



经济及社会理事会

Distr.: General
15 May 2026
Chinese
Original: English

供决定

联合国儿童基金会

执行局

2026 年年会

2026 年 6 月 16-19 日

临时议程* 项目 11

2025 年儿基会评价职能年度报告

摘要

本报告概述了 2025 年儿基会评价职能的情况，提供了背景概况，重点介绍了主要成果与挑战，最后就未来数年进一步强化该职能及其扶持环境的优先事项进行了总结思考。

供执行局审议的决定草案要点载于第五节。

* E/ICEF/2026/9.

说明：本文件全文由儿基会负责处理。

26-06992 (C) 150526 180526



请回收



一. 概述

1. 2025 年是一个关键转折点：2022 - 2025 四年期即将结束，是儿基会《订正评价政策》（[E/ICEF/2023/27](#) 和 [E/ICEF/2023/27/Corr.1](#)）实施的第二个完整年度，同时也是儿基会面临前所未有的压力时期。未来重点倡议及其相关的方案预算审查带来了重大机构重组与财政紧缩，评价职能在压力下既要持续开展工作，又要适应变革，同时还要维护作为其公信力基础的独立性。
2. 在此背景下，评价职能核心指标的执行情况总体保持稳定——这一成果并非必然，也来之不易。2025 年评价提交数量较 2024 年有所增加，反映出在方案周期要求下，评价职能着力扩大评价覆盖面；但数量仍低于本四年期前期峰值，这与评价职能始终坚持质量与战略相关性优先于数量的原则相一致。评价质量保持稳定，2025 年平均评分略有提升。管理层回应的及时性逐步改善，但仍有相当一部分回应超过 60 天时限。这一长期存在的挑战，反映了评价职能的分散化特点，以及仅靠评价工作所能推动的成效存在局限。评价支出绝对值保持稳定，但始终低于 1% 的基准，表明扶持环境资金紧张且分布不均。
3. 尽管对评价工作在组织决策中总体贡献的量化衡量仍是一项持续挑战，但有关影响的定性证据不断增加。《2025 年有影响力评价研究》系统地分析了评价发挥作用的方式与原因，识别出评价在机构改革、方案设计、国家政策和人道主义应对等领域产生影响的常见模式。各类评价、综合报告和咨询说明为重要机构决策提供了依据，包括儿基会《2026 - 2029 年战略规划》的编制，以及儿基会对未来重点倡议进程的统筹推进。这些事例表明，评价职能的影响力远非绩效指标所能全面体现。
4. 三项跨领域战略举措——影响评价、国家评价能力建设和全球评价综合分析——共同推动了评价职能最具雄心的目标：助力儿基会及其合作伙伴侧重于有效举措，大规模为儿童实现影响级的成果。这些举措通过就有效干预措施生成严谨证据、建设各国政府用以对成果进行自我问责的国家体系，并将积累的知识转化为可供高层政策论坛使用的决策参考，使评价工作的覆盖范围与影响力远超儿基会单独所能实现的水平。
5. 在内部，该职能在成熟评价体系的体系构建与关系营造两方面均取得显著进展。首部《儿基会评价手册》编制完成，基于风险的规划工具实现制度化，并制定了一套方法指南。全球评价领导小组（GELT）正式成型，成为在评价职能三个层级间保障一致性与集体管理的平台。《2026 - 2029 年全球评价计划》——包含 23 项评价（上一四年期为 44 项），已完成全额成本核算，并附有评价办公室首个含成本核算的年度工作计划——于 2026 年 2 月获执行局核准，体现了基于风险的审慎优先排序，显著提升了评价规划与资源配置的透明度。
6. 评价职能在进入新四年期时，虽面临资源缩减和机构格局调整，但其基础较以往任何时期都更为坚实。填补《2026 - 2029 年全球评价计划》中约 790 万美元的资金缺口、维护保障独立性的制度安排，以及在质量、应用与战略影响力领域保持既有势头，将是未来面临的核心挑战。即将由经济合作与发展组织发展援助委员会（发援会）开展的儿基会评价工作独立外部评估（将于 2027 年提交执行局），将为

该职能的运行状况与成熟度提供及时且完全独立的视角，并为下一个四年期规划方向提供重要契机。

二. 评价职能的主要发展

A. 背景和 2025 年的优先事项

2022 - 2025 四年期最后一年的不确定性与变革

7. 作为 2022 - 2025 四年期的收官之年，2025 年是一个重要节点：《2022 - 2025 年全球评价计划》执行进入收尾阶段，也是儿基会《订正评价政策》（2023 年）实施头两年的结束。本四年期面临诸多重大挑战，此期间提交执行局的评价职能年度报告均有阐述；《订正评价政策》也作出了重要战略调整，旨在重新定位评价职能，以应对这些挑战，并在未来数年优化评价职能在支持组织学习与问责方面的价值。本年度的工作重点体现为持续推进政策落实，监测相关核心绩效指标，同时为制定儿基会《2026 - 2029 年战略计划》作出贡献。

8. 这种延续四年期既有工作势头的思路，很快因儿基会日益严峻的运行环境而难以维系。随后启动的未来重点倡议带来了重大机构改革与财务不确定性。对评价职能而言，该倡议导致 2026 - 2029 年评价办公室资源额度大幅削减，降幅超出全组织 25% 的基准水平，且办公室迁至罗马。评价集合基金可支配资源亦有所缩减，治理安排相应调整，包括评价管理层在评估资金需求方面的职责。对人员配置的影响同样显著。评价职位被裁撤、调岗，或由机构资源经费转为其他资源经费供养，导致人员配置低于政策规定的最低编制标准。上述情况对评价职能整体的覆盖范围、工作能力与资源配置构成压力，资源额度缩减意味着 2025 年及下一个四年期计划开展的部分评价工作无法获得足额资金，或需推迟实施。

9. 因此，未来重点倡议从约束与明确方向两方面，定义了 2025 年的运行环境。它要求评价职能在承受大幅削减的同时，维持工作开展、保障独立性，并为处于前所未有的变革中的组织提供支持——多重需求叠加，使得优先排序与适应性成为本年度的核心要求。

在不断变化的运行环境中重新调整工作重点

10. 面对组织层面普遍存在的不确定性与变革，评价职能在 2025 年转向四项并行工作重点：（1）调整评价职能自身结构与运行方式以适应新现实，同时陪伴儿基会度过前所未有的机构压力期，并维护其独立性；（2）继续履行核心承诺，产出及时、相关、可信且实用的评价成果，为儿基会及其合作伙伴创造价值；（3）巩固既有工作势头，通过内部强化评价实践、对外构建面向儿童的全球证据体系，从而带来改变；（4）防止评价职能既定关键绩效指标出现倒退。在资源紧张与组织不确定性的双重压力下同步推进这些重点工作，构成了评价职能全年工作的核心特征。

11. 关于第一项重点，评价职能不仅面临规模缩减，还通过加快落实优先排序与增效措施以适应日益收紧的财务约束，其中部分措施在未来重点倡议启动前已在规划或推进中。这些措施包括选择性采用混合实施模式，即将内部评价专业力量与个体

咨询相结合，替代通常成本更高的机构咨询服务。同时还对业务流程进行调整，在提升时效性、聚焦实质内容与方法的同时降低成本，并对精简后的评价办公室人员配置进行重组，以最大限度提升团队内部的一致性与凝聚力。富有创造性与创新性的方法和技术解决方案（下文第二 C 节将详细阐述）进一步表明，在部分情况下，成本效益可与评价的时效性、相关性、质量与公信力形成相互促进。

12. 对评价工作进行优先排序变得更为关键，以确保日益有限的资源投向具有最高战略重要性的领域，进而在支持学习与问责方面发挥最大潜在价值。这一转变体现在《2026 - 2029 年全球评价计划》中。该计划经编制形成草案，提交执行局 2025 年第二届常会审议，并于 2026 年第一届常会核准。该计划采用基于风险与机遇的方法确定优先事项，并根据关键程度对评价工作实行结构化分级。最终形成的 23 项评价组合——较上一四年期规划的 44 项大幅削减——也体现出对机构吸收能力限制的更高敏感度。应执行局要求，首次提交了 2026 年含成本核算的年度工作计划，显著提升了评价资源配置以及资源与核心评价优先事项之间缺口的透明度。

13. 评价职能在适应不断变化的运行环境与资源缩减状况的同时，积极通过循证分析支持儿基会统筹推进未来重点倡议。在改革进程初期，评价办公室即编制两份咨询说明（其中一份与内部审计和调查办公室联合编制），依托 2017 年以来的评价、审计与战略审查成果，梳理出对决策具有参考意义的共性挑战与经验教训。评价职能相关人员还参与与区域管理团队的磋商，旨在推动落实政策承诺、保障区域评价能力，并持续巩固支撑评价工作取得成效的良好合作关系。

14. 评价职能的独立性取决于充足且可预期的资源、不受报复地开展工作的自由，以及参与直接影响其自身的决策与行动的权利。因此，在面临前所未有的机构变革与资金约束的情况下，保障独立性的这些核心特征，成为 2025 年的另一项关键优先事项。执行局第 2025/14 号决定重申了评价职能的独立性，并要求自 2026 年起编制列明成本的年度评价工作计划。此外，2025 年下半年已采取初步措施，为开展儿基会评价工作独立外部评估做好准备。该项评估系执行局在第 2023/12 号和第 2023/20 号决定中提出要求，将为儿基会管理层和执行局提供独立、公正的视角，审视评价职能及其扶持环境的运行状况与成熟度，包括与独立性相关的安排。

15. 总体而言，围绕第一项战略重点开展的这些工作，旨在夯实评价职能的机构基础，使其在未来的儿基会体系中尽可能高效有效地运行。后续各节记录了其他三项优先事项的进展情况：持续为儿童和儿基会创造价值的评​​价工作与证据成果（第二节 B）；在资源受限条件下拓展职能覆盖面的创新做法与方法学改进（第二节 C）；以及通过各项战略举措，推动构建更广泛的儿童问题循证问责体系（第二节 D）。

B. 评价对组织成果的突出贡献

16. 致力于理解并展现评价工作如何发挥实效，是本四年期该职能的一大特征。2025 年，这一工作方向取得了一项重要里程碑成果——《有影响力评价研究》。¹该

¹ 儿基会，《2025 年有影响力评价研究：最终报告》，儿基会，纽约，2025 年，可查阅 <https://www.unicef.org/evaluation/media/7386/file/The%202025%20Study%20of%20Influential%20Evaluations.pdf.pdf>。

研究剖析了评价结论如何转化为儿基会内部及外部政策环境中的具体决策。研究确定了评价工作为儿童改善成果的四条常见路径：为机构改革与战略定位提供参考；指导方案重新设计与规模化决策；影响国家政策与体系建设；以及在脆弱和高风险环境中强化问责与业务学习。下文案例分别体现了上述路径。

为机构改革与战略定位提供参考

17. 《儿基会人力资源管理评价》显示，在数字化人力资源系统与员工福祉举措的支持下，招聘、性别均等和员工队伍多元化方面取得进展，同时也指出了战略人力规划、职业发展和组织文化方面存在的不足。文化“特征印记”分析表明工作与儿基会核心价值观相符，但也凸显了风险规避与层级化决策相关问题，为人力规划、招聘和绩效管理改革提供了依据。《儿基会人力资源保护工作基线评估》指出，人力资源全流程中保护工作融合存在缺口，并制定了将保护工作纳入招聘、培训和绩效体系的路线图。《儿基会创新工作评价》则发现，治理、评价组合管理和伙伴关系方面取得进展，但同时存在监管分散、规模化创新成果证据不足等问题，为完善创新治理与资金跟踪提供了参考。评价办公室与人员及文化司、创新办公室之间密切协作，使得三项评价成果均得到充分采纳。

18. 聚焦部门的评价工作也为机构定位提供了支撑。《儿基会初级卫生保健工作评价》发现，超过 80% 的国家办事处参与初级卫生保健相关干预工作，在服务提供结合政策、治理与资金等上游工作的领域成效最为突出。这些结论为加强跨部门融合、明确儿基会初级卫生保健整体工作思路提供了参考。在区域层面，《欧洲及中亚区域免疫方案编制多国评价》指出了可通过加强数字化建设、外展服务与区域协作来维持免疫接种覆盖率。

指导方案设计与规模化决策

19. 《加速普及卫生设施与全民供水方案评价》发现，该方案在八个国家大幅提升了民众获得改良饮用水、环境卫生和个人卫生服务（水卫）的机会，覆盖学校、卫生机构和社区，同时也指出了可持续性方面的挑战。管理层回应完善了监测框架、跨部门协作与可持续性核查机制，为方案下一阶段的设计与筹资提供了依据。

20. 同样，《2018 - 2023 年儿基会儿童早期发展与学前教育工作评价》（[E/ICEF/2026/3](#)）分析了儿基会在加强儿童早期体系建设方面的贡献，证实其在支持各国政府制定多部门政策与计划中发挥了重要作用。该评价还指出，儿基会助力提升了各国政府为儿童早期方案筹措资金的能力，但同时也提到，公共筹资战略尚未转化为对儿童早期服务的大规模投资。如 2026 年第一届常会所述，该评价的建议正用于指导战略计划周期内相关指南、指标与成果链的更新。

21. 国家层面形成的证据也为方案调整提供了参考。《南苏丹可持续水卫韧性方案评价》指出了可持续性方面的差距，同时证实了在脆弱环境中扩大安全饮用水和卫生服务方面取得的进展，为方案设计调整提供了依据，并加强了与负责维护供水系统的国家及社区行为体之间的协作。《尼日利亚 - 儿基会水卫国家方案合作评价

《2018 - 2022》》记录了在扩大水卫服务覆盖方面取得的进展，同时指出了可加强与政府体系协调、探索创新筹资方式的方向，为下一阶段方案规划提供了依据。

22. 《尼日利亚东北部儿童韧性综合教育方案编制评价》与《教育不能等待多年期韧性方案二期评价》则为受冲突影响地区以韧性为导向的教育方案编制提供了证据，为改进教师培训、社区参与以及将教育干预与更广泛的韧性举措相结合提供了参考。

影响国家政策与体系建设

23. 《危地马拉慢性营养不良综合战略评价》为国家政策对话提供了证据，并加强了儿童慢性营养不良问题的协调机制。《马来西亚国家营养行动计划评价》推动了部门战略与筹资安排的协调统一。《厄瓜多尔儿童慢性营养不良预防与减轻区域战略评价》为在国家以下层级加强国家营养政策实施提供了证据。在市级层面，《阿根廷市镇联合关爱儿童与青少年方案评价》展示了儿基会如何支持加强地方层面以儿童为中心的政策和服务规划与协调，为后续国家方案的重新设计提供了依据。

在脆弱环境中强化问责与业务学习

24. 《儿基会在苏丹的三级应急响应评价》凸显了儿基会在维持人道主义行动中的作用，同时指出了在备灾、早期预警和人员配置缺口方面的不足。《儿基会在海地的三级应急响应评价》强调需加强备灾工作并实施更具统筹性的人道主义战略，为儿童提供基本服务。《儿基会人道主义行动应急准备工作全球评价》也指出在加强备灾体系方面取得进展，同时凸显了持续存在的缺口，如备灾职能资金不足、风险分析不一致等。

25. 在分散层面，《儿基会对乍得东部难民危机的应对评价》为调整人道主义协调工作、支持难民儿童与收容社区提供了依据；而《索马里中南部联合韧性方案评价》则表明，韧性方案编制相关证据可加强保护与社会保护干预措施的规划，支持脆弱环境中儿童的粮食安全、教育与保护成果。

26. 《有影响力评价研究》梳理了能够增强评价影响力的有利条件。在改革与财政紧缩背景下，这些条件使评价证据能够指导优先排序与决策。本节所列案例正体现了这一点：及时、面向决策的评价在多种环境下强化了惠及儿童的方案、体系与政策。

C. 内部带来改变： 强化评价职能的“体系构建”与“关系营造”

27. 在资源缩减且有证据表明仍有领域亟待加强的背景下，随着未来重点倡议的启动，本四年期前期为提升评价工作的相关性、公信力、实用性与独立性而启动的各项举措变得更为紧迫。《订正评价政策》（2023 年）将这些措施归类为健全评价职能的“体系构建”与“关系营造”。2025 年两方面均取得显著进展。

28. “体系构建”指将政策转化为更具体的指南和程序等实质性成果。一项里程碑式成果是编制完成首部《儿基会评价手册》，该手册整合了整个评价职能的方法标准与业务指南。基于风险的规划方法已实现制度化，最终形成的《2026 - 2029 年全球

评价计划》比以往任何时候都更具数据驱动与循证特征，该计划通过专门平台识别高关键性评价主题。含成本核算评价计划的指南得到完善；编制了关于性别平等主流化与青年参与的指南说明；制定了一系列高级方法学指南；更新了质量保证工具；推出了影响评价评估模板；评价政策实施程序已进入机构审批阶段。儿基会评价工作独立外部评估的筹备工作也已启动。其中多项标杆性资源的完善工作将在 2026 年持续推进。

29. “关系营造”指评价职能内部，以及评价职能与儿基会其他部门伙伴之间的协作关系，这是将评价政策所阐明的原则内化所必需的。对该职能而言，这一环节尤为重要：评价职能直至 2019 年才正式组建为全球性职能（即集中与分散相结合），2020 - 2021 年出现管理层空缺，2022 年从新冠疫情期间的远程办公模式逐步恢复，且在分散层面采用矩阵式管理架构。通过全球评价领导小组（GELT）的正式组建，职能内部的关系营造得到加强。该小组由评价办公室管理层及区域评价顾问组成，旨在强化规划、政策实施与战略定位的一致性，包括全球评价计划的制定。与区域评价顾问的直接沟通和专项协作也进一步深化，加强了职能的协同性、问责性与集体管理，以及分散层面的独立性标准。

30. 除在推进上述政策承诺方面取得进展外，该职能还采取了多项其他措施，以加强自身在评价工作的规划、实施和沟通方面的运作。为制定《2026 - 2029 年全球评价计划》而采用的基于风险、突出重点、与战略保持一致的方法便是此类进展之一，该方法需要搭建一个定制化的数据驱动平台，用于识别高重要性评价主题。

31. 在开展评价工作时，该职能寻求具有创造性和创新性的解决方案，以应对复杂的评价挑战。儿基会人力资源管理评价运用了新方法，借助案例研究和数据可视化工具分析人力趋势和组织文化。同样，创新评价通过借鉴多个区域的案例研究，考察治理安排、组织体系和伙伴关系方面的变革如何影响创新解决方案的推广。这些分析有助于找出影响绩效的制度性障碍和有利因素。

32. 在准入和安全限制制约传统实地工作的情况下，适应性研究方法尤为有效。例如，在对巴勒斯坦国三级应急响应开展评价期间，在加沙开展的早期启动工作团实现了快速反馈，为业务调整提供了支持。在斐济、黎巴嫩、利比里亚和土耳其等其他脆弱环境中，尽管业务受到干扰，仍通过分阶段和适应性方法保持了分析的连贯性。在多个区域，针对应急响应型社会保护的评价结合了远程数据收集（包括手机调查）和方案行政管理数据，在无法开展实地走访时及时形成证据。

33. 人工智能助力儿基会快速扩大整合和综合大量评价证据的工作规模。人工智能和混合方法的更广泛应用也要求加强保障措施，包括人工监督、多方验证和针对性质量保证。评价办公室更新了《评价证据差距图谱（2018 - 2025）》，该图谱系统整合儿基会各方案的评价证据并明确知识缺口，为未来评价工作提供指导并减少重复劳动。阿富汗、卢旺达和乌拉圭的办事处试点运用人工智能辅助证据挖掘和微观综合技术，从既有评价中提取经验；肯尼亚、津巴布韦、加勒比地区、洪都拉斯和萨尔瓦多开展的评价组合层面综合分析则为方案审议和国家方案评价提供了依据。

34. 这些进展还体现为该职能为更广泛的联合国评价界在研究方法上作出的贡献。通过文本挖掘技术，该职能梳理了 600 余份联合国评价报告所采用的研究方法。这项工作形成了全系统范围的评价方法证据基础，为联合国评价界提供了一项公共产品，并支持各方持续努力提升全系统方法的透明度与质量。伙伴关系也推动了方法创新，包括在东部和南部非洲以及南亚开展具备气候适应能力的饮用水、环境卫生和个人卫生评价，通过与环境研究机构合作，运用地理空间分析、卫星影像和气候风险建模等技术。

35. 其他创新性进展则聚焦于评价中同样重要的定性方面：在评价中纳入儿童声音与儿童能动性，以及优化评价相关沟通工作。在伯利兹、不丹、马尔代夫、尼泊尔和索马里开展的评价，进一步将儿童和社区视角融入评价设计、数据收集与验证环节。相关工作还延伸至知识转化领域：在巴巴多斯、马里和委内瑞拉玻利瓦尔共和国，儿基会搭建了交互式监测、评价与学习仪表盘，提升评价结果的可及性，支持其在方案审议和伙伴对话中的应用。

D. 外部带来改变：推动形成更广泛的儿童工作循证问责运动

36. 大规模、可持续的影响层面成果需要集体行动，而非单一组织独自能够实现。在 2022-2025 四年期内，评价职能着力发挥儿基会独特的比较优势——其法定职责、全球布局以及与各国政府的可信伙伴关系，助力推动形成更广泛的儿童工作循证问责运动。三大互补支柱推动了这一议程：影响评价、国家评价能力建设以及全球证据综合。2025 年，各项工作均持续推进。

影响评价

37. 影响评价是一类严谨性极高的评价形式，有助于儿基会及其合作伙伴明确哪些干预措施能够为儿童带来可衡量的成果，从而值得继续实施、调整优化或推广扩大。因此，它是一项重要的战略工具，可将有限资源聚焦使用，实现投资回报最大化，推动规模化影响并降低风险。本年度内，儿基会完成了五项影响评价，尽管面临财政限制，仍保持了稳定的分析产出，同时为国家政策对话增添了价值。在乌干达开展的《西尼罗次区域卫生系统强化干预措施影响与经济评价》，形成了在脆弱环境下现金转移支付与系统强化干预综合效果的证据，为资源配置决策提供参考。在马拉维开展的《社会现金转移支付方案影响评价（第一次中期）》，提出了阶段性成果，为方案调整提供指导。在秘鲁开展的《联合方案高中转移支付（TAS）影响评价》，评估了其对于中等教育成果的作用，为国家政策对话提供支撑。

38. 为鼓励加大对影响评价的投入，儿基会设立了影响催化基金，用于支持重点方案领域的影响评价工作。该基金通过为可行性研究、基线数据收集和方法筹备提供资金，助力国家办事处形成证据，为政策制定和推广决策提供依据。该基金支持了童婚、社会规范及适应性社会保护等专题方向，已在布基纳法索、加纳、尼日利亚、索马里和土耳其完成基线数据收集。巴基斯坦、乌干达和也门的影响评价筹备工作也在推进，同时在心理健康与社会心理支持、气候适应性饮用水环境卫生和个人卫生服务、针对移民和流离失所者的展望伙伴关系等领域开展多国可行性评估，在全面启动影响评价设计前评估实施条件与评价可行性。来自 14 个国家的技术援助请求表明，尽管资源有限，对严谨证据的需求依然持续存在。

39. 影响评价工作也深化了联合国系统内的协作。作为联合国评价小组影响评价工作组联合主席，儿基会牵头制定了适用于联合国各机构的影响评价统一定义和共用方法指南。这一全系统公共产品将进一步提升面向儿童的影响证据的严谨性与可比性。

国家评价能力发展

40. 影响评价旨在形成有效干预措施的相关证据，而国家评价能力建设则旨在提升各国政府及其他合作伙伴自主需求、生成并运用评价证据的制度能力。这一区分意义重大：儿童相关工作的可持续问责最终依赖国家体系，而非外部体系。儿基会依托自身法定职责、全球覆盖范围及与各国政府的可信合作关系，推动将评价工作纳入国家规划与预算流程，从单一培训转向系统性变革。

41. 在全球层面，儿基会通过评价领导者高级研修课程（ExCEL）持续强化国家评价体系，面向副部级官员、议员、总统办公室人员及其他政策制定者等高级别政策负责人培养评价能力。2025 年新增两期课程，累计共举办 12 期，为 83 个国家培养 408 名获证学员。在多个国家，课程学员后续参与推动了本国主导的评价工作、国家评价战略制定，以及评价结果在政府决策中的更广泛应用。一项重要里程碑是巴西首次在全国范围内开设该课程，与巴西国家公共管理学院、新加坡国立大学合作以葡萄牙语授课。该方案还在刚果民主共和国开展次区域培训，汇聚了 39 名高级官员。日益增长的需求与重复参训情况表明，该课程正推动各国政府形成更广泛的儿童工作循证问责运动。

42. 评价体系强化中级指导方案（IMPrESS）为政府官员及国家评价从业人员提供支持，助力完善相关制度与实践，推动评价结果应用于公共决策。该方案将系统化培训与实践应用相结合，帮助学员将评价工作融入政府规划与问责体系。自启动以来，该方案已在 101 个国家开展 41 期培训，惠及 1695 名学员；仅 2025 年就举办 13 期，惠及 694 人。相关筹备工作推动了评价工作的国家制度化进程，包括贝宁国家行政学院的方案本土化调整，计划于 2026 年启动。

43. 这些方案辅以国家层面的参与，将新掌握的技能与形成的网络转化为制度改革成果。在蒙古，儿基会推动与发展伙伴及国家评价协会协作，强化评价工作在政府各机构中的制度化建设。在泰国，儿基会提供技术援助，助力制定国家教育部门评价框架。儿基会还在孟加拉国、埃及、摩洛哥和南非支持国家评价政策制定与体系强化工作。在贝宁，儿基会支持举办第七届国家评价日，加强政府、民间社会与评价网络间的对话交流。

全球评价综合

44. 全球评价综合通过整合经验教训，日益成为对单项评价的补充，用以识别工作中反复出现的模式、共性挑战以及相关举措切实为儿童带来成效的领域。综合评价汇集多项评价结果，为决策者提供更宏观的视角，把握共性主题、现存差距与改进方向。阿富汗、肯尼亚、黎巴嫩、卢旺达、乌拉圭和津巴布韦的国家办事处运用人工智能辅助技术，对既有评价开展快速证据梳理，整合各方案领域经验教训。这些工作将既有成果转化为契合规划与预算周期、可直接用于决策的参考，同时减少对新增独立研究的依赖。

45. 儿基会与联合国开发计划署联合牵头全球可持续发展目标综合联盟，召集联合国系统及多边开发银行共 45 个独立评价机构参与。在该平台下，围绕可持续发展目标 1 至 5 社会救助干预措施开展的“人员支柱”综合评价，整合了 155 项同行评审研究与 94 项联合国评价的证据，并在 2025 年 11 月第二届世界社会发展首脑峰会等高级别论坛上发布了具有政策参考价值的结论。这三大支柱——形成严谨的有效干预证据、建设可自主生成与运用证据的国家体系、为全球政策层面整合既有知识——共同构成了一套连贯且具有前瞻性的策略，推动评价工作规模化服务于儿童权益。

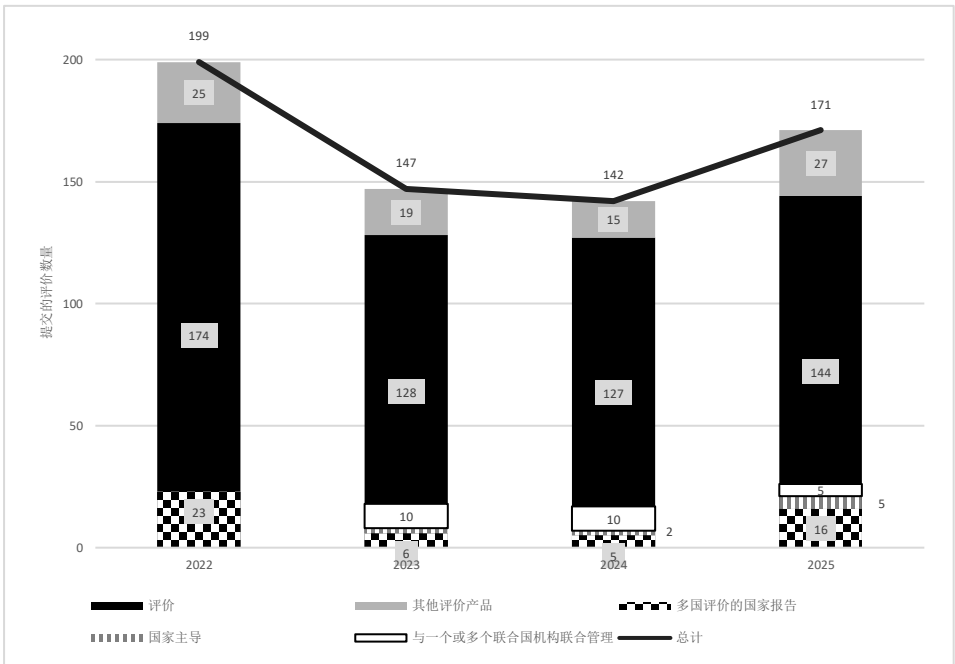
三. 业绩概览：评价职能的持续进展和主要经验教训

46. 在整个四年期内，儿基会对照为评价职能设定的关键绩效指标，对进展情况进行监测和报告。总体情况呈现出在压力下保持总体稳定的态势：往年取得的进展以及 2025 年面临的各类冲击，均未显著改变该职能的运行轨迹。在这一机构变革前所未有的年份里，能够维持这种连续性本身就是一项重要成果。相关指标从报送情况、质量、管理回应和支出等方面均印证了这一点。

A. 提交的评价数量和地理覆盖范围

47. 正如以往报告中所强调的，其目标并非单纯以增加评价数量为目的，而是确保各项评价具备战略相关性、可信度和实用性，同时使各国家办事处达到最低覆盖要求。在连续两年受控下降后，2025 年评价提交数量有所回升。2025 年共提交评价及其他评价类成果 171 项，较 2024 年的 142 项有所上升。

图 1
2022–2025 年提交的评价数量



资料来源：证据信息系统整合（EISI）数据库。

48. 2022年提交数量处于异常高位，主要是此前扩大覆盖范围的工作及新冠疫情相关评价大幅增加所致。2023年和2024年数量下降，反映出机构为明确评价界定标准采取了审慎措施，同时战略上从重数量转向重质量与相关性。这也反映出部分地区资源紧张、评价领域投入不足的情况。

49. 2025年，机构重组、人员更替以及预算收紧进一步对评价工作形成下行压力。因此，本次提交数量回升是结构性和管理性对冲因素共同作用的结果。标准方案周期要求发挥了重要作用：对方案设计和新一轮周期规划具有关键意义的国家方案评价数量翻倍以上，从2024年的9项增至2025年的20项。多国评价数量同样有所增加，体现出对跨国方案成效开展审视、提炼跨区域适用经验的工作力度。

50. 此前为弥补覆盖缺口开展的工作也推动了本次数量上升。2024年有39个国家办事处未提交任何评价材料，这些办事处在2025年均提交了至少一项评价，说明评价工作覆盖到了此前参与度较低的办事处，相关工作取得进展。同时，评价可行性评估数量有所增加，表明更多采用前期诊断方式统筹评价工作安排，确保在委托开展大型评价前，方案已具备相应条件。

表 1
2024–2025 年提交的评价和其他评价产品的数量，按地点分列

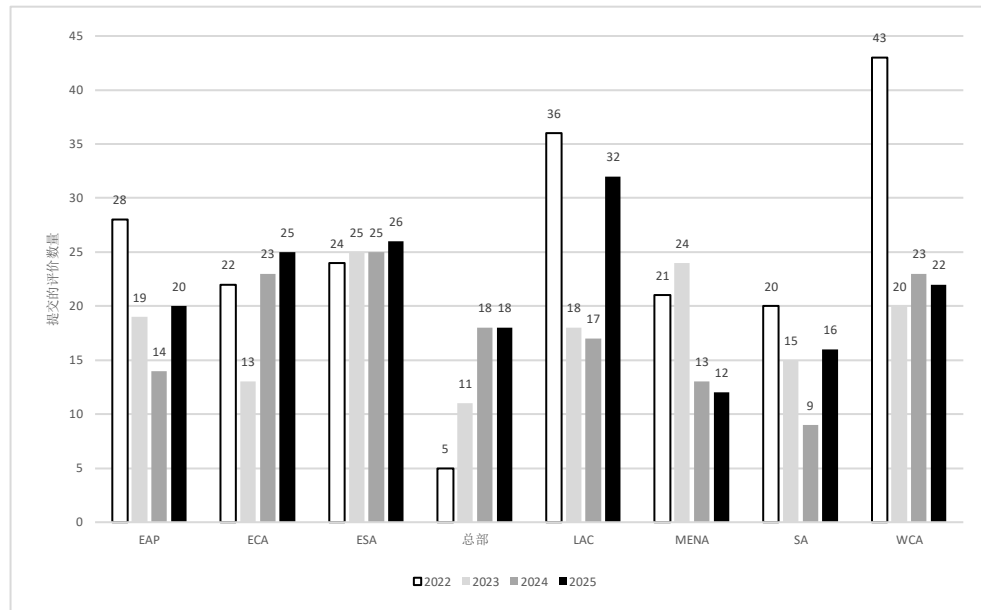
地点	提交的产品																			
	评价		其中包含的国家主导的评价数量	其中包含的联合国可持续发展合作框架/联合国全系统的评价数量	其中包含的与一个或多个联合国机构联合管理的评价数量	其中包含的多国评价国家报告的数量	可评价性评估		实时评估		综合		影响可行性		基线评估		审查		总计	
	2024	2025	2025	2025	2025	2025	2024	2025	2024	2025	2024	2025	2024	2025	2024	2025	2024	2025	2024	2025
EAP	10	17	3	0	0	1	3	1	0	0	0	1	0	0	1	1	0	0	14	20
ECA	23	23	1	0	1	6	0	1	0	0	0	0	0	0	0	0	0	1	23	25
ESA	25	24	0	0	1	2	0	0	0	0	0	0	0	0	0	2	0	0	25	26
总部	14	15	0	3	1	1	0	0	0	0	2	0	1	2	0	1	1	0	18	18
LAC	16	20	0	0	0	2	1	11	0	1	0	0	0	0	0	0	0	0	17	32
MENA	11	10	1	0	1	0	2	1	0	0	0	1	0	0	0	0	0	0	13	12
SA	7	13	0	0	0	2	0	2	0	0	2	0	0	0	0	1	0	0	9	16
WCA	21	22	0	0	1	2	1	0	0	0	1	0	0	0	0	0	0	0	23	22
总计	127	144	5	3	5	16	7	16	0	1	5	2	1	2	1	5	1	1	142	171

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 数据库。

51. 四年内，各区域的提交情况各有不同，反映出评价周期、覆盖水平和工作环境存在差异。同时，有13个国家办事处在2024年和2025年连续两年均未提交任何评价，表明覆盖缺口依然存在。区域差异主要体现为评价周期和工作条件的不同，而非整体转向更高的提交数量。恢复进度还受到人员稳定性、委托实施时间安排以及内部协调情况的影响。规划保持稳定的地区，提交量回升更快；持续开展机构重组或人员更替的地区，恢复则较为缓慢。

图 2

2022-2025 年提交的评价，按地点分列



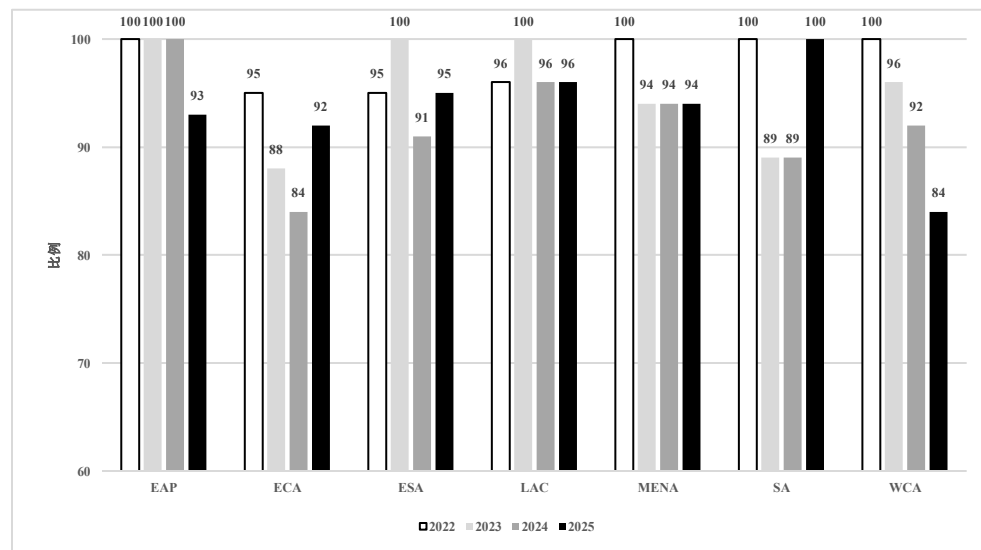
图例：EAP：东亚和太平洋地区；ECA：欧洲和中亚地区；ESA：东部和南部非洲地区；HQ：总部；LAC：拉丁美洲和加勒比地区；MENA：中东和北非地区；SA：南亚地区；WCA：西部和中部非洲地区。

资料来源：EISI 数据库。

52. 2025 年，大多数区域的评价覆盖率——即三年内至少提交一项评价的国家办事处占比——仍保持在较高水平。南亚实现全覆盖；拉丁美洲和加勒比、东部和南部非洲、中东和北非、东亚和太平洋（93%）以及欧洲和中亚基本实现全覆盖；西非和中非覆盖率虽略低于其他区域，但仍保持高位。

图 3

各区域的评价覆盖率，2022-2025 年



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 数据库。

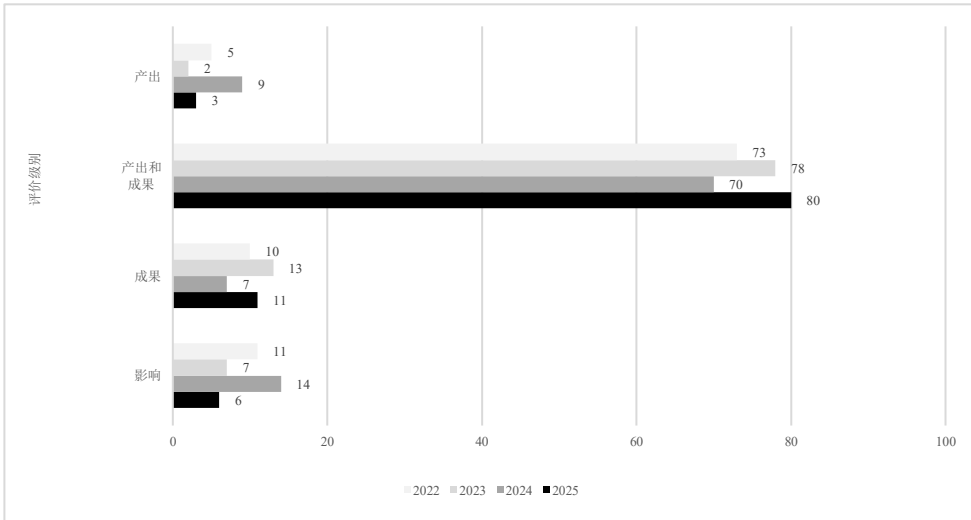
B. 评价概况

53. 儿基会评价政策承诺构建均衡的评价组合，使评价工作更多聚焦于儿基会的结果与影响级成果，兼顾总结性与形成性评价方式，实现各方案领域广泛的主题覆盖，并在适当且可行时积极开展联合、机构间、全系统及国家主导的评价。

目标与主题重点的均衡性

54. 2025 年，以成果或影响为重点的评价数量占比达到 97%，创历史新高，延续了 2023 年以来的上升趋势。与此同时，专门聚焦影响评价的数量仍然有限，反映出此类深度分析存在技术难度与成本方面的制约。

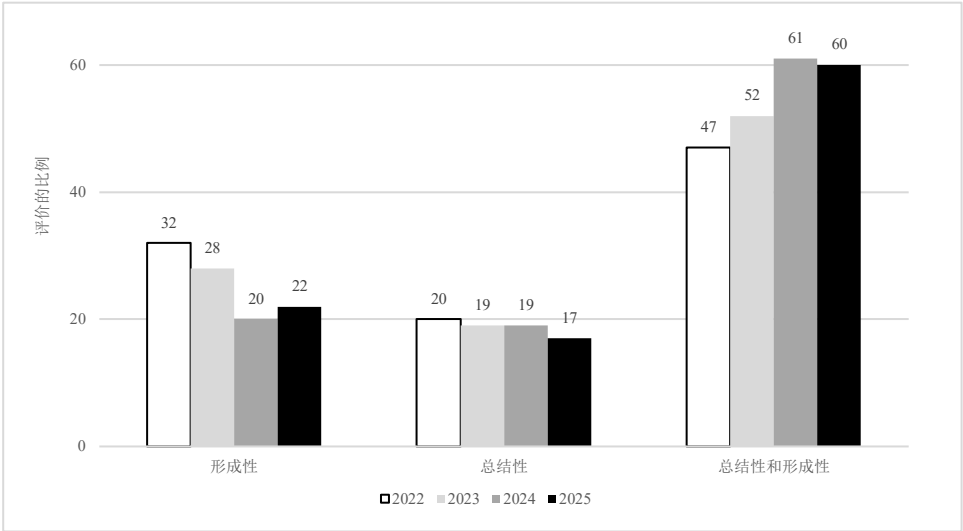
图 4
2022 - 2025 年按层级和年度划分的评价



资料来源：全球评价报告监督系统（GEROS）。

55. 形成性评价重在通过学习实现前瞻性改进；总结性评价则通过评估较长时期内取得的成果，强化问责。大多数评价应兼顾两者目标。四年间，评价目标的构成发生明显变化，综合型评价方式占比不断提高。这一趋势体现了修订后评价政策所作出的明确战略转向，也反映出现实层面的考量：在预算紧张的情况下，综合设计可使各办事处在满足绩效评价要求的同时，为方案调整提供依据，国家方案评价尤其如此。在复杂或人道主义环境中，这种综合方式显得尤为关键。

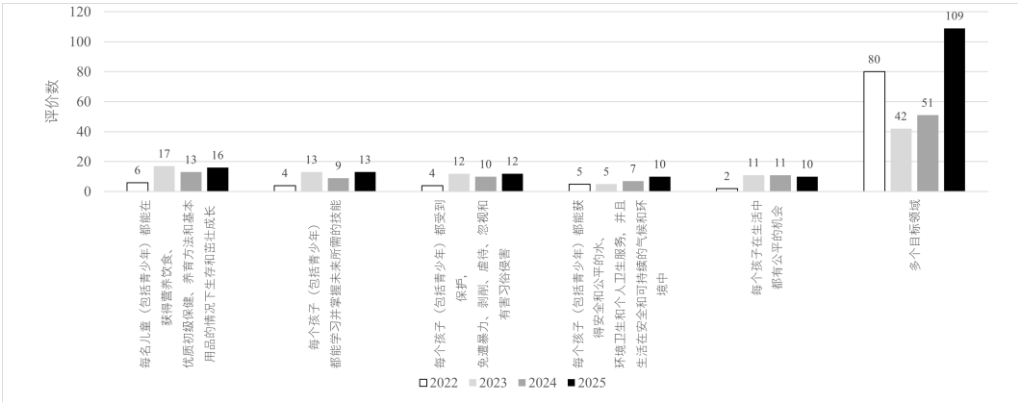
图 5
2022 - 2025 年评价的主要目的



资料来源：GEROS。
注：总数不等于 100%，是因为元评价占比不足 1%，已予以剔除。

56. 各方案领域的评价主题分布依旧保持均衡，覆盖多个目标领域的评价数量大幅上升，并在 2025 年达到历史最高水平。这一增长体现了儿基会有意推动跨部门综合方式，与儿基会日益重视一体化方案的导向相一致，也反映出国家伙伴对综合政策参考依据的需求不断上升。

图 6
2022-2025 年按儿基会各目标领域分列的评价重点



资料来源：EISI 数据库。

联合评价、机构间评价、全系统评价和国家主导的评价

57. 2025 年，儿基会继续开展联合评价、全系统评价和国家主导评价，但数量仍然有限。共有五项评价与其他联合国机构联合管理，三项为全系统评价，五项为国家主导评价。与往年相比，联合评价数量有所下降，国家主导评价和全系统评价则略有上升。这一模式主要受结构性因素制约，包括规划周期不匹配、方案重点各异、协调成本较高、资源紧张以及评价委托工作的分散化特点。尽管如此，儿基会与各

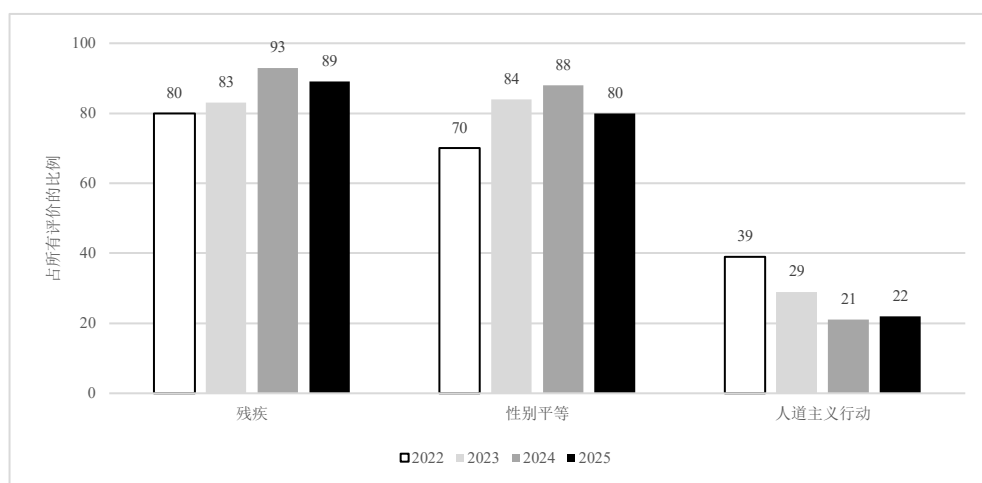
国政府及联合国机构的伙伴关系依然稳固，尤其在能力建设、联合方案实施和证据生成方面。因此，国家主导评价数量的上升，表明国家层面的参与度不断提高。

评价对残疾、性别平等问题和人道主义行动的覆盖

58. 总体而言，将残疾、性别平等和人道主义这三项跨领域视角纳入评价的程度依然保持在较高水平。尽管如此，进展情况并不均衡。在四年期前三年持续上升后，2025 年残疾视角与性别平等视角的融入程度略有回落。

图 7

2022-2025 年涵盖残疾、性别平等问题和人道主义行动的评价



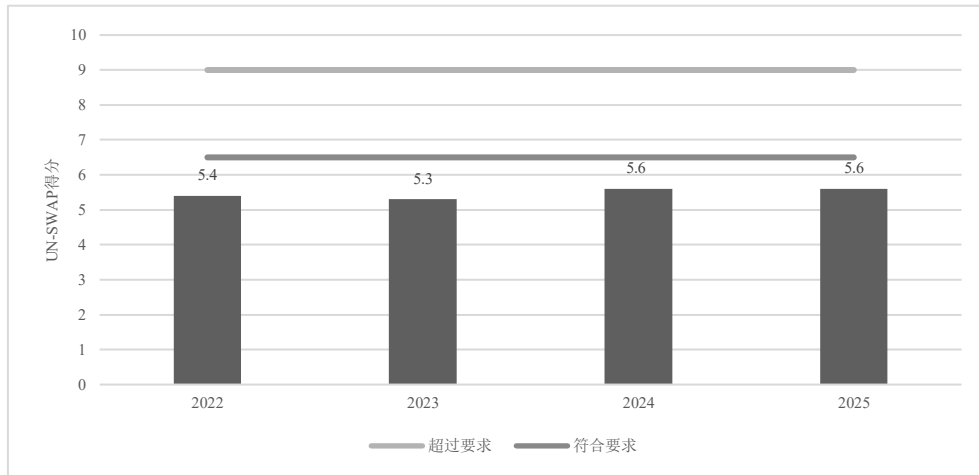
资料来源：GEROS。

59. 人道主义视角与另外两个跨领域专题完全不属于同一类别。这类视角在评价中的融入程度历来偏低，且自 2022 年达到峰值后持续下降，主要原因有三点。第一，2022 年属于异常情况，当年大量与新冠疫情相关的评价工作拉高了基数。第二，部分国家以发展为核心工作重心，并不适合普遍适用人道主义视角，在这类环境下系统纳入人道主义议题相关度较低。而另一些国家本身就属于人道主义情境，相关工作要么由纯人道主义范畴的评价覆盖（人道主义视角从一开始就作为评价核心），要么属于大规模复杂紧急状况，更多开展实时审查以支持业务调整 and 风险管理，而非事后续效评价。第三，有效融入人道主义视角需要专业能力，即便在应当纳入的场景中，也常因人员能力有限、指导不均衡、与人道主义伙伴协调难度大等因素受到制约。

60. 联合国全系统性别平等与妇女赋权行动计划指数，用于衡量评价中性别视角的有效融入情况。尽管 2024 至 2025 年持续发力改善该项指标长期偏低的状况，但表现仍与 2024 年持平，始终接近但未能达到“达标”门槛。制约因素包括：分性别和年龄的统计数据不足、监测体系未按性别平等分析设计、专业技术支持不均衡等。性别平等视角往往只在方案设计中提及，或单独列为一个章节，未能系统贯穿于评价问题、研究方法、调研发现和政策建议之中。依赖通用团队、时间安排紧张等问题，进一步制约了促进性别平等方法的系统性应用。

图 8

《2022-2025 年联合国全系统性别平等和增强妇女权能行动计划》中的评价业绩



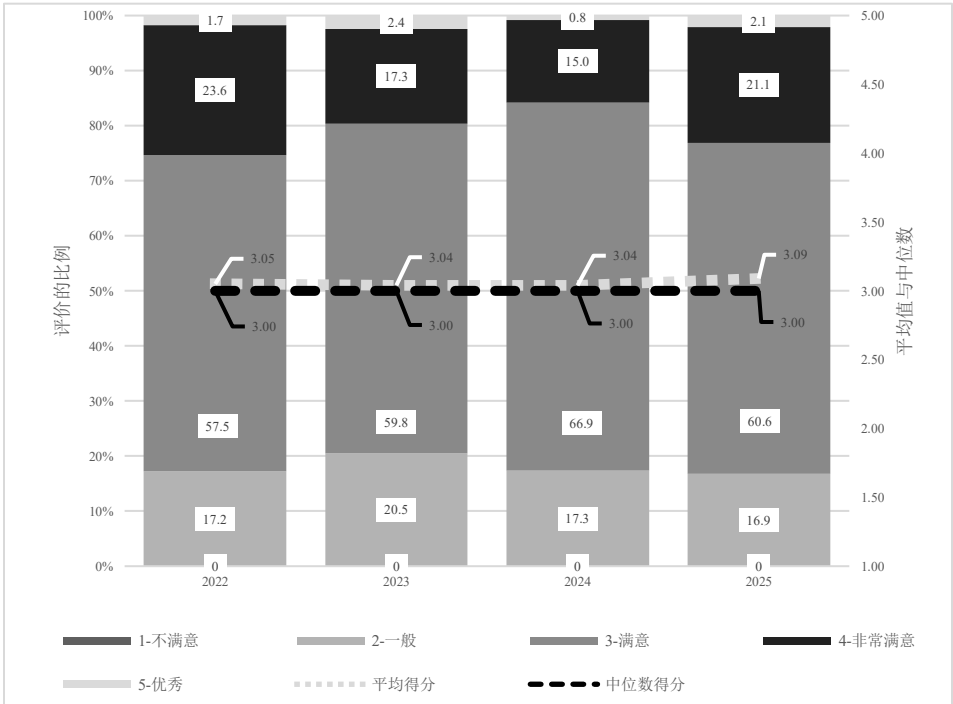
资料来源：GEROS。

C. 评价质量

61. 2022 - 2025 四年间，评价质量保持稳定，并在 2025 年略有提升。平均质量评分从 2024 年的 3.04 升至 2025 年的 3.09，中位数评分四年间稳定在 3.0，如图 9 所示，整体质量保持一致。

62. 多数评价达到合格质量标准，获评“高度满意”和“优秀”的比例持续上升，表明整体质量扎实，同时仍有空间进一步提升分析深度、方法严谨性和建议的清晰度。质量差异与保障条件密切相关：工期紧张、管理能力有限、技术反复评审机会较少，往往会降低分析深度；而在启动阶段提供有力的方法与战略指导、明确利益相关方职责、及早获得管理层与技术支持，则有助于提升质量。

图 9
2022-2025 年总体评价质量评级



资料来源：GEROS。

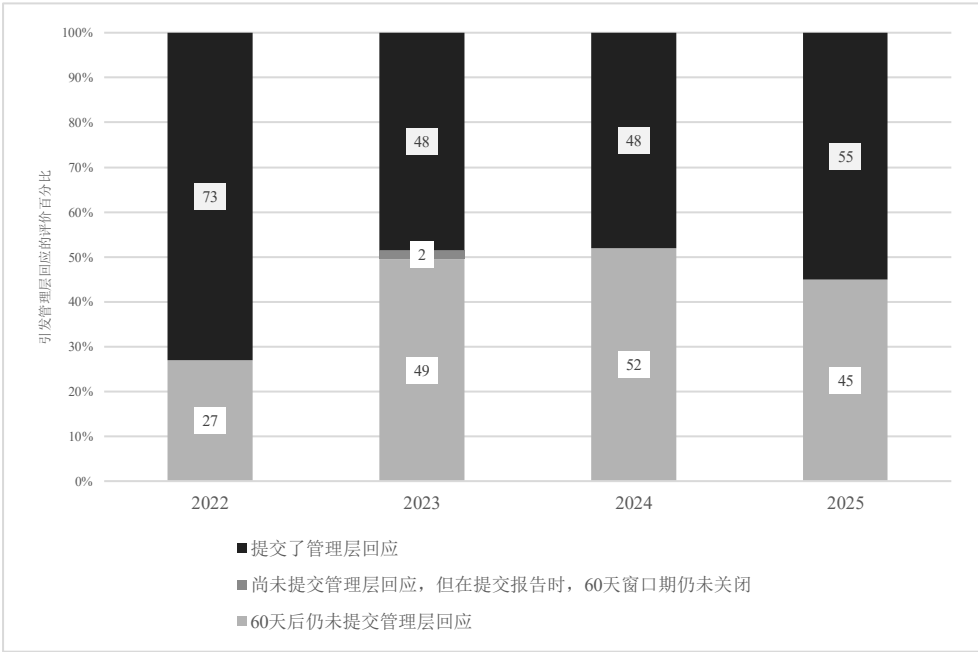
D. 评价管理层回应的提交和执行

63. 根据评价政策规定，管理层回应需在最终评价报告发布后 60 天内提交，内容应说明管理层对各项建议的接受程度，并列明建议落实方案。及时提交管理层回应虽能在一定程度上反映评价的应用情况，但并不代表相关建议已得到切实执行。反之，未能及时提交管理层回应，也不意味着建议未被有效落实；事实上，部分最具实质意义的行动有时会在管理层回应发布前就已开展。

64. 本年度的回应情况较往年有所改善，但仍有相当比例的评价报告超出 60 天时限，或在报告编制时仍处于待处理状态。看待这一情况还需结合未来重点倡议相关的组织调整带来的扰动，此类变动影响了人员稳定性与协调工作。总体来看，数据显示回应规范性正逐步提升，同时也表明仍需开展更持续的监测，并建立更完善的机制，以跟踪评价建议的实际落地应用情况。

图 10

2022–2025 年管理层回应提交的及时性

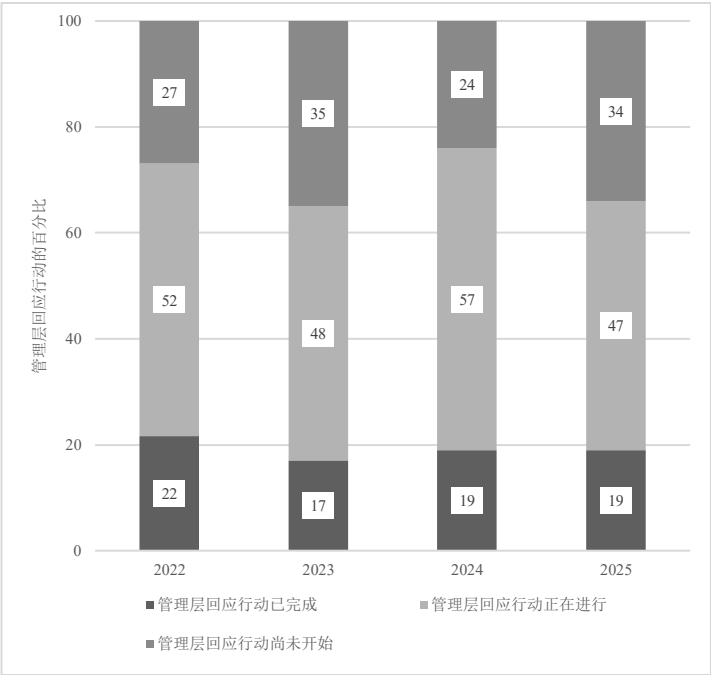


资料来源：EISI 数据库。

65. 对管理层回应落实情况的跟踪，可以反映商定行动是否转化为评价建议所要求的实际改进。四年间，落实情况整体保持稳定，多项建议已完成落实，部分仍在推进或尚未启动，这体现出将评价结果转化为业务和政策调整需要一定周期。

66. 总体而言，落实评价建议的相关机制已在运行，但仍需强化优先级设定、监测与问责机制，以保持推进力度。

图 11
2022–2025 年管理层回应行动及落实情况

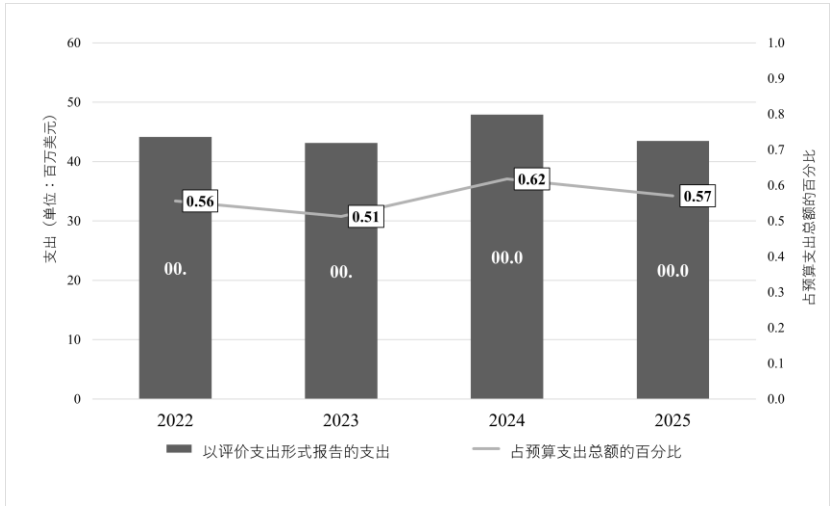


资料来源：EISI 数据库。

E. 评价支出

67. 作为一项政策承诺，儿基会目标是将方案支出的 1%用于评价工作。四年期间，相关支出总体保持稳定，但仍低于这一基准水平。

图 12
2022-2025 年以评价支出形式报告的儿基会支出（单位：百万美元）以及占支出总额的百分比



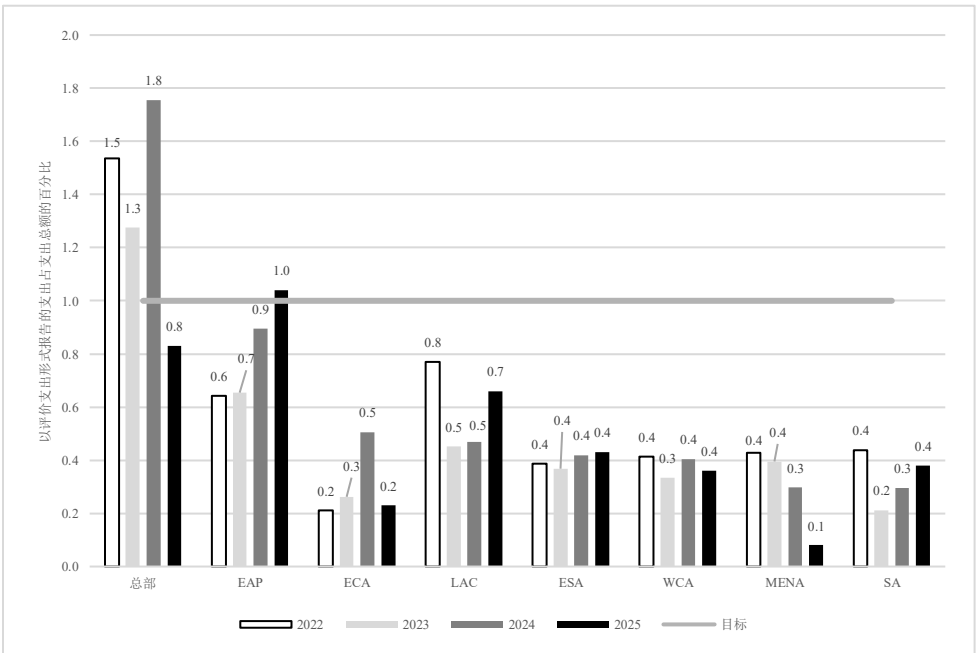
注：柱状图显示每年报告的评价支出（百万美元）。折线图显示评价支出占总支出的百分比，并标注各年度数值。
分子：评价支出；分母：方案支出净额。
资料来源：VISION 数据提取（2026 年 3 月 6 日）。

68. 报告支出与完成的评价数量并非直接对应。尽管 2025 年报告的评价数量有所增加，但评价支出的变化与完成的评价数量并非简单的一一对应关系。这反映出各办事处及机构层级在评价规模、范围、时间安排、复杂程度、组合管理以及成本编码做法上存在差异。

69. 该领域整体情况仍存在差异，部分地区阶段性达到机构基准，而大多数地区仍低于该水平。年度间的波动表明，变化主要源于组合实施时序、成本结构和报告方式的不同，而非单纯是评价业务量的变动。有两个单位（评价办公室及欧洲和中亚区域）在 2024 年出现支出峰值，但即便 2025 年提交数量增加，该峰值并未持续。这也印证了上文所述的情况：在相近或更为紧张的资源条件下，评价数量实现了增长。

图 13

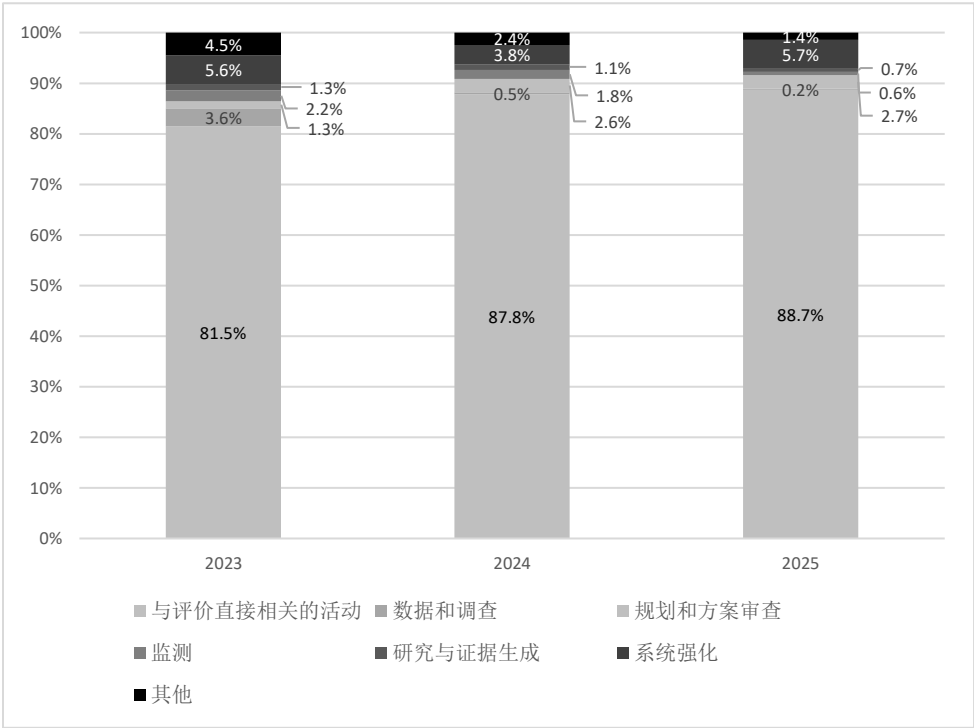
2022-2025 年以评价支出形式报告的支出占支出总额的百分比，按区域分列



图例：东亚及太平洋地区（EAP）；欧洲及中亚地区（ECA）；东部和南部非洲（ESA）；总部（HQ）；拉丁美洲和加勒比地区（LAC）；中东和北非地区（MENA）；南亚地区（SA）；西非和中非地区（WCA）。
资料来源：EISI 数据库。

70. 年度间变化部分反映了支出的编码方式，而指导原则执行不一致，限制了全机构报告投入的可比性。支出编码趋势显示，评价成本记录的准确性正逐步提升。四年间，以往易出现较高误编码的多数类别准确性有所提高，被正确编码为与评价相关的支出占比相应上升。这表明为改进编码规范所开展的专项指导与监测工作已开始见效。一个明显的例外是“体系强化”类别，其规模依然较大，并在前期下降后出现回升。该类别部分可能反映了用于国家评价能力建设活动的投入；后续将开展进一步分析，以明确其构成，并确保编码规范能准确反映与评价相关的投入。

图 14
2023 年至 2025 年以评价支出形式报告的活动，按活动类型分列



资料来源：评价支出数据取自 VISION 系统（2026 年 3 月 6 日）。

71. 综合来看，2022 - 2025 四年周期的绩效指标呈现出清晰一致的态势：评价工作在其最直接可控的领域取得了实实在在的进展，包括评价质量、聚焦成果层面、综合评价方法运用以及编码准确性等；而在依赖更广泛机构支持条件的领域，仍持续面临结构性制约，例如管理层回应及时性、评价支出水平以及跨领域视角的系统性融入等。尽管 2025 年面临巨大压力，上述整体态势仍保持稳定，表明本四年周期奠定的制度基础较为稳固。同时，持续存在的差距也明确指明了评价工作进入新四年周期后，需要持续关注的优先事项。

四. 结论和未来的方向

72. 2022 - 2025 四年周期伊始，评价工作正处于艰难节点：长达一年的领导空缺、新冠疫情期间长期远程办公刚结束不久，且儿基会内部各级分散化评价体系也才刚刚建立。在新一届领导班子带领下，本四年周期的核心工作是大刀阔斧地重整并激活分散的团队，将新近全球化布局的三级评价体系统一在共同愿景之下，并制定和实施修订后的评价政策，重新定位评价职能以实现最佳成效。这一举措确保各级评价工作具备相关性、及时性、可信度与实用性；确保评价工作保持支撑其公信力的独立性；并确保获得充足、可预期的资源，以履行其职责。

73. 正当工作势头逐步向好之际，未来重点倡议的启动带来了前所未有的制度变革与财务不确定性。未来重点倡议要求评价工作做出调整并重新设定优先事项：积极

参与变革进程，提供评估依据以支持机构渡过艰难转型期，同时持续推进政策落实、兑现《2022 - 2025 年全球评价计划》承诺，并制定适配新机构现实的后续计划。

74. 本报告所载成果反映了评价工作在这一严峻背景下取得的成效。从量化指标来看，体现出极强的韧性：报告提交、质量、管理层回应及支出等方面整体保持稳定，即便在 2025 年面临多重压力也未出现明显滑坡——这一成果并非理所当然，也来之不易。从质效层面来看，则实现了实实在在的进步。评价核心工作更加注重成果应用，所开展的评价切实推动了机构改革、方案设计、国家政策制定与战略决策，其中首次以如此直接且实质性的方式，参与制定《2026 - 2029 年儿基会战略规划》。日趋成熟的评价工作在体系构建与关系营造上取得显著进展：首部《儿基会评价手册》编制完成；全球评价领导小组（GELT）成为凝聚共识与集体管理的重要载体；在影响评价、国家评价能力建设及全球综合分析等方面的战略举措持续推进，进一步构建起以证据为基础、服务于儿童权益的问责体系，这也是评价工作最具雄心的长期目标。而《2026 - 2029 年全球评价计划》——规模更精简、优先次序更清晰、透明度高于以往任何一版——为下一个四年周期奠定了坚实基础。

75. 在此基础上，评价工作迈入新四年周期，明确了一系列优先事项：

- (a) **落实《2026 - 2029 年全球评价计划》中最高优先级承诺**，同时为其余重点任务争取资源；
- (b) **在评价中持续聚焦成果与影响层面的结果衡量**，并探索创新方法使之更加扎实可靠；
- (c) **通过推广已编制完成的《儿基会评价手册》**，确保评价始终具备高度相关性、及时性、可信度与实用性；扩充技术指导工具以提升高阶方法技能；制定模板；从评价设计阶段起强化质量保证流程；并依托新成立的全球评价学习委员会，以全体系技能梳理为基础，开展更贴合需求的内部能力建设；
- (d) **与管理层协同，强化“有用一用好”的转化链条**，包括提升管理层回应的及时性与监测力度、加强落实问责、更系统地跟踪记录评价的实际应用、对重点评价开展专项跟进，并持续运用评价综合分析，将反复出现的问题提请管理层关注；
- (e) **在财政环境日趋紧张背景下，巩固 2025 年在推动儿童事业发展、影响力超出儿基会自身方面取得的重大进展**，持续在国家评价能力建设、影响评价和评价综合分析方面发挥引领作用，响应各国政府对国家主导评价日益增长的需求；
- (f) **通过编制已计算成本的多年计划、改进支出跟踪与报告、按风险优先级统筹区域及国家办事处资源配置**，提升评价经费使用透明度；以及
- (g) **在全球评价领导小组（GELT）已显著增强内部凝聚力、协调性与协作性的基础上继续深化**，由其统筹全球评价实践，作为统一战略方向、应对机构内外部持续变革的平台。

76. 评价工作还将持续跟踪儿基会内部结构性变革，包括卓越中心设立、战略与证据办公室组建、区域办事处整合等。有关治理改革及内部监督安排的持续讨论，可

能进一步影响评价工作运行的制度环境。随着这些变革重塑整体证据生态，以及评价工作与方案领导团队协作、支持分散评价并促进机构学习的制度安排，评价职能将继续如 2025 年一样进行适应性调整。重要的是，《2026 - 2029 年全球评价计划》中载明的机构转型评价，将把未来重点倡议作为儿基会开展的若干变革管理举措之一纳入评估范围。由执行主任办公室委托开展的内部事后审议也将于 2026 年进行，评价办公室将继续与审议小组密切协作，最大限度减少两项工作重复、提升互补性。

77. 即将到来的一个尤为重要的机遇是，儿基会评价工作即将接受独立外部评估，该评估将由经合组织发展援助委员会（发援会）根据执行局第 2023/12 号和 2023/20 号决定开展。本次评估将以完全独立、公正的视角，审视评价职能的运行状况、成熟度及其支撑环境，有助于强化评价实践，并为 2027 - 2028 年评价政策的审议与修订提供依据。发援会在全过程中将广泛听取儿基会工作人员、执行局及审计咨询委员会的意见，并于 2027 年第一届常会上向执行局提交评估结果。对于一个在过去四年大部分时间里于艰难条件下夯实基础的职能领域而言，此次评估既是一次问责检验，也是一次重要契机——借此全面回顾职能建设取得的进展，明确需要进一步加强的方向，并为评价工作制定未来路线，使其在未来数年持续为儿童事业发挥价值。

五. 决定草案

执行局

注意到关于儿基会评价职能的 2025 年年度报告（[E/ICEF/2026/17](#)）及其管理层的回应（[E/ICEF/2026/18](#)）。