

肆. 如何做监测评估



中国资助者圆桌论坛

2017年2月

资助工作者应知道如何用好监测评估

近年来，公益领域开始重视监测评估。然而，由于监测评估的内涵丰富，且仍在发展之中，如果不把其中的概念和原则弄清楚，不但不能发挥出其应有作用，反而可能会带来很多困扰。

本册帮助你了解监测评估的基本知识，从而知道在何种情境下应该开展何种类型的评估，这对自身以及伙伴的工作都将大有裨益。

《资助指南》其它四册内容，可在CDR官方公众号下载：

- (1) 资助指南总述
- (2) 如何做项目分析
- (3) 如何做机构分析
- (5) 如何做领域扫描

致谢！

感谢所有CDR成员机构以及数十位一线资助工作者，将他们的经验、智慧甚至烦恼无私地分享给我们。

感谢澳门同济慈善会、凯风公益基金会、南都公益基金会、招商局慈善基金会、浙江敦和慈善基金会（按首字母序）对CDR的机构支持。

感谢亿方公益基金会对本系列手册开发工作的支持。

未来，我们将继续和资助工作者们密切互动，同时汲取国际同行的相关经验，每周补充案例和工具，并定期更新手册内容，欢迎大家关注。

文责自负，指正请联系：

李志艳

18611645023

zhiyan@cdr4impact.org

如何做监测评估 · 目录

1. 为什么要监测评估

2. 评估的四种应用场景

3. 评估的原则与注意事项

问责的需要

公益本质上是受出资人/理事会之托，为服务对象带来正向改变。

这就意味着我们理应：

- 了解服务对象的变化，检验自身工作是否有效，并向出资人/理事会汇报。

目前，公益机构问责主要停留在两个层面：

- 资金问责——我没贪污钱
- 活动问责——我做了事情

由于没有常年积累的监测评估数据，在成效问责层面，绝大多数机构都遭遇到了困难。

改进的需要

- 所有的社会问题解决方案并不是“设计好了，然后执行”这样的两段论，而是持续深入的“构想——执行——监测/评估——优化”的循环过程。监测评估不仅是在验证方案，更是在设计与发展方案。
- 问责与改进，一体两面。只有真诚地验证有效，直面“无效”，才可能做出改进。问责，只有是为了持续改进，而不仅是简单的监督或奖惩，也才能收集到真实的信息。
- 问责与改进，两者之间也存在一定张力，评估本身是一场“政治事件”，资助工作者需要体会并把握好其中的分寸。

建立关心服务对象、关心成效的机制和文化



理论上，出资人、公益从业者都是为了服务对象而工作，但现实中，由于服务对象处于弱势，不能或不愿把真实的意见反馈出来，这常常导致公益机构发生服务对象偏移、工作目标游移的现象。

因此，追求成效的公益机构应有意识地在机构内建立一种关心服务对象，关心社会成效的工作机制和文化，有意识地收集服务对象的建议和反馈，并使之成为组织进一步发展的智慧和动力来源。

如何做监测评估 · 目录

1. 为什么要监测评估

2. 评估的四种应用场景

3. 评估的原则与注意事项

监测评估的四种场景

资助工作者
跟进
项目进展

严谨回答干
预是否有效

帮助被资助
方建立内部
评估机制

资助机构
自身的评估

1 资助工作者跟进项目进展

资助出去的每一个项目都需要跟进，由于项目众多，资助官员要能判断，在有限的时间内，需要收集哪些关键信息，以及用什么方式来收集。

常见内容

- 服务递送的种类和数量
- 抵达人群的数量和特征
- 对服务对象变化的观察
- 被资助方的学习与成长
- 项目有哪些重大风险

对外

- 收集进展报告
- 每年走访项目点1-2次，与团队成员、服务对象等对话
- 实际参与被资助的项目的关键环节

2 严谨回答干预是否有效：影响力评估

无论是让自己更清楚，还是对外传播更可靠的信息，在某些时候，我们总是希望能够严谨的回答：“我们的项目有效吗？”

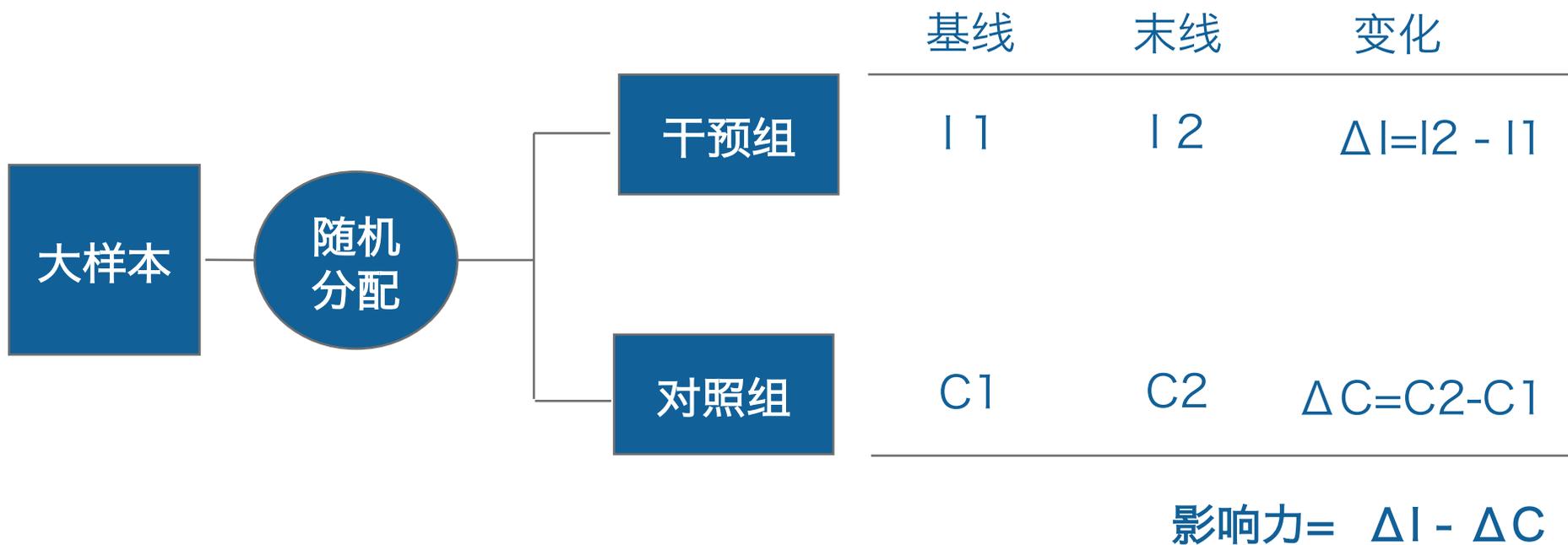
影响力评估，即：完全归因于项目的，排除其它因素之后，纯粹由项目给社会带来的带来的“净影响”。

社会影响力评估 = 归因分析 = 净影响

归因分析是公认的评估难点，目前受到学术界认可的严谨的评估方法即“社会实验法”，其中的随机控制试验法（RCT），被称为评估的黄金标准。

影响力评估的方法

社会实验法，是将大量样本按照随机或者其它分配方法，分成干预组和对照组，如果试验设计得当，干预组与对照组除了有无待评干预之外，其它方面统计意义上无差异，两组分别测量前后变化，两项对比即项目的“净影响”。



影响力评估的应用场景：稳定干预大规模推广之前

影响力评估，是对特定干预进行大样本定量评估。它有两个要求：

- 干预要稳定，不可反复变化，因此适合成熟的稳定干预
- 耗费资源，一场评估大约要几十万至数百万人民币

因此，成熟稳定干预大规模推广之前才需要做影响力评估，当然也有必要做，如果干预被证明无效甚至是负效，越是推广，对社会的影响越不好。

典型的影响力规模化情景：

- 向政府提出政策建议（以事实为基础的政策建议）
- 公益机构推广特定类型的产品或服务

案例：歌路营《新一千零一夜》影响力评估

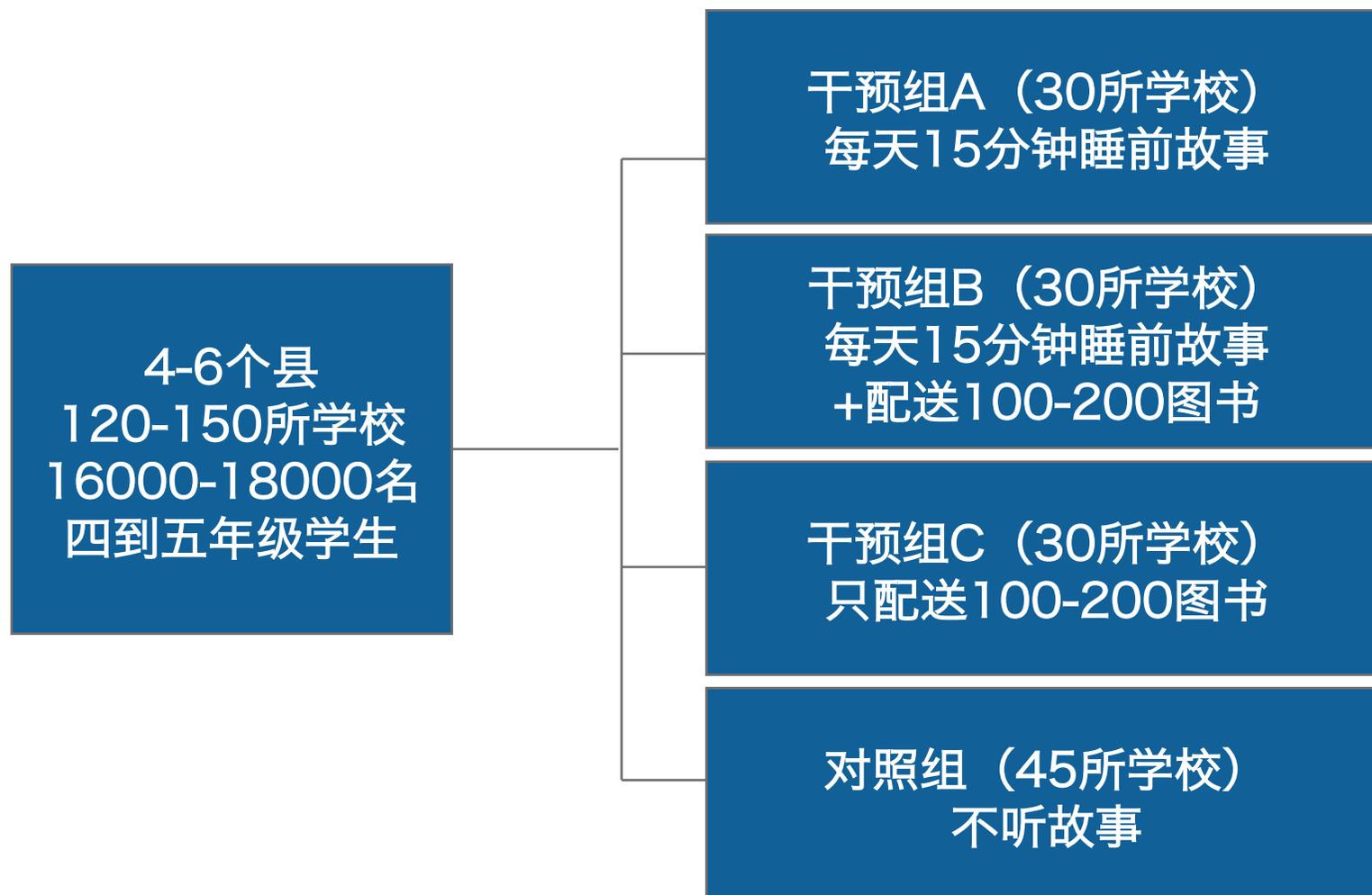
为提升农村寄宿制学生的身心健康水平，歌路营推出了“新一千零一夜：乡村学校住校学生睡前故事干预”项目。“新一千零一夜”目前已经服务1947所学校，未来希望能够拓展至近万所学校。在大规模推广之前，歌路营希望对干预效果做全面而科学的评估。

评估主要回答四个问题：

1. 每天15分钟的睡前故事对农村住校生的阅读能力和阅读习惯有什么影响？
2. 每天15分钟的睡前故事能提高农村住校生心理健康水平吗？
3. 每天15分钟的睡前故事有助于提高他们的学业成绩吗？具体地，能提高他们的语文成绩和数学成绩吗？
4. 每天15分钟的睡前故事有助于改善他们的平时人际关系吗？具体地，能改善同学们之间宿舍关系吗？能有助于减少校园暴力（同学间欺负行为）吗？

案例：歌路营《新一千零一夜》影响力评估

评估采用了随机控制实验法，评估仍在进行中，下图介绍评估设计：



阅读已发表的科学评估报告

我们无力为每个项目做影响力评估。幸运的是，国际上已有大量科学评估报告发表，大家可以阅读与自己工作领域有关的报告，用于优化和改善相关干预。

常见的影响力评估报告下载地址包括：

<http://www.3ieimpact.org/> （综合议题）

<http://evidencebasedprograms.org/> （综合议题）

<https://www.povertyactionlab.org/> （综合议题）

<http://ieg.worldbankgroup.org/> （综合议题）

<http://www.educationinnovations.org/> （教育议题）

<https://www.unicef-irc.org/> （儿童议题）

<http://ceee.snnu.edu.cn/> （含中文政策简报）

3 帮助被资助方建立内部评估机制：精益评估

- 开展影响力评估有特定条件，并不适合所有机构、所有项目。尤其是一些初创和发展期的机构，需要一种更加简单实用的评估工具，在这里我们将介绍一种“精益评估”。
- 有些资助机构要求一些短期（如一年期）项目做评估，实际上如果不是持续开展的产品或服务，开展评估的必要性不是很大，还不如在机构层面提高评估能力，建立核心业务的评估机制。

精益评估的特点与内容

主要特点

- 它是机构工作流程的一部分
- 可以由非学术人士操作
- 成本低、反馈快
- 可以为机构战略以及业务产品的优化提供针对性建议

主要内容

- 记录服务对象的特征，帮助界定用户定位
- 记录用户参与活动的状况
- 记录发生在用户身上的变化
- 满意度调查
- 记录用户建议与反馈

案例：格莱珉银行PPI (progress out poverty index)

格莱珉银行推出小额信贷之后，世界银行、孟加拉发展银行等机构为其提供了影响力评估以验证其效果，与此同时格莱珉银行推出了机构自身的内部评估工具PPI，它由以下指标组成：

- 会员及其家庭成员居住在有马口铁皮屋顶或价值超过25000塔卡的房子里，家庭成员睡在床上而非地上。
- 会员饮用自管道井中抽上来的干净水、开水或者用明矾、净水药片或罐壶过滤器净化过的无砷水。
- 会员家庭超过6岁的孩子都在上小学或者小学毕业。
- 所有家庭成员都是用清洁卫生的厕所。
- 家庭成员有足够的衣服满足日常需要。
- 家庭有其他收入来源，例如蔬菜园子或果树，以满足额外之需。
- 借款人在储蓄账户中维持在年均5000塔卡的余额。
- 借款人全年有能力供应家庭成员一日三餐，家庭没有食品匮乏之虞。

案例：格莱珉银行PPI (progress out poverty index)

- 小额信贷在操作中很容易发生两种偏移：一是服务对象上移，即偏向于服务贫困地区相对不贫困的人口，二是慢慢局限于经济扶贫，忽略了对社会因素的关注。PPI指标体系，清晰地说明格莱珉银行想要服务真正的穷人，同时关注用户的教育、健康等社会状况。
- 信贷员可以方便地在贷款前入户调查时，使用PPI识别真正的穷人，还款时用PPI记录贷款户家庭变化。格莱珉银行长期在年报中使用PPI对外披露机构扶贫进展。
- 格莱珉银行后来成立了专门的机构对外推广PPI，并且根据不同国家不同地区的国情，推出了60多个国家或地区的PPI当地版本的指标。

4 资助机构自身的评估

资助机构不能只是要求他人做评估，自己做评估是对合作伙伴最好的倡导。

资助机构也可以使用“精益评估”的理念，建议侧重于：

- 了解被资助机构/项目的发展状况；
- 机构对被资助机构的承诺是否兑现，兑现程度；
- 收集被资助方对资助战略与实践的反馈与建议。

案例：美国有效慈善中心（CEP） “伙伴反馈报告”

美国前50强资助机构中的绝大多数，每年都会聘请CEP开展“伙伴反馈报告”，邀请被资助方（也包括被拒绝的申请方）对资助机构以下方面做出评价与反馈：

- 资助机构战略及政策的清晰度
- 对被资助方所在领域的理解力和影响力
- 对被资助方服务对象的理解力和影响力
- 对被资助方机构的理解力和影响力
- 资助筛选、监测、互动等流程
- 非资金支持

一些基金会如盖茨基金会，还设置了网上申诉渠道，用于被资助机构随时匿名反馈意见。

如何做监测评估 · 目录

1. 为什么要监测评估

2. 评估的四种应用场景

3. 评估的原则与注意事项

原则与注意事项

- 资助机构首先自己做评估，如果我们不做评估而力推评估，于理有亏，而且我们也很难体会评估在真实操作中的细节与困难。
- 以被资助方为中心，只有被资助方认同评估的价值，评估才会成为有用的工具。
- 不管何种类型的评估，都要一定的资源和技术，如果我们要求被资助方开展评估或提供评估数据，最好同时匹配资金和技术支持。
- 评估要求客观性，因此不宜成为筹资或者说服理事会的工具，它对筹资或说服有用，但如果仅定位于此，评估则容易变成双方合谋寻找证据的工具。
- 评估也是沟通工具，开展评估提供了一个利益相关者共同研讨和提升工作成效的机会。

参考文献

- 李志艳, 《公益的方法》、《一种真正在用的评估》
- Acumen Foundation, The Lean Data Field Guide: Tips for Collecting Customer Data to Build More Impactful Business
- 北京大学中国教育财政科学研究所、歌路营, “新1001夜农村住校生睡前故事项目评估计划”
- 阿西夫·道拉, 迪帕尔·巴鲁阿, 《穷人的诚信: 第二代格莱珉银行的故事》, 中信出版社
- Grant Makers for Effective Organizations, How Can We Help Our Grantees Strengthen Their Capacity for Evaluation

出品单位



中国资助者圆桌论坛是中国资助机构的平台机构，其使命为“服务中国资助者，探索有效资助，推动社会问题解决”。

资助单位（排序不分先后）



亿方公益基金会成立于2013年，其使命是关注公益研究与社会企业，支持民间公益组织，推动社会创新发展。



招商局慈善基金会成立于2009年，以“关注民生、扶贫济困、热心公益、和谐发展”为宗旨，主张通过理性的思考、实事求是的态度、创新和可持续的做法，给人提供向上的阶梯，推动平等合作，建设更加富强、公正、美好的社。



南都公益基金会成立于2007年，使命是支持民间公益，愿景是社会公平正义，人人怀有希望。



澳门同济慈善会在中国内地的资助工作，以让每个中国儿童都能接受公平优质的教育为使命。帮助中国处于弱势的儿童不仅能上学，还能上好学，充分发展个人潜能，提升儿童对未来学习、生活及工作的适应能力，进而能够平等地参与并且促进中国的健康发展。



凯风公益基金会是段伟红博士于2007年成立的全国性非公募基金会，通过支持思想的生成和传播，致力于推动人类进步，促进社会发展。



浙江敦和慈善基金会成立于2012年，以“弘扬中华文化，促进人类和谐”为使命，秉持“尊道贵德”的价值观，在国学传承、公益文化、公益支持等领域开展资助。