

工具 | 如何基于变革理论 开发监测评估框架



资助者圆桌论坛 (CDR)

2020年5月

背景信息

- 很多伙伴都知道变革理论，也知道监测评估，但却不一定对两者之间的紧密关联有深切感知。无论基金会或资助伙伴，我们都希望用监测评估来了解项目或服务是否改变了人们的生活。如果是，这一过程是如何发生的？如果没有，为什么没有？而变革理论正是指导监测评估回答以上问题的基础。
- 新慈善资本（New Philanthropy Capital，简称NPC）是一家英国的慈善智库，多年来NPC开发了一系列简单易用的变革理论、监测评估等慈善工具，为公益组织提供了诸多学习资源。本次我们将编译一份实用性操作指南：[如何运用变革理论开发监测评估框架](#)，供公益伙伴们，尤其是不熟悉监测评估或数据收集能力有限的小型公益组织伙伴，参考学习。
- 原文参见 www.thinknpc.org/resource-hub/understanding-impact/

本文目录

□ 本文目录:

1. 你评估的真实动机是什么？
2. 监测评估难在哪？
3. 我们可以怎么做？
 - 3.1 构建变革理论
 - 3.2 从变革理论出发分析可收集的数据
 - 3.3 为不同类型的数据选择合适的收集方法
 - 3.4 确定监测评估的优先级

1. 你评估的真实动机是什么？

- 通常公益组织希望开展评估是源于两个主要动机：
 1. 从评估发现中学习，提高有效性，为受益人带来更多成果和影响。
 2. 向资助者和支持者展示所取得的成果。
- NPC认为应当将“学习”作为我们的真实动机而非“展示”，因为关注“学习”更有助于公益组织：
 1. 保持对新信息和变化的开放态度，不断改进工作，提升效率；
 2. 通过学习了解哪些干预最有效，从而做出更好的决策；
 3. 即使面对复杂的社会问题，也能更明智地响应；
 4. 认识到事情总会变得更好，并接受反馈，从而变得更加负责和可信；
 5. 对于成功和失败，怀抱更加开放的心态。
- 而过多关注“展示”则会诱导你回避以上内容，只展示最佳结果，陷入自我陶醉，而不去想怎样会更好。

2. 监测评估难在哪？

□ 通常我们希望监测评估能够回答三个问题：

1. 我们做了哪些事？
2. 我们做得有多好？
3. 有人变得更好了吗？

□ 很多伙伴认同监测评估的重要性，但发现以上三个问题越来越难回答，为什么？

1. **数据收集困难**：公益组织经常旨在实现个人的改变，而这些改变是不明显的，且需要受益人的自我反省，他们可能不愿意或没办法做到（比如老人、小孩或重度残障人士），因此可能需要用复杂的方法去收集这些信息。同时，为了得出有统计意义的结果，往往需要足够大规模的数据，这也是一个挑战。
2. **归因困难**：即假设你衡量出了受益人发生了多大的变革，接下来如何确定这些变化有多少是你的干预措施带来的呢？虽然越来越多的从业者认识到归因的重要性，但对于大多数公益项目来说，从干预组和对照组获得足够的数据来证明统计学上的显著差异既不经济也不可行。

□ 以上两个困难，我们将在后续讨论成果数据和影响数据如何收集时予以回应。

3. 我们可以怎么做？

- 尽管存在上述困难，但所有公益组织都面临这些问题，每个致力于开展监测评估的组织可以一同正视、解决他们。
- 与此同时，以上困难只在回答“有人变得更好了吗？”这一问题时存在，而另两个关键问题“我们做了哪些事？”、“我们做得有多好？”则非常容易回答，并会为第三个问题提供大量有用的信息。
- 因此，公益组织可以从前两个易于回答的问题开始，制定对工作有积极影响的监测评估框架，促进对受益人的更大影响。**这并不难。**
- NPC认为评估什么以及如何评估，取决于你的优先事项、你希望评估结果能够引起谁的兴趣、以及评估计划的实际可操作性来决定。2014年NPC发布了开发监测评估框架的四个要素：**构建变革理论、确定评估的优先事项、选择所需要的证据水平、选择合适的资源和工具**（原文参见 <https://www.thinknpc.org/resource-hub/npcs-four-pillar-approach/>）。本次我们编译的是NPC于2020年发布的最新指导手册，以更简洁、实用的方式讲述了这些操作过程。

3.1 构建变革理论（1/2）

- ❑ 制定监测评估计划之前，要先制定变革理论的原因是：变革理论中突出强调了与评估相关的内容。它代表了你想要实现的目标，以及你希望目标群体如何从中受益，它能够帮助你从大量的信息中聚焦，定位出你关注的评估问题。
 - 例如，一个青年公益组织希望在受益人身上优先看到两个预期成果：（1）对社区中正在发生的事情有更好的认识；（2）更愿意参与到社区活动中。此时，预期成果中没有“更好地理解如何使用社交媒体”，因此前两个问题与评估相关，而这个问题不相关。
- ❑ 但需要注意不能完全依赖于变革理论开展评估，因为你不可能事先预估到一切状况，倾听用户反馈的声音能够更好地帮助我们发现预期外的成果。
 - 比如在上述例子中，可能在实际信息收集时发现一些受益人反馈说他们知道了如何更好地使用社交媒体，那么这会成为一个计划之外的评估点。如果在评估时不给受益人自主反馈的机会，就无法发现这一点。因此，我们需要保持开放，倾听伙伴故事。
- ❑ 这也同时强调了评估应由“学习”而非“展示”来驱动。由“展示”驱动的评估可能只会关注变革理论中的内容以展示变革理论的有效性，而以“学习”驱动的评估除了学习变革理论中的内容外，同时对其他可能发生的事情保持开放心态。

3.1 构建变革理论 (2/2)

- **思考监测评估框架的最好起点便是变革理论。** 变革理论是慈善机构思考和描述他们的目标是什么以及他们希望如何实现目标的过程。
- **NPC提出十步构建变革理论**（CDR已编译www.jianguoyun.com/p/DXP5_-8Q6YDCBhjW6JwD）
 - 第1步：情境分析。** 你希望解决什么问题？你的组织提供什么？
 - 第2步：目标群体。** 你能在最大程度上帮助或影响的人是谁？
 - 第3步：影响。** 你希望看到的持续或长期变化是什么？
 - 第4步：成果。** 对你的目标群体来说，什么样的短期变化能够进一步推动影响的发生？
 - 第5步：活动。** 你打算做些什么？
 - 第6步：变革机制。** 你的活动将如何产生你想看到的成果？
 - 第7步：排序。** 成果和影响可能发生的顺序是什么？
 - 第8步：变革理论路线图。** 你如何向他人展示你的计划？
 - 第9步：利益相关方和影响因子。** 谁（以及什么）是你达到预期影响的促进因子和阻碍因子？
 - 第10步：假设。** 你的变革理论经得起推敲吗？

3.2 从变革理论出发分析可收集的数据

- 从变革理论出发，可用于监测评估的主要有以下六类数据，主要从机构外部进行收集（下表按各类数据在项目生命周期中出现的顺序排列）：

对应的变革理论步骤	对应的数据类型
第1步：情境分析	背景证据：从已有知识中你了解到哪些关于问题和人们需求的信息？
第2步：目标群体	用户数据：服务是否被有效地递送给预期的目标受益群体？
第5步：活动	参与数据：服务在多大程度上吸引目标受益群体持续参与？ • 产出数据：你提供了多少服务？（从机构内收集） • 质量数据：你认为自己做得有多好？（从机构内收集）
第6步：变革机制	反馈数据：人们对于服务有什么看法？
第4步：成果	结果数据：短期内，服务如何影响或帮助人们？
第3步：影响	影响数据：上述取得的成果是否有助于改善人们的生活？

- 接下来将分别介绍每类数据：是什么、为什么重要、如何收集、以及关于该类数据的注意事项。

背景证据

是什么？	现有的任何关于问题和人们需求的知识： • 任何现有的关于问题大小和原因的信息； • 关于类似问题，以前有过的可能解决方案以及实证研究； • 与利益相关方交谈，获取新信息。
为什么重要？	变革理论中任何一个得到已有研究支持的方面都不需要再次验证，因此熟悉背景证据有助于节省时间和精力。比如如果你的干预措施旨在让人们多锻炼，则你不需要再证明锻炼如何改善幸福和健康，因为已有研究充分证实了这一点。
如何收集？	可以收集任何现有的外部数据、研究、证据或信息： • 收集来自官方的统计数据。 • 收集以前对类似干预措施的研究或评估，如在线搜索已有实证研究（比如3IE提供了多个领域的实证集）、文献综述等
注意事项	• 如果缺乏学术背景，可以与对相关议题感兴趣的学者合作。 • 即使不能对议题做系统性综述研究，也要尽量寻找潜在的相关证据。 • 关注已有成功案例的背景与你的项目背景差异，因为不同环境下同样干预措施可能导致不同结果。因此要密切关注实现你的变革理论中的预期成果的要素，如政策环境、需求层次、合作伙伴、基础设施等。

用户数据

是什么？	你计划服务的人的背景信息：他们的特征、他们面临的问题。
为什么重要？	• 用户数据帮助我们核查是否找到了合适的人。 • 能够有效地接触到合适的人，这本身就是成功的第一个指标。 • 初期收集的用户数据有助于用于后期对变化的分析，及对于成功和失败相关因素的探索。 • 借此机会，用更广泛的问题来深度理解用户需求，改进服务。
如何收集？	最好在与目标用户早期/第一次接触的时候进行收集。也可以通过合作伙伴或转介人那里直接获得用户数据。 • 客观特征（如年龄、性别）； • 主观特征（如态度、愿望）； • 其他你感兴趣的更广泛的问题（可能与基线调研重叠）
注意事项	• 需要哪些用户数据，应该直接由你的目标群体的特征来决定，这些特征是在变革理论的第二步中定义的。目标群体可以是项目的最终受益者（如学生），也可以是你希望通过其间接接触到最终受益者的人（如教师）。 • 过量的问题可能阻碍人们参与项目的热情，或干扰你的研究，你需要在必要的信息与不影响项目本身之间找到平衡，或者在建立更牢固的关系之后再开始追问。

是什么？

- 参与数据：人们参与活动的程度
 - 参与的频率、时间、持续时长、参与多个活动的情况
 - 参与的质量（是否积极参与等）
- 产出数据：你提供服务/物资的数量
- 质量数据：你对于服务质量的自我评价

为什么重要？

- 参与数据：①如果要达成预期目标**人们必须按计划参与你设计的活动**，因此参与数据是人们参与的证明；②与用户数据一样，参与数据也是成功的早期指标之一；③参与数据可与用户数据结合，探索哪类群体更愿意参与活动，以及与用户反馈和成果数据结合，可以分析参与时长/质量与成功之间的关系。
- 产出数据：通常不重要，也不具有启发性，但它可以提供服务概况，让出资人或其他方了解到项目进展情况、影响规模。
- 质量数据：有助于核验项目是否达到了你的/行业的标准和目标，并探索解决方案。

如何收集	• 直接在机构内部开展服务或物资递送的过程中记录收集。 • 如果你的服务涉及多个项目点，最好有一个数据管理系统。可能的话，参与数据尽量与用户数据一同管理，这便于分析服务与用户参与度的关系，以及可以采用怎样的策略留住你认为最受益的人。
注意事项	• 如果参与数据显示人们没如预期一样参与到活动中，说明有问题，需要引起关注。 • 产出只需要统计你提供了多少服务/物资，不需要统计有谁参与了。 • 为了判断质量，可以用定量和定性相结合的方法，如数据收集+同行/专家评审+员工自我反思。 • 为了更好地收集数据，建议： ➢ 让数据录入工作尽可能简单； ➢ 询问员工或志愿者关于如何收集数据的建议； ➢ 建议及时录入产生的数据，而不是隔段时间批量录入； ➢ 确保员工输入数据的一致性； ➢ 尽量采用电子数据。

是什么？	人们对项目的看法： • 活动是否按计划进行？人们接受服务时的想法和感受是什么？ • 活动质量如何？ • 人们对活动的总体评价如何？他们会向朋友推荐吗？ • 具体喜欢/不喜欢哪里？怎样可以做得更好？
为什么重要？	反馈数据帮助你测试变革机制，以确定你的服务是否按预期进行，是否产生了期望看到的反应（指人们对服务的态度、看法，如人们感觉舒适、信任、被倾听等），并帮助你改进。
如何收集	反馈问题中的部分内容由变革理论中第六步的变革机制决定，从此出发，首先问封闭式问题，然后再问一些开放式问题（关注人们对服务的看法），可参考“是什么”中列出的问题。反馈数据可以是正式的或非正式的： • 非正式反馈：鼓励用户随时、以任何方式反馈观点；这有助于让用户感觉自己被重视、被倾听，并帮助你察觉成功/失败的早期迹象；鼓励在项目全周期都以非正式形式收集反馈。 • 正式反馈：使用结构化的方式（如问卷或访谈等）开展调查；这有助于将人们引向你感兴趣的特定问题，并且结果更有说服力；频次不要太多，可以与后面的成果数据一同收集。

注意 事项

- 反馈不仅仅来自用户，也可以来自员工、志愿者和其他利益相关方，以及不参与或退出项目的人。**从不同的人那里可以学到很多关于如何改进的内容。**
- **关键的挑战是如何获得有代表性的反馈。**乐于反馈的往往是最积极和最消极的人，努力与不愿反馈或退出的人接触，并设计简短有效的反馈形式很重要。

- 余下的成果数据、影像数据往往最让人头疼。这就是前面提到的**数据收集困难**和**归因困难**这两个挑战真正出现的地方。理解它们之间的区别是很重要的，尤其是在做变革理论的时候。
- 接下来，我们将**通过成果数据来讨论如何应对数据收集挑战，通过讨论影响数据来讨论如何应对归因挑战。**

是什么？	目标用户因项目发生的短期（几天或几周）改变（如知识、态度、行为变化） • 人们出现你预期的正向变化了吗？ • 人们因项目出现了哪些你预期外的正向改变或受益？ • 服务的哪些方面、在什么情况下、为谁、提供了帮助？
为什么重要？	• 成果数据能够显示你是否真的为人们带来了改变。 • 出资方可能对成果数据更感兴趣，因为这能让他们看到他们的钱正在发挥作用。 • 成果是产生长期影响的好兆头，如果你的变革理论中已经实证支持已有成果与影响之间的关系，将会增加你的信心。
如何收集（未完）	• 与反馈数据相似，从变革理论出发，先问一些与你的预期成果相关的封闭式问题，然后再问一些开放式问题（关注人们的变化），参考“是什么”中描述的问题。 • 通过与正式反馈相同的方式进行收集，也可以与正式的反馈数据同时收集。

如何收集 (续)	为了应对数据收集挑战，以下建议供参考： 1. 借用前人用过的成熟工具或框架有助于保证信效度。 2. 采用前后测对比的方法有助于衡量变化，前测数据可与用户数据一同收集。 3. 根据你实际能够接触到用户的方式和频率来决定选择怎样的调研工具更合适。 4. 让员工或志愿者通过观察的方法收集数据（如人们的互动变多了吗）。 5. 酌情引入小规模定性研究，如访谈、焦点小组等。 6. 选择合适的样本量和调查频率，不需要调查所有人，也不要频繁调查。
注意事项	• 收集成果数据比收集用户、参与和反馈数据更难。即前文提到的数据收集挑战。我们希望看到的改变通常是内在的（如思维方式转变）或看不见的（如产生新的行为），这些“软性成果”往往难以被意识到是如何改变的。 • 为了避免各种偏差，你需要努力确保人们理解数据收集的目的，并且尽可能地鼓励对方保持诚实和坦率，同时保持调研问卷的简洁。 • 与反馈相似，成果数据的代表性也是一个重要挑战。

是什么？	持续、有意义的改变： • 是否出现了持续性的改变？ • 更广泛的社区层面有所受益吗？
为什么重要？	我们非常需要影响数据。我们都想知道在何时创造了改变，我们也想让其他人知道这一点。
如何收集（未完）	前文关于成果数据的收集建议同样适用于影响数据。为了应对归因挑战，以下建议供参考： 1. 充分挖掘、利用已有的实证研究为你的项目提供支持，尤其是能够证明你的短期成果与长期影响之间的关联的研究，你可以不必再次证明这些关联。 2. 开展小规模¹的定量和定性研究，获得关于潜在的成果和影响间关联的信息。定量研究尽量随机选择样本。定性研究可以通过一些简单的问题探索因果关系，比如询问：如果没有我们的服务，您认为会发生什么？是否有其他的因素让您的生活得以改善，比如重大的生活变化？ 3. 在不方便设置严格的对照组的情况下，你可以搜索与你的目标人群相同类型的群体的统计数据²作为对照组基线数据，这些数据可能来自官方统计网站/相关研究。随着项目开展，对比你的目标人群与该类人群的差异。

- | | |
|-----------------|--|
| 如何收集 (续) | <ol style="list-style-type: none">4. 努力寻找没有接受你的服务但与你的目标人群情况相似的人作为对照组，根据你的情况选择不太严谨的准实验或是更严谨的随机控制实验。5. 委托专业的第三方开展影响评估。这种方法成本较高，通常只有当一个项目影响范围足够广，并有足够的预算开展稳定的研究时，才是必要或可行的。 |
|-----------------|--|

以上1-3点很适合小型公益组织自己操作，4-5点则需要在专家帮助下实现，或通过大型资助项目获取共享评估框架的共建机会。

- | | |
|-------------|--|
| 注意事项 | <ul style="list-style-type: none">• 收集影响数据后，还需要考虑我们如何知道是我们的服务带来了变化？即前文提到的归因挑战。• 一方面，收集影响数据比收集成果数据更难。因为影响发生在未来，要想和人们保持联系并开展调研会更加困难。但另一方面，影响往往是更实质性的容易看到的重大积极变化，更容易测量；同时政府或官方可能可以提供一些影响数据（如学校出勤率）。• 我们往往希望知道发生了哪些正向影响，但同时也应收集负向影响数据。 |
|-------------|--|

3.3 为不同类型的数据选择合适的收集方法

数据类型	定量方法	定性方法
用户数据	• 报名/注册/转介表	• 报名时直接访谈用户 • 在转介表中设置开放式问题 • 访谈转介的人
参与数据	• 统计产出（如会议数、发布文本数等） • 记录人们出席及再次出席的情况 • 监测数据（如出席率、网站点击率、下载率等） • 监测质量（每人参与的时间、用户选择的多样性、训练有素的志愿者服务的比例等）	• 使用简单的量表或定性标签，由员工/志愿者评价服务提供质量及个体参与质量
反馈数据/ 成果数据	• 非正式方法：偶遇调查、举手/使用标签投票 • 正式方法：调研/问卷、核心项目人员对于人们变化的评估	• 非正式方法：意见箱、社交媒体、案例研究、直接询问意见 • 正式方法：焦点小组、深度访谈、访谈其他利益相关方
影响数据	• 二手资料 • 跟踪/纵向研究	• 用小样本进行跟踪/纵向研究。

3.3 为不同类型的数据选择合适的收集方法——示例（1/2）

假设有一个青少年慈善组织致力于提升困境青少年的教育成效，其监测评估方法与指标示例如下：

数据类型	方法	指标
用户数据	报名时填写的用户表格或由转介方发送的核心信息	出生日期、性别、居住地、种族、宗教、家庭情况、学校所在地、成绩
参与数据	电子化个案管理系统	- 青少年多久来参与一次活动/课程、每次参与多久、每次个案服务持续多久、第一次正式服务前转介所用总时长 - 青少年对于不同类型活动的兴趣大小
反馈数据	- 调研或简单的满意度评分系统 - 非正式方法如社交媒体、一对一交流或意见箱	- 对活动/服务的满意程度，是否愿意将项目推荐给其他人 - 对于活动设施的质量反馈、询问大家对于改进的建议（如如何让课程中的活动设计得更有趣）

3.3 为不同类型的数据选择合适的收集方法——示例（2/2）

数据类型	方法	指标
成果数据	• 焦点小组 • 调研	• 教育内容相关的知识、看法、态度和行为的变化 • 与教师、家长或其他相关方相比，青少年认为服务在促成这些变化的过程中发生了怎样的作用？（即归因） • 服务的哪些方面在什么情景下帮助了哪类青少年？服务的哪些方面没有帮到青少年？
影响数据	选取样本进行调研并设置对照组	与对照组相比，接受该项目服务的青少年成绩有所提高。

3.4 确定监测评估的优先级（1/6）

上述内容讲述了如何从变革理论出发设计监测评估框架，但在最终确定监测评估计划之前，还应确定优先级。因为你不可能收集所有数据，你也不需要花精力收集你不会分析或使用的数据。尤其是对于资源有限的小型组织来说。具体步骤如下：

1. 确定变革理论中的关键要素：虽然我们不会把不重要的东西纳入变革理论，但在其中总有重要性排序，**找出最重要的内容**。你可以询问自己、志愿者、以前服务过的人。

举例：在一个促进青年就业的项目中，你的预期成果可能有很多，比如青少年的自尊、自信、同理心、对工作的积极态度等，但最重要的是，**所有参与该项目的青少年必须明确他们接下来想要做什么工作，以及围绕该工作计划有哪些积极的期望。**

3.4 确定监测评估的优先级 (2/6)

2. 用“假设”来突出需要研究的问题：如果我们不考虑优先级问题，可能会直接评估变革理论中各环节的所有内容。但在构建变革理论的第9步和第10步中，我们会尝试提出对于自己和其他利益相关方来说关键的一些“假设”，即找出变革理论中薄弱、有风险或尚未得到实证检验的环节，这些是我们需要优先重点研究的问题。

举例：想象一个学校早餐俱乐部项目的简化版变革理论是：提供健康的早餐→儿童吃上健康的早餐→儿童在上课时注意力更集中→儿童成绩提高。关键假设有：

- 我们可以用小厨房和有限的预算制作“健康早餐”。
- 目标家长会首选带孩子去早餐俱乐部吃早餐。
- 目标家长不会事先给孩子们吃不健康的早餐。
- 孩子们会吃我们的健康早餐。
- 我们观察到，缺乏健康的早餐是一些孩子注意力不集中的原因。
- 更好的专注度将会提高孩子们的成绩。

这些是比简化版变革理论的描述更具体的一些问题，这些问题为我们确定评估内容提供了更明确的方向。比如，我们现在知道要**观察谁吃早餐、吃了多少，以及需要一种方法来测量注意力水平。**

3.4 确定监测评估的优先级（3/6）

3. 充分回顾、复盘已有信息：无论是外部研究或是你已经收集到的项目相关经验数据等，都有助于帮助你聚焦，缩小必要的研究范围，避免注意力分散。**你能得到多少信息将取决于你的服务或活动开展多久或有多创新。**具体要如何收集，参见前文对“背景证据”数据类型的介绍。

举例：在考虑绩效管理和评估时，我们通常是针对现有的服务和活动。**通常需要考虑收集项目历史数据和现有数据**，可以问问自己：我有哪些数据，它们保存在哪里？谁负责收集和管理这些数据？数据目前是如何使用的？这些数据有用吗？数据是否准确？有哪些数据遗失了吗？哪些数据收集工作进展顺利？哪些数据没有？

3.4 确定监测评估的优先级（4/6）

4. 考虑什么信息对你的决定最有帮助：区分优先顺序的最佳方式可能是预测你和你的利益相关者需要做出的最重要的决定，以及哪些信息对于做出更好的决定最有帮助。

举例：依旧使用前述的学校早餐俱乐部项目的例子。

- 首先我们需要迅速做出决定的有：早餐俱乐部是否吸引来了目标儿童，父母是否带孩子来，他们是否吃了早餐。所以，**我们需要保留关于参与的详细信息。**
- 假设在项目实际开展过程中，我们没能立刻接触到目标受众中的每一个人，我们要如何增加项目的吸引力？此时，**我们需要从非用户那里了解他们对早餐俱乐部的看法以及参与的障碍。**
- 最后，项目的资助方和我们自己都需要决定是否继续运行这项服务，这一决定要基于项目的成本和成果大小。因此，**我们需要信息来告诉我们目标学生的参与度、注意力水平，以及成绩的变化情况。**

3.4 确定监测评估的优先级 (5/6)

5. 制定一个可用于与其他人分享的评估计划：列出你的主要评估问题，即你的评估优先事项，并建立在你在第一部分所做的工作的基础上，使用不同类型的数据。这是一种向他人展示你计划使用哪些方法和指标来回答上述问题的有用方式。

6. 认真考虑是否要做影响评估，以及如何做：

- 影响评估的目的应该是更新现有的证据基础，而不是重复已有研究。因此，是否开展影响评估取决于你工作的行业或领域，以及你的工作有多创新。我们要充分挖掘、利用已经存在的所有信息，以确保我们只回答需要回答的问题。如果在变革理论的一个部分中已经有了关于因果关联的证据，这就不是进一步评估的重点。
- 影响评估是一件道德正确的事情，但因为影响评估的难度、成本均较高，因此我们只应在最紧迫的时候，有选择地偶尔尝试回答影响问题。也就是说，**偶尔做一次，但要做得好，这样做是为了支持整个行业知识的增长，而不是仅仅为了满足组织的需求。**

3.4 确定监测评估的优先级（6/6）

- 我们不鼓励为了获得资助而开展影响评估，因为这更像是为了“展示”而不是为了“学习”而评估。这时你可能会更关注有利于展示影响力的方法，而忽视能帮助你学习或做得更好的方法（比如定性研究）；同时，以筹资为目的的评估未必会产生你想要的效果，你可能花费了大量时间和金钱做出统计图表，最终资助方却对结果很失望，建议提前与内部和利益相关方确认需求、评估目的和需要报告的内容。

NPC写给资助者的话：

资助者应该认识到自己所扮演的角色，对于推动公益项目发展方向的潜力。同时也要意识到影响评估的挑战性。看似简单的评估要求可能会演变为困难、压力和资源浪费。

我们认为，资助者和理事会应该确保他们理解什么是影响评估，**确保只有在准备好支持评估并且清楚评估的目的时，再提出评估要求。**此外，有可能的话，把影响评估的重点放在行业普遍感兴趣的问题上，帮助建立整个行业的证据基础，而不是仅仅关注某个组织是否产生了影响。

总结建议

1. **用户、参与和反馈数据应尽可能多地收集（最好收集每个人），并将其作为长期的例行工作和高质量服务的一部分。**它们既是最容易收集的，也能为公益组织提供关于“我们做了多少”、“我们做得有多好”等问题的真实见解，这将有助于改进服务。重要事项：尽量让每个个体的前后变化数据能够联系起来，随着服务的增长，可能需要电子管理系统。
2. **成果和影响数据不需要一直收集，可以定期偶尔收集（如每两年一次）。**考虑到收集成果和影响数据所涉及的挑战 and 成本，公益组织应该一直问自己是否需要这样做。NPC认为对于较小的组织，尤其是那些提供轻度干预的组织，甚至可以说，**成果和影响数据不是必须收集的内容。**重要事项：如果要收集，样本必须有代表性，必须包括那些退出服务、没有真正参与的人，甚至那些不喜欢你的人。而不是仅关注成功故事。
3. 将定量数据（如参与度）和定性数据（如变化的故事）结合起来。
4. 最后，不需要一上来就考虑上述所有建议来改进你的评估方案，**最好的开始是通过试验找出对您和您的组织最有效的方法。**



资助者圆桌论坛（CDR）

CDR是面向中国资助者的技术支持平台，其使命是“服务中国资助者，探索有效公益，推动社会问题解决”。

cdr4impact.org.cn

版权声明：本报告版权归原作者所有，CDR将本报告编译，为公益同行学习分享所用。