开始，评价办公室还将委托对当前管理层的回应及其执行情况进行审查，以便为管理层提供重要和有用的评估。与此同时，儿基会加强了 EISI 平台，纳入了就尚未做出的管理层回应和临近截止日期的行动发出通知这一功能。此外，在任何管理层回应结案之前，必须附上证明行动已经得到实施的书面证据，由相关办公室以外的审查人员核实这些行动已经得到完全实施。

52. 必须认识到，儿基会的其他部分，包括工作人员、合作伙伴、执行局和审计咨询委员会，也是遵循管理层回应程序和进一步加强整体组织问责制的关键行为者。因此，在当前信息超载的环境下，尽可能有效地宣传评价结果至关重要。为此，评价办公室扩大了其伙伴关系并加强了技术，以便使人们更容易获得评价证据。2020年9月启动的“全球共享发展平台”是一个可公开访问的平台，公众可通过公共领域获取评价证据。儿基会支持的这一平台的核心内容是提供评价证据，同时提供关于儿童和青年的外部合作伙伴证据。评价办公室参加了儿基会数字平台工作组，该工作组的目的是为数据、分析和知识管理提供数字解决方案。

53. 评价职能还维护着一个评价资源中心，在那里可以找到与开展评价有关的经验教训、技术文件和其他工具。同样，评估职能还维护着全组织范围的协作网站，这些网站位于 SharePoint 中。这些协作网站既是共同创造的空间，也是与评估网络研讨会和学习活动相关的文件、文章和所有材料的数字化参考文献库。

54. 2021年，评价办公室开展了一系列学习活动，包括网络研讨会（针对内部和外部受众）、全球学习活动（2021年有两次活动）、多媒体2019冠状病毒病学习周，以及汇集了儿基会、大学、非政府组织和来自全球南方的学生的各种“全球共享发展平台”活动。“全球共享发展平台”系列研讨会包括来自儿基会内部的专家，以及外部合作伙伴，包括国际劳工组织、联合国可持续发展解决方案网络、可持续发展目标青年学院以及印度、肯尼亚、巴基斯坦、秘鲁和美国的所大学。

方框 4

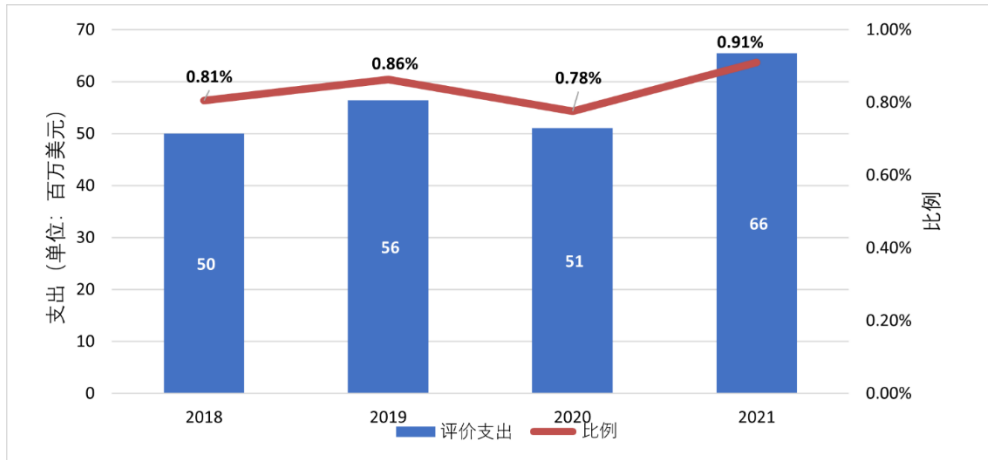
2021年有影响力的评价：通过评价加强儿童保护政策

1998年的《儿童法》对保护加纳儿童的权利至关重要，它明确规定了社会工作者在处理抚养费案件中的作用。然而，这导致社会工作者处理的案件数量过多。《加纳政府/儿基会儿童保护方案（2012-2019年）的形成性和总结性评价》概述说，应改用家事法庭系统。因此，修订了儿童保护和家庭福利部门间标准作业程序，并且儿童抚养案件现在被视为“低风险”，这意味着它们不应由社会工作者管理。由于这次评价，还提出了对《儿童法》的额外修订。

E. 财政资源和评价支出

55. 从 2018 年到 2021 年，儿基会的总体方案支出有所增加（从 2018 年的 62 亿美元增至 2021 年的 72 亿美元）。如图 15 所示，这意味着评价总支出增加。图 16 使用 2018-2021 四年期间使用的现行公式，概述了趋势。

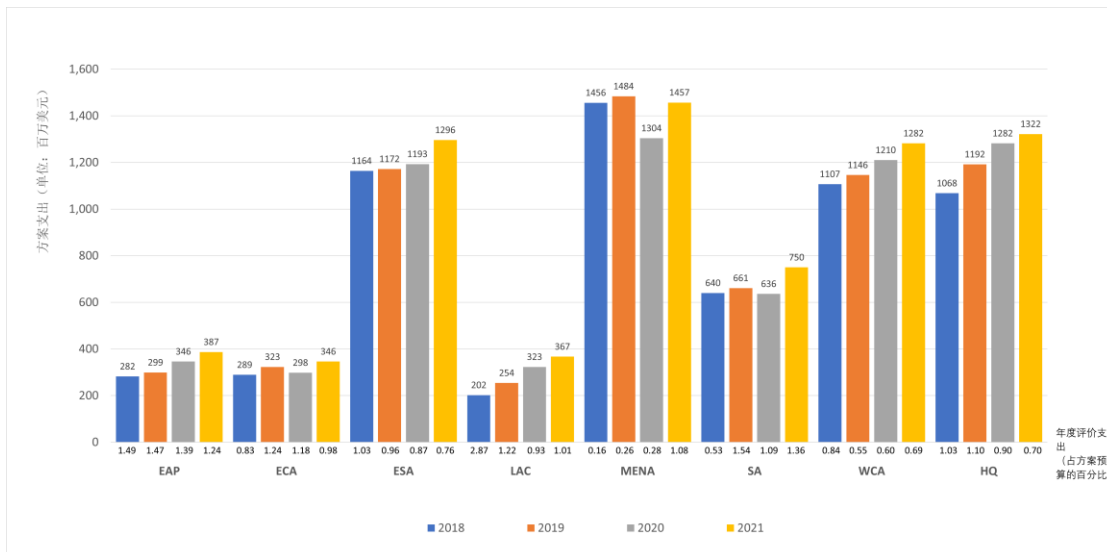
图 15
2018-2021 年儿基会评价支出，以美元计，占方案预算总额的百分比



资料来源: Insight 评价职能业绩仪表盘。

56. 与 2020 年相比，2021 年评价支出方面的业绩有所改善，评价支出占总体方案支出的 0.91%。在分散层面，所有地区都突破了 0.50% 的大关。中东和北非地区的支出增长尤为显著，该地区将 0.80% 的方案支出用于评价，而南亚地区的评价支出也大幅增加。

图 16
2018-2021 年方案支出和评价支出百分比，按地点分列*



*条形图和条形图上方的金额显示每个地区和年份的方案支出。每个条形图内的数字表示所示年份中，各地区用于评价的方案预算百分比。

图例：EAP：东亚和太平洋地区；ECA：欧洲和中亚地区；ESA：东部和南部非洲地区；HQ：总部；LAC：拉丁美洲和加勒比地区；MENA：中东和北非地区；SA：南亚地区；WCA：西部和中部非洲地区。

资料来源：Insight 评价职能业绩仪表盘。

57. 除了需要继续注意确保评价职能获得充足的资源外，在本四年期，还必须以其他联合国机构的趋势和现行做法为基准，衡量儿基会跟踪评价支出的进展情况。评价办公室已经开始汇编这些信息，为这一讨论提供信息。

58. 本报告概述的大部分进展在很大程度上归功于近年来评价工作绝对支出的总体增加（特别是在分散层面），这是因设立评价集合基金而促成的。展望未来，如果要使在分散层面取得的覆盖率增加趋势持续下去，就必须确保这一资金来源持续充足。

1. 人力资源能力

59. 为评价职能提供的财政资源有助于确保有足够的人力资源开展评价。所有地区都成功地招聘了多国评价专家，这些专家对评价工作的开展至关重要，特别是在资源有限的情况下。由这些专家服务的办事处现在可以集中资源进行评价。为了落实该政策的覆盖面基准，一些预算超过 1 亿美元的办事处也在确保有一个专职的评价专家，因为它们每年都要开展评价。

方框 5

有影响力的评价：对欧洲和中亚地区儿基会 2019 冠状病毒病疫情期间幼儿发展对策的多国评价

“对欧洲和中亚地区儿基会 2019 冠状病毒病疫情期间幼儿发展对策的多国评价：克罗地亚国家案例研究”提供了快速、循证的见解，以便逐步改进根据 2019 冠状病毒病疫情应对措施而调整的幼儿发展干预措施。该评价有助于改善克罗地亚的政策环境，重点是在所有应急计划中支持最脆弱的家庭和儿童，并确保政府承诺推广这一做法。该评价还为克罗地亚制定新的儿基会国家方案文件提供了宝贵的参考。

60. 根据其继续建设工作人员评价能力的承诺，评价办公室在 2021 年为工作人员和合作伙伴开展的混合评价学习课程方面取得了重大进展，该课程包括通过互动平台促进中级在线学习，随后由新加坡国立大学提供为期两周的面对面高级评价培训。2021 年举办了一门试点课程和两批在线协助课程，参与者来自四个地区（东亚和太平洋、东部和南部非洲、欧洲和中亚以及南亚）。完成课程的 53 名学员获得了结业证书。目前正在招聘将于 2022 年教授法语和西班牙语课程的教员。新加坡国立大学开发了面对面培训课程，将于 2022 年为大约 100 名员工和合作伙伴提供培训。此外，两门自定进度的在线课程正在进行测试，将在今年下半年推出。

五. 结论和未来的方向

61. 2021 年，2019 冠状病毒病疫情形势依然严峻，带来重重挑战。评价职能适应了这些前所未有的挑战，找到了新的工作方式、新的评价产品来帮助为正在进行的应对提供信息，并建立了新的评价伙伴关系来处理与这一前所未有的突发公共卫生事件相关的具体评价领域。
62. 尽管有诸多限制，但评价职能继续在既定的业绩指标方面表现良好：评价活动数量、评价覆盖面、评价支出均有所增加。评价质量总体保持较高水平，管理层回应要求几乎得到普遍遵守，这表明就评价建议采取有意义行动的一些必要（即使不是完全充分）的先决条件似乎已经具备。正如本报告通篇展示的评价结果所表明的那样，评价在许多重要方面有益于儿基会的工作。
63. 这场疫情既带来了挑战，也创造了机遇。远程数据收集成为克服访问限制的必要变通方法，而实时评价、实时评估、综合和审查均表明，可以相对轻松的方式快速产生有用的分析结果。即使其中一些变通办法并不完全是可持续的，但它们说明可以简化评价，而不会严重损害评价职能工作的独立评价性质。
64. 与此同时，评价职能面临持续挑战时的适应过程也为当前的四年期提供了宝贵的经验教训。《2022-2025 年全球评价计划》总结了这些经验教训，并初步展示了影响 2021 年评价职能战略方向的一些主要趋势。这包括高度重视结果和影响级别的证据生成，还包括扩大供管理人员使用的成套评价产品，以帮助更及时、更迅速地地为儿基会的工作提供信息。
65. 全球评价计划还为大力加强评价职能内部以及与合作伙伴的协调和整合指明了方向。这包括提升各评价组合的一致性，以采取更全面的整体儿童方法；更一致和有意义地纳入性别平等、残疾问题和其他脆弱性来源，以确保没有任何儿童掉队；加强评价职能三个层面的一致性；加强与监测、研究和审计等不同但互补的职能的协调和合作；以及以加强与可持续发展目标有关的联合评价和机构间评价的形式，加强与传统和非传统评价伙伴的合作。
66. 本年度报告中介绍的业绩数据提供了额外的知识来源，从中可以汲取经验教训，为今后的工作指明方向。这项分析的一些方面强化了《2022-2025 年全球评价计划》中传达的经验教训和战略方向，而其他方面则提供了在制定四年期计划时尚无的新数据和信息，从而为评价职能制定战略方向提供了进一步的参考。
67. 最突出的经验教训之一是，在使用传统评价和不太传统的评价活动（如可评价性评估、综合、实时评估和实时评价）方面仍然存在不平衡现象。虽然评价职能扩大了其评价工作的范围，以应对 2019 冠状病毒病疫情带来的独特需求和挑战，但这些非传统工作仍然只占 2021 年所开展工作的一小部分。如果评价职能要在“行动十年”期间最大限度地符合的目的，就有必要在传统评价和不太传统的模式之间重新平衡其提供的服务。一方面，有必要更加重视结果和影响级别的评价，执行局在其第 2022/5 号决定中确认了这一点。另一方面，评价职能还必须纳入更多现有的早期、

快速和轻便的评价方法，如可评价性评估、实时评价和实时评估、审查和证据综合。前一种模式对于从问责角度衡量为儿童取得的实际成果至关重要，而后一种模式对于帮助儿基会取得这些实际成果至关重要。在所有这些努力中，将更加注重联合评价和机构间评价工作，并以可持续发展目标和《2030 年可持续发展议程》为基础，更加坚定地开展评价。

68. 第二个经验教训是，不能把稳定的评价质量看作理所当然；相反，如果要保持高质量，甚至提高质量，就需要保持警惕并制定创造性解决方案。虽然 2021 年的质量总体上仍然相对较高，但与前几年相比，在一些关键领域的质量有所下降或持平。虽然与 2019 冠状病毒病疫情有关的适应过程可解释其中一些不足出现的原因，但并不一定能解释所有不足之处。2022 年，将进行更细致的分析，找出质量变化的确切原因。此外，还将对全球评价报告监督系统平台进行审查，以确保该平台是最新的，在应用上是一致的，并对其进行优化，以反映目前正在进行的各类评价工作可能需要的不同质量保证方法。

69. 质量本身并不能保证评价将实现其最终目标：影响儿基会及其合作伙伴的行动、决定和总体战略方向。虽然令人鼓舞的是，2021 年对管理层回应要求的遵守率仍然很高，但这并不意味着管理层的回应同样有力，也不意味着管理层的回应正在得到积极落实和持续监测。因此，从这一分析中得出的第三个经验教训是，必须更加重视组织针对评价采取的后续行动。2022 年，将进行一项关键分析，以确定任何可能需要改进的“供应侧”领域——即评价本身的各个方面，如建议的制定、全球评价报告监督系统对建议质量的评级方法、评价管理、管理层回应流程或其他因素。与此同时，这一重要分析将探索最终用户的潜在改进领域，如关于管理层回应的制定和实施的平台和治理结构。最后，将委托进行一项综合，以确定评价建议反复针对但尚未采取有意义行动的儿基会业务领域。通过这些方式，评价职能将与儿基会的其他部门一道，按照《2022-2025 年战略计划》的规定，更广泛地转向更有力的结果和影响级评价。

70. 要完成上述重大战略转变，同时保持迄今取得的成果，将需要资源。在这种情况下，本分析显示，为评价职能提供充足的资源——这是执行局在其第 2018/10 号决定中提出并在其第 2022/5 号决定中强调的一项指示——是当前四年期需要持续关注的一个领域。虽然整个评价职能必须有资源保障，但特别令人关切的是分散评价职能是否拥有充足和可预测的资源——因为大多数评价是在这一级进行的，2021 年在这一级取得的进展最大，而且需要在这一级衡量对可持续发展目标的影响。近年来，分散评价职能的非工作人员费用（以及部分工作人员费用）由评价集合基金支付。2022-2025 年综合预算将该基金调整到 2150 万美元的拟议水平。展望未来，有必要对该基金进行持续监测，以确保这一数额足以维持迄今取得的重大成果，并为本四年期未来开展雄心勃勃的工作做好准备。

71. 为了使评价职能继续发展并始终符合目的，有必要超越内部分析模式，如本年度报告或《2022-2025 年全球评价计划》。为此，将在 2022 年进一步分析本报告的结论，以期揭示出需要解决的持续挑战的根本原因，以及需要在整个评价职能中更

广泛地推广的成功案例的根本原因。此外，2022 年，评价办公室将委托进行一次独立的外部同行审查，从独立、公正和客观的角度进一步探讨上述问题以及整个组织评价职能的总体成熟度。这一独立的同行审查将为评价职能提供进一步指导，而且重要的是，将为修订儿基会评价政策提供重要参考；该政策将于 2023 年提交执行局。

72. 通过采取上述众多措施，可以预见，评价职能将继续发展、改进并应对我们这个时代前所未有的挑战。通过将评价视角转向其 2022 年及以后的工作，以确保评价工作尽可能及时、相关、高效、有效和有影响力，评价职能将致力于以更好的方式支持组织学习和问责，使儿基会、其合作伙伴和执行局掌握在“行动十年”中实现世界各地每名儿童的权利所需的评价证据。

六. 决定草案

执行局

1. 注意到关于儿基会评价职能的 2021 年年度报告（[E/ICEF/2022/17](#)）及其管理层的回应（[E/ICEF/2022/18](#)）；
2. 还注意到对儿基会作为群组牵头/共同牵头机构的作用的评价、其摘要（[E/ICEF/2022/19](#)）及其管理层的回应（[E/ICEF/2022/20](#)）。