

Praxis 论文之二

面对挑战：组织能力建设影响评估

作者：John Hailey, Rick James and
Rebecca Wrigley

国际非政府研究中心

2005年2月

目次

摘要	3
引言	4
1 挑战面面观.....	5
1.1 项目和程序设计模糊.....	5
1.2 权力, 控制与所有权: 谁的需要和议程?	7
1.3 测量复杂无形的变化.....	7
1.4 表明因果关系和变量.....	8
1.5 对环境与文化的回应.....	9
1.6 关于投资费用	9
2 克服挑战: 实践的意义	9
2.1 利益相关者与优先性.....	10
2.2 自我评估	11
2.3 整合方法	12
2.4 不同方法与工具的平衡.....	12
2.5 简单而系统化	14
2.6 接受可能的联系, 而不是直接的因果	15
2.7 确认投资的层面	16
2.8 组织学习: 把评估与行动联系起来.....	17
3 结束语	17
4 下面的步骤.....	18
参考文献	20
缩略词表	22

面对挑战: 组织能力建设影响评估:

作者: John Hailey, Rick James , Rebecca Wrigley

关键词:组织能力建设,影响评估,监测与评价,组织学习

摘要

评价影响一个复杂的过程,特别是对像组织能力建设这样错综复杂,难以捉摸经常错误定义过程的影响的测量。这是一个飞速变化、日新月异的领域,虽然在评估方面有了一些进展,但却受困于定义上的麻烦、方法论上的争执、各执一词的批评和对这种评估过程的主要目的的不确定。

本篇论文概述了组织能力建设干预的影响评价方面的思考与实践。强调了一些概念、方法论和实践上遇到的挑战(包括清晰,权力,文化等)然后从总体上介绍非政府组织和公民社会组织在克服这些挑战的实践中采取的办法。论文是经过深思熟虑的,旨在让实际工作者(特别是那些在发展中国家和处于转型国家的实际工作者)介入这种有成果的辩论中。

本论文确认了 INTRAC 的 Praxis 项目在未来将关注的关键挑战,它们包括;改善组织能力建设影响评价的特征的理解的需要,产生并纪录变革性的、可采用的、可实现的方法。最后的挑战是考虑如何提升组织能力建设效果评估的形象,使人们将其看成帮助组织学习的重要工具而不是浪费时间金钱的负担。

引言

过去十年里，发展援助部门增长迅速，同时，人们也越来越关注这个部门的效果和影响。援助预算的紧缩使许多主要的捐助方对资助对象越来越挑剔，对诚信、业绩和结果要求得越来越严。非政府组织也越来越迫切地意识到对其本身效果评估的重要性。

但是，由 OECD/DAC 于 1997 年进行的有关 NGO 发展计划和项目的研究却得出以下结论：

尚缺少“足以得出以下方面的坚实结论的数据和信息：有关计划（projects）的影响、它的效率和有效性、它的可持续性、计划在性别和环境方面的影响及对加强民主力量机构、组织和构建公民社会方面的贡献。甚至缺少超越具体项目之上对非政府组织发展干预的影响的评价的数据，更不用说有关建设和加强体制能力方面的数据了。（Riddell et al. 1997）

作为对发展干预是否产生长期的、可持续的变化的测量，影响评估显然是个复杂的程序。如果测量内在摸不着的、流变的、反复的过程，影响评估和复杂性就随之增加。能力建设恰是一种这样的过程。捐助者和非政府组织都非常重视能力建设，因为他们意识到社区小组、地方非政府组织组织和其他公民社会组织处于与贫穷和社会不公战斗的前沿。因此他们关注不同的能力建设干预的动力机制和效果，关注如何找到评价其长期影响的方式。比如红十字会国际联盟能力建设框架 2001 认为：所有能力建设干预必须要测量并记载下来，以确保可问责的程度并促进跨组织学习。

组织能力建设可以被看作是有意识的、全面的干预，旨在改善与其使命和内容相关的组织的有效性和可持续性。¹ 干预集中于确认和发展组织内的能力要素：诸如技巧、系统、领导力，项目操作、对外关系。干预可以在不同层面发生，比如：为个别的成员提供培训课程、团队建设、高级经理人员辅导以及组织层面上的愿景展望和战略计划。这些在组织功能和过程上的投资目的在于对项目操作引起实际变化，最终，是为了改善穷困、最边缘化的人群的生活。这就对影响评估提出了特别的挑战，尤其是它是否能表明在一项特殊的组织干预和较大的过程变化间存在因果关系。比如，在建立成员行为评价程序和使最易伤害的群体状况改善之间是否能发现一种关联？

对行为和结果的强调也使得能力建设影响的评估成了件对参与机构有风险的事，因为除非结果与资助与否不发生关联，否则会将当参与机构推到易受损失的位置。因此，权力动力在效果评价上或明或暗充当着重要的角色，这使许多人将影响评估视为外部强加的控制和问责的机制。而且，许多效果评估办法偏于机械的、线性的、标准化的和象征性的。它们还经常与实际决策相脱节。一旦获取了足够信息，又往往不能恰当分析，及时地、以易用的方式予以传播。这就难以让人获得有意思的思考，难以将得到的成果用于实际，改善实际行为。

但是，近来在影响评估上取得了进步，特别是在发展和使用更多的合适的指标上。这些指标曾被看成是生硬的、不灵活的，但人们开始试图解决上述问题，比如，

¹ 定义基于 James 2001.

通过采用参与式评估过程发展出一套与不同的利益相关者有关的指标.它增强了对评估结果以及评估过程的拥有感.

在这种情况下,**Praxis** 论文致力于从总体上提供有关组织能力建设影响评估的新思想,确认有哪些障碍使之难于付诸实践.它不够精确,也不够全面,却有启发性,能激发人们深入论辩,引起不同反响,促使人们发现对能力建设实践者来说有用的,恰当的,可操作性强的解决方案.

第一部分提供了评估组织能力建设干预影响时来自概念、方法和实际方面的挑战;第二部分转入对一些非政府组织和公民社会组织(CSOs)为克服影响评估中遇到的这些的挑战所采用的程序和实践;结语的评论部分强调了**Praxis** 项目中为了产生和记载更多更有创意的、更适当的方法未来要重点做的工作.最后的“下一步骤”部分讨论了怎么应付论文中提到的各种挑战.

1 挑战面面观

评估影响是个复杂的过程,而测量时内在的复杂性、不可触摸和经常错误定义的程序使这一过程更加困难.组织能力建设就是如此.任何在这个领域工作的人都会体会到这种艰难.

领域内一直存在着关于组织能力建设影响评估的有效性和价值的辩论,人们也普遍担心它对被评价组织提出太高要求,是浪费时间.评估往往被设计成满足捐助者的体制需要,或者让外界更好地了解这个组织,而不是帮助发展组织成员的视野和技能.长久以来,评估被看作外部施加的活动,不是旨在使组织成员获得反思和新知,而是向外界提供信息.

但是,对与测量效果相关的概念、方法论和实践挑战方面,人们的理解有了发展,下面将对此重点详述.这些挑战与多数发展项目有关,但是,人们对这些挑战在组织能力建设过程中的特殊表现显然所知不多:

1. 不清晰的项目和程序设计
2. 权力,控制和所有权,谁的需要和议程?
3. 测量复杂和无形的变化
4. 表明原因和结果
5. 回应环境和文化
6. 考虑投资成本

1.1 项目和程序设计模糊

影响效果评估过程有效性的最重要问题之一是:对项目和程序本身的目的有多大程度的清晰把握.通常,在设计组织能力建设项目时,人们对提升意识,消除人们对能力方面的疑惑强调不够,对在利益相关者建立有关项目目

标旨在实现哪些方面的变化的共识关注不够。因此，项目目标往往表述得含糊其词，并以此作为共识的基础，其原因部分是由于参与方有不同的文化、组织和社会视角。（Morgan 1999）。

另一方面，项目的逻辑框架虽然有助于澄清目标、实施战略和监测与评价的标准，但如果结构过于死板，就很难表达能力建设过程流变、反复的特点。（ECDPM 2003）。

如果项目的概念框架不清楚，就会使影响的监测、评价和评估体系的设计复杂化---试想，如果你不能确定你的起点在哪里、到哪里去，你又如何能知道怎么能到达那里？或者正如《艾莉丝漫游仙境》中的一句话，“如果你不知道到哪里去，那么，条条道路都能抵达。”影响评价框架还要对寻求哪个层面上的变化（比如：内部组织变化，外部项目操作，或者提高的操作引起的人们生活的变化）给出定义，作为一个相对较新的术语，影响评估经常与项目产出监测或是直接目标与结果的评价相混淆。下面的列表是消除产出、结果与影响间混淆的方法之一。（选自：Fowler 1997）：

测量的要点	测量种类	测量内容	指标
产出	监测	努力	活动的实施
结果	评价	有效性	产出的利用和收益的持续产出。
影响	影响评估	变化	与当初有问题状况有多大不同。

不弄清楚项目过程要扮演的角色，会增加设计清楚的影响评估过程的难度，这些不同角色包括：

- *传统角色*:确认、分析和记载发生了哪些变化，以纪录影响。
- *交流角色*: 交流成就，庆祝长期成功，以激发员工和利益相关者（包括捐助者），鼓励其他人采用类似的过程或是进行类似的投资。
- *治理角色*: 通过对特定活动和产出的投资监督，以系统的、透明的方式追寻结果和影响，确保一定程度的可问责性；
- *学习角色*: 产生有关变化的信息和视角，并进行分析，以使所有利益相关者都能从相关的关系和过程中学到东西,并相应调整自己的行为 and 干预方式；
- *政策角色*::产生可用以改革政策的数据和分析,发展新的战略,改善政府或者捐献者的实践,或者加强宣传运动。

这使得下列决策复杂化：即测量什么，基线数据是不是必要—用什么指标示是比较恰当，还有数据搜集和分析中的数量、方法和可行性。缺少清晰性会让参与者糊涂，在把这些方法用于实践时缺少关注的焦点，白费力气，使人们质疑钱花得是不是值得。反过来：对要获得的信息的相关性和实用性考虑不够会导致：1、使用不够（即：在某人书桌子上搜集灰尘），2、误用（即：与现实计划和决策系统无关联）；3、过度使用（即：数据作用过大，以至淹没了判断和直觉的功能）（Morgan 1999）。

1.2 权力、控制和所有权：谁的需要和议程？

不同的利益相关者、参与者和听众，有不同的需要和目标，对决策也拥有不同程度的权力和控制。捐助方能施加问责性和管理控制以满足他们以及他们身后的支持者的需要，这些支持者或者是政府部门或者是非政府组织的总部。它往往导致从上到下的集权，这与常规的动力结构相同，它存在技术上的偏见，倾向于外向的，抽取的，非参与的。即便捐助者鼓励参与式的方法，这种方法也往往是在捐助者主导的问责性框架内进行。

另一方面，人们越来越感觉到需要当地的伙伴、社区和受益者真正参与到过程中---采用多重利益相关者的办法。这会推动地区所有权和权力，构建反应能力，学会改善行为，最终实现自治。在这种模式下，捐助者起支持和协调的作用。还有一种办法是两种极端方法的综合：特别是在组织行为者成份比较复杂的时候。

挑战是辨认和理解参与者的需要、动机和议程。通过协商取得真正的认同。

由谁评价影响的问题提出了另一个重要的挑战。近来由基于 Maastricht 的 ECDPM(2003)进行的比较研究得出如下结论：人们越来越意识到在能力发展领域中外主导的评价往往是不恰当，或者是适得其反的。他们提倡更多的自我评价，但经常对热情评价过高，对推动这种学习的复杂性评价过低。然而，内部评价也存在偏见和缺少信誉的问题。

题。

人们也强调了人为操纵和人类出错对自我评价的负面影响(Paton 2003)。在成员或是利益相关者不感到真正的拥有或承诺时更是如此。他们可能觉得，引入体系或指标是种威胁，并将其与惩罚联系起来，或是将其视为官僚控制机制。外部评价往往夸大负面因素，低估正面的，这一事实使事情更糟。它部分是由于在提案阶段项目的正面影响往往被高估（尤其是为了寻找和确保资金），也由于这一影响在一个更广的背景下所引起的变化比较小。结果，人们很可能有受挫的感觉，可能会操纵数据，使其偏向有利的方向。他们也可能在为搜集、分析数据和支持审计追踪的文件所花费的时间和资源持批评态度。

因此，挑战就在于认识和理解参与各方（无论是隐藏的与否）的需要、动机和议程，通过协商就影响评价过程是为了什么、怎么满足不同需要达成共识。它可能因议程中途发生改变变得复杂化，也可能因被那些故意紧抓大权影响程序以满足他们自己组织需要而复杂化。有效的共识得基于对权力关系的理解，这包括一些前提条件和激励，使权力较少的参与者真正理解公开性不会带来惩罚，不与资金决定相挂钩。然而，如果成功，这种认同会保证关键利益相关者的适当所有权，内部和外部于这一过程的各层面的参与也会变得切实可行。

1.3 测量复杂的无形的变化

寻求可见的、数量化的结果是可以理解，特别是幕后捐助者要确保他们的钱花得合理、能引起可以测量的变化。但这导致了一种趋势，即开发高度复杂、综合的框架，强调评价能力建设项目的影响的挑战，这种框架测量的内

容过多，有过多的量化的行为指标。与此相矛盾的是，这种新的体系和框架的发展是在官方机构（包括 USAID）日益呼吁更简单的、更友善的评价发展干预是否成功的体系之时。这样复杂的框架并不能解决如何测量无形的过程或者在特别的输入与较广泛范围内的变化之间是否存在因果关系，这一点，会在 1.4 部分详细讨论。

还有一种观点：要捉住对发展的实践者最重要的变化，我们不能不顾质量，只注重数量。导致我们只关注那些重要的容易被测量的部分(Taylor 2003)。依赖数量和计算会对内在不确定和变化的过程提供虚假的精确，过分依赖于数量化的数据可能意味着真正变化的实质没被纪录或理解。正如爱因斯坦指出的：“不是所有有用的东西都是可计量的”，甚至 USAID，传统上是一个最强烈要求致力于综合的、可测量的指标的，也警告人们这种测量可能带来的虚假精确的危险。他们注意到就其本性而言，组织能力的测量是主观的。他们严重依赖于个人的角度、判断和解释。因此，这是相对而不是绝对的测量，人们可能被误导。(USAID 2000)。

要评价这样复杂的过程，需要综合的、多方位的框架。因此，有人开始使用更系统的或是基于系统的方法，在这种方法中，能力和变化的关系能被勾勒出来，随着人们的行为、关系、活动和行动的变化。(ECDPM 2004)。用系统思考来理解和发展能力是必要的，以便使能力建设与已存在的内部操作相结合，并加强组织在其运作的大系统中的地位。(ECDPM 2004)。这一方法揭示了不同要素和环境间的相互关系，包括内部和外部的，以及组织生活的不同方面（它的内部功能，工作的项目，关系和进化）。

1.4 表明因果关系和变量

怎样把某一项特殊的输入变量与较广范围内的变化联系起来，或是将变化归功于它，这构成了主要的挑战，对评价组织能力建设影响是个非常重要的障碍。内部因素和外部因素的相互作用，以及变化着的环境，使在干预和较广大的变化和归属间建立因果关系的任务复杂化。(Paton 2003; Engelhardt and Lynch 2003)。举例来说，能发现在一个司令部制订的新的或是修改过的战略与现实动作中发生的实际变化间的关系吗？

我们是不是在无任何休止的寻求也许存在也许不存在的基于因果关系的影响时迷失了主要问题？

一个项目或者干预是不是导致长期的变化，这个评价要求相当灵活的方式，能够容纳在更广的范围内的往往是无关的变化。如果遇到多样的行为者、复杂的伙伴关系—现实中的关系通常是这样的，就更困难了。也许更现实的挑战是表明哪种变量导致了变化的结果，而不是试图定义因变量(Mayne 1999)。这样一来，就提出了个人崇拜问题，我们是不是由于不停地搜索基于可能存在也可能不存在的因果关系的影响而迷失了要点？把两者连结起来是不是既灵活又有价值 是基于这样一个假定，即：在他们之间存在一种线性的因果关系。但在现实中，相互间的关系往往更复杂，更有偶发性。干预的效果可能不是自动发生的，也可能不是直接的。它可能要潜藏一段时间后才起作用，也可能引起多重反响。因而需要更基本的辩论，即在现实意义上，是不是存在制造因果关系的辩论，这种连结是不是只是由于捐助者提出要求才成立？

1.5 对环境和文化的回应

越来越多的组织能力建设项目和影响评估是为组织的特殊需要而设计的，并反映他们工作于其间的特殊文化氛围。例如，有证据表明，有的组织(**that organisations operating in high power distance cultures or ones with high levels of uncertainty avoidance**)会更坚持表现测量的概念,更信任测量代表的意见。人们也越来越意识到需要更有包容性的和文化上适当的程序。不同的办法适合不同的环境。比如，在小组和行为者有共同的较一致的利益时，参与的办法更适合，而在有几种目标时就不行，能促进利益相关者间的信息流动的方法更合适。(Morgan 1999).

关键在于恰当地应用基于几个核心原则并适应与大多数环境的影响评价框架，再采用对环境的认识和反应等更灵活的因素相补充。这既有赖于进行评价的个人的跨文化的能力，也有赖于他们为了使该框架在当地文化和氛围内能运作所做的变通是不是有效。文化和氛围的不同可能在比较不同背景下的同类的组织的工作表现上限制我们的能力，。因此，有必要在搜集可比较信息以获得一般性知识和开发灵活和适应特定环境的方法间进行妥协。

有必要在搜集可比较的信息以得出一般性结论，和开发一种灵活的、适应于特定环境的方法间妥协。

1.7

1.6 关于投资费用

影响评估的实践会对被评价的组织提出巨大要求,经常是既费时间又费钱.他们可能受限于没有足够的体制能力和缺少投资、缺乏复杂的技巧以及缺少有效搜集和分析信息的基础条件。越来越多能反映当地文化和社区需要的个性化框架的设计变得昂贵，要求利益相关者的参与，需要动用更多的时间和财政资源才能有效实施。因此，影响评估框架的运转成本很高，要想成功运作，捐助者和经理人需要考虑直接和间接的管理成本及相关经费。否则，评估过程就会管理不善，得到很少有可操作价值的信息，会引起抵制、不满、士气低下。得到的结果还会缺少信度。如果组织和其捐助者没能为进行有效的、适当的测量过程准备充分的费用，就得问问自己，这是否可行。

2 克服挑战：实践中的含义

前面我们讨论了公民社会组织在评价组织能力建设干预影响时面临的种种挑战，这一部分将列举并详细讨论如何克服这些挑战并将影响评价用于实践。这里我们要回顾克服挑战各种方法,将其列举出来，并在后面相应章节中详细论述。很明显，在评估开发项目的影响上已经取得了进步，但人们还不太清楚它们与组织能力建设干预的某些特征的关系。

1. 利益相关者的介入与优先性
2. 自我评价
3. 整合方法
4. 平衡不同的方法和工具
5. 简单与系统化
6. 接受可能的联系，而不是直接的因果性
7. 认清投资层面
8. 组织学习：将评价与行动联系起来

2.1 利益相关者的介入与优先性

经验表明，参与过程能确保当地民众更积极的参与、更大程度的地方所有权，也增加了可信度和质量保证。它还能克服相关的伦理问题，包括就规模和深度，哪些人介入，哪些人可以接触数据等取得一致意见。更直接的介入能保证关键的利益相关者（无论是个人受益者，社区小组，公民社会组织还是捐助者）都能参与解决关键问题，这些问题包括：

- 这一系统主要是为了谁？（即：优先排列所有可能的利益相关者）
- 他们需要什么信息？目的是什么？
- 谁拥有这些信息，什么时候、怎样搜集 这些信息？
- 谁会分析这些信息？
- 这些信息由谁提出，怎样提出？
- 他们会怎么用这些信息？
- 我们会从中学到什么？
- 它如何能对未来的能力建设提供帮助？
- 不同的利益相关者需要的信息都是什么？

对这些问题的回答可能取决于关键行为者对能力建设及其作用的不同想法。利益相关者的不同文化和环境对他们的想法有重要影响，反过来又他们对能力建设会发生什么的假定的概念上的理解发生影响。不同的利益相关者（甚至在同一个利益相关小组的不同个人）对变化怎么发生有自己各自不明确的理论。这些理论往往是下意识的，它意味着，不同的利益相关者对能力建设有非常不同的理解，虽然他们都用同样的术语描述它。为了有效地评价能力建设的影响，有必要在一开始就能力建设的过程和目标及其如何发生达成一些共识。

要有效地评价能力建设项目的
影响，有必要就较广泛的过程、
能力建设的目的以及怎么实现这一
目的达成某种共识。

经验还表明：有必要弄清为什么要进行影响评估。这样的评价活动往往因目标不清或是目的自相矛盾而受影响，这是由于不能弄清角色和恰当地运用术语（见 1.1）。实践证明介入设计影响评价过程需要

区分这些不同角色，以保证证据的搜集、分析是以有益于该评价所期待的。

最后，在这个设计阶段，不同的利益相关者有必要参与决定什么时候测量、测量什么。人们往往对计划的行动、产出、短期结果和长期影响产生混淆。² 因此分清系统是为了谁、需要什么样的信息、怎样获得这些信息、在什么时间内获得这些信息，这些问题相当重要。只陈述说它是为了所有利益相关者往往意味着这样的系统仍然是由外来的捐助者把持着，并不是所有的伙伴都有使用结果产生的信息。它也往往意味着捐助者的需要占统治地位，并由此导致低效率的、官僚式的系统占主导，这就影响了捐助者想要支持的开发工作的质量。每一个利益相关者都有责任提出他们自己需要的信息、要面对这些需要的负担，不仅仅是授权给相关的伙伴代行此事。除非这些区别能被理解并执行，否则影响评估就不会成功。

2.2 自我评估

如果评价过程是要贡献于而不是破坏能力建设的最终目标，能力建设的自我估价是很重要的，过程的所有权越来越视为能力建设的 *sine qua* (James 2002) 因此必须被优先化。人们越来越认识到，除非有一些自我评价的形式，否则，任何影响评价体系的开发都是无价值的或是无效的，这甚至包括合作伙伴对捐助者提供的支持的评价。正如 ECDPM (2003) 所得出的结论：‘自我评价被推荐为能力发展。。使用包括组织的成员和外部利益相关者的参与式的评价方法至关重要。保证评价过程产生有用结果的唯一最有效的方式就是让潜在的用户参与评价的全过程。’

影响评价过程要由这一过程所影响的人和能最好地参与进来的人进行。

类似地，加拿大 IDRC 注意到“接受方从捐助者施加的评价中获得的经验非常有限...在处方和表现变化间有鸿沟...这要求相关方和受影响者的积极介入...这不仅是组织参与自我诊断这样一个简单的问题,它与一个组织应该掌控评价过程这一更基本的问题有关.’ 人们越来越认为，除非有参与式的自我评价，否则这个过程很可能失败。

有各种各样的组织能力评价的工具和方法。通过一个引导的过程，让组织能反思自己的表现，然后选择工具，做出建设能力、扩大影响的战略计划³。这包括 Pact 组织能力评价工具⁴，以讨论为导向的组织自我评价⁵，也包括区域性、与周围环境相关的工具，比如由马拉维 CADECO 开发的工具，它用非洲的谚语加强交流。这一工具以及使用它的经验教训在由 Chiku Malunga 撰写的 PraxisNote 6 “在组织能力建设中使用谚语”中有记载。⁶然而，组织自我评价工具被设计成系统过程，以确认组织能力的当前水平，并不一定适用于评价影响。

² 见 1. 1 中关于区分这些术语的框图。

³ 更多的有关组织评价的信息，可在以下网址找到：

http://www.impactalliance.org/ev.php?ID=7442_201&ID2=DO_TOPIC

⁴ 更多信息可参看：http://www.pactworld.org/services/oca/index_oca.htm

⁵ 更多信息可参看：<http://www.edc.org/GLG/CapDev/dosafile/>

⁶ 该文可从 INTRAC 的网站上免费下载，网址是：www.intrac.org

INTRAC's 在中亚和非洲的经验认为：影响评价过程要由它所影响的人和能最好地介入此事的人所拥有。社区小组、成员和志愿员以及当地的利益相关者的积极参与激发了对这一过程的拥有感，有利于对发现结果的解释。它保证这一体系适合当地的环境，保证了知识的灵活性，能够评价负面的和不曾预料的结果。但最重要的是，它使人们聚焦于关键问题上，即：当地利益相关者当前和长期的需要及现状。

增加参与和自我评估会使发现更准确、更符合实际，可以改善学习，更好利用资源，但它也有一些明显的缺点。包括缺少外部参照点，整个过程可能过于主观、有偏见。还有，人们担心潜在的组织自我欺骗或积极操纵数据。结果。任何自我评价过程都要由其他证据的整合方法，以确保避免产生错觉。因为任何研究过程要想可信，必须不仅有某种程度的力量，还要有基于透明度和整合方法的质量保证。

2.3 整合方法

用最有限的数据获得最可靠的结果的一个重要原则是整合。整合能解决可信度和有效性的问题，因为通过采用不同的来源，方法和观点（比如：以小队为单位工作，使用多重信息源），能抵消任何一种方法带来的偏见。

言及使用不同方法并不意味着这些方法是偶然选中、拼凑的，他们是经过筛选的，以能分别抵消对有效性威胁。这里的关键是对质量方法的系统应用。

办法的准确性来自系统的应用，但一种研究办法的不准确很少能被另一种研究方法的数据补偿。(Fielding and Fielding 1986).

为了保证能有较广的视野，一种办法是采取小组的办法，该小组包括能带来新视野的外来成员，也包括内部成员。小组的办法能提高可信度和有效性，特别是有当地利益相关者的参与。然而，要花费时间和精力保证小组内的不同成员都能参与进来，他们的主张要都能被考虑进去，而不仅仅把他们当作是数据搜集者，甚至只把他们当成代表。为了充分发挥不同观点的作用，在影响评价的一开始就要花时间开发一组参考术语和指标、分清角色、精化数据收集计划、指定特定责任（按照其技巧和经验），这是非常重要的。

2.4 不同方法和工具的平衡

为了帮助整合，需要平衡方法和工具，以满足不同利益相关者的需要。有许多比较简单的方式实现这一点。比如，在非洲工作时，INTRAC 用回归能力评分体系使客户对能力随时间变化的程度能用量化的方式衡量。它依赖于在一开始就搜集到可依赖的、有效的基线数据，对当时的全貌有一个描述。这可以通过传统的问卷，记录叙说证词或者标明对某一特定陈述的赞同程度的问题。它可能测量出变化，但经验告诉我们，这种方法是否相关或精确也存在问题。

I

人们已经担心过分依赖计量测量及其指标，结果，人们越来越接受这样的观点，有必要采用混合的方法和工具以捕捉足够的、适当的和可信的数据。人们意识到计量方法在影响评价领域有局限，因为他们不能做到下述事情：

- 解释为什么有的事情发生了；
- 捕捉不同成分间的关系，这往往比各成分本身更重要；
- 确认在组织内和组织与外来机构间的权力和关系的转移；
- 评估变化的细微差异 – 然而，这样的组织评价工具是缩减的，倾向世俗的，不容易计量的东西就被忽略了；
- 衡量与重要的环境变化相关的变化。

看到这些局限性，人们更加认识到任何可信的能力建设项目的评估都应该采用一系列质量方法论- 包括反思评论，案例，角色扮演，特征分析，画图示及陈述。比如，位于南非的能力建设组织，社区发展资源协会（简称 **CDRA**）有必要进行叙述“故事”以充分探讨人类变化的深度和复杂性。这些故事是最简单的媒介，借此可以描述复杂的情形。通过培养我们讲故事的能力，我们能转达和保护要寻找和已经得到的要点。这种叙述对通过帮助员工和志愿者了解新概念、参加新实践的方式促进组织学习也很重要。（**CDRA 2001; Taylor 2003**）。

Any credible attempt to assess the impact of capacity building initiatives should incorporate a range of qualitative approaches – including reflective commentaries, case studies, role-plays, characterisation or drawing and narratives

CDRA 认为，开发这样的故事是门艺术，是发展实践者需要开发的。有规律的观察，把它织成生动、深刻的故事，然后进一步获得经验和教训，形成概念，这个过程是非常有价值的。实践者得能读出组织变化，也能理解到，自己就是任何过程的一部分。他们需要有反思的能力，甚至在很投入身其中的时候也要这样。换句话说：发展实践者本身就是故事的一部分。但是他们也要认识到，要让这样的叙事

评价 对捐助者来说可信，一定要考虑并解决捐助者焦虑的事情。一旦捐助者确认款项按计划花费了，（.即：通过更传统的框架或表现测量），他们就会放松下来，倾听有关所发生的人和关系的变化的故事。（**Taylor 2003**）。

许多业已进行的实验并不使用事先确定的指标。孟加拉的基督教开发委员会与 **Rick Davies** 合作进行的有关参与监测的实验就是这样一个重要的案例。（**Davies 1998**）。这项实验包括搜集数据，从参与者那里搜集故事，这些故事是基于他们自己项目实行期间发生的“最有意义的变化”的理解。这一方法能确认并肯定那些可能并不是人们事先期待的或能预见的，但对参与者有重要的影响的变化。

这种对捕捉质的变化的态度在“行动援助乌干达”的项目中得到充分体现，该项目对组织变化影响进行了评价。这种评价包括对小组讨论、访谈、和由一位咨询员进行个人叙述的方法，这种叙述能表现出叙述者对身处其间的变化过程的感想。（**Wallace and Kaplan 2003**）。讲故事的作用越来越被捐助者接受，例如，马拉维的 **USAID** 就雇佣了一名成员专门书写这样的故事。

有证据表明，人们越来越接受这样的观点：对评价能力建设和评估影响时，有必要采取混合的办法。（**Roche 1999**；

反思和参与方法的这种使用引出了当地文化的影响问题

Taylor 2003). 但越来越多地使用这种反思和参与的办法引发了如下问题：当地文化对所使用的方法会发生什么影响？在什么程度上，影响评估过程受文化制约，以致不再标准化地使用通用的指标。

实践者越来越认识到使用多种多样的办法，其利远大于弊--高成本。然而，这些评估体系的价值和信誉依赖于他们如何适应当地环境，同时又能被国际社会接受，并能与它者进行比较。要实现上述目标，要考虑以下几点：

- 通用框架和在特定环境下的测量方法间的平衡；
- 使用质量或叙述的过程在设计、搜集和分析的努力和资源上的代价；
- 信息搜集的潜在的主观性；
- 与开发分析和强化不同来源的信息的方法的挑战，这些信息得有一定程度的连续性和可比性。

2.5 简单和系统化

人们对这一系统的代价和应用性一再关心，强调任何影响评估体系都尽可能简单、实用。这样的能力要能提供一种洞察力，让人们能把握能力建设的结果，它需要足够简单以提取发生变化的实质。

地方发展人员对简便的评价体系的要求出人意料地得到了 USAID 的支持下(至少是在原则上，虽然可能在实践上不是如此)，这个机构一贯以坚持严格的评价过程著称。USAID 承认影响评估的局限和复杂性，并宣称：

最佳的测量体系是设计得尽可能简单—不要太浪费时间，不要毫无道理的昂贵，但又能给管理者提供足以满足其需要的良好信息。(USAID 2000:6)

这往往需要优先性---比如被要求从一个多于二十个选项的菜单上选择构成有效领导的三个指标。人们认识到这样的体系不应该依赖复杂的金字塔式的指标，这样的体系应该足够灵活，能随时调整，并能真正反映业已发生的变化。

另一个方法是从系统的角度开发能力建设的效果评估。传统上，人们是从缩减角度进行评估的，着重于单个成分和活动层面上改善的测量。例子应该是某个培训项目或结构重组的效果。目前时兴的也是确定业绩目标然后通过衡量计划和实际成就之间的差异评价影响。

基于系统的办法有长远的观点，探寻一个系统内能力的不同要素间的关联，以及作为其它系统的部分与该系统的关系。

采用体系的方法要有长期的观点，探寻能力建设系统内的不同要素间的关联，以及与其它系统的关系。它实现了简单化，拒绝把能力断裂成互不相关的部分，每一个都有独立的因果指标。能力和影响被看作使结合的结果，它来自内部复杂的关系及内部活动与外部环境间复杂关系。简而言之，焦点是在起作用的不同动力机制，它可能使许多当下的研究方法显出问题。系统的角度发现，基于计划、控制和可预期性

的评价过程的许多方式是错误的。复杂的组织体系不能被控制，他们只能被干扰和动摇。人们在聚焦于事先设定的结果时，能力发展中各方面普遍存在的无意的后果被忽略了。也许，最麻烦的是，对业绩目标的聚焦可能恰好削弱了人们要寻求的目标。反过来，如果整个系统各要素间的关系能得到理解，那么业绩也就呈现出来了。

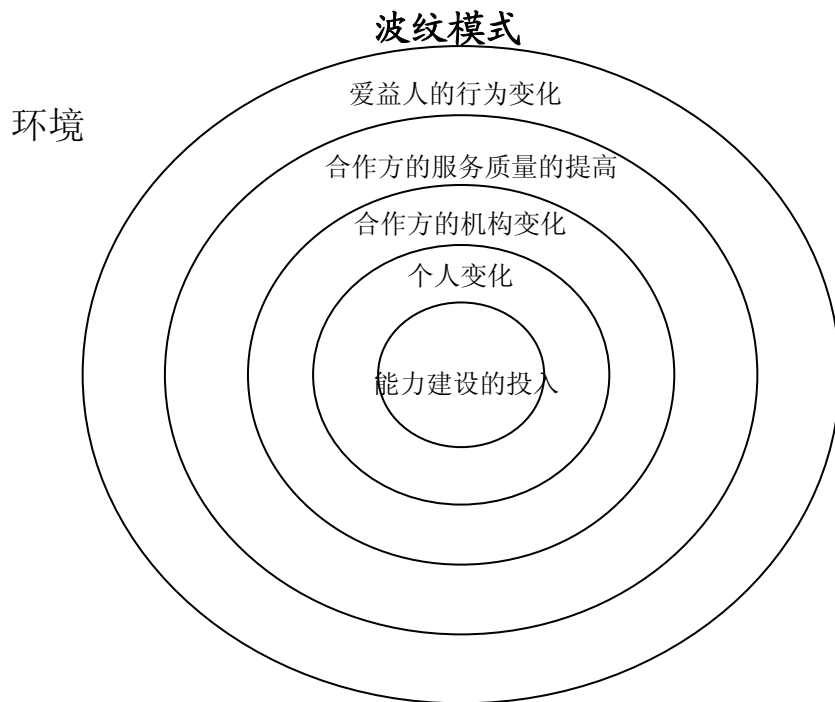
ECDPM 使用的一个分析框架被用来帮助理解能力是怎么产生的，有意的干预怎么支持着它的出现，出现的能力怎么有利于业绩，它使用了下列要素：

<p style="text-align: center;">能力, 变化和行为</p> <p>能力: 在一个既定的环境和指定的系统内，能力采取什么形式？</p> <p>变化与适应: 支持能力建设的变化过程怎样在一个组织或系统中发生？</p> <p>行为: 能力和行为间的相互关系是什么样的？</p> <p>外部环境 外部环境—历史，文化，政治和体制环境以及他们创造的局限和机会，怎样影响了组织或系统的能力和行为？</p> <p>周边系统: 能力建设干预和周边系统是什么关系，特别是组织和体制系统？</p> <p>外部干预: 外界怎么影响了变化的过程？</p>

2.6 接受可能的联系,而不是直接的因果

设计和实施影响评估系统的一个最大挑战是如何解决因果关系问题.在任何一个复杂的、多侧面的,办法几年中在不同时间段不同地点发生许多干预的项目里，很难确认直接的因果联系。这部分是因为干预和变化的多重性，部分是由于人员的变化（特别是领导或治理层的变化），部分是由于外部政治、经济环境的变化,它们产生了许多预料不到的后果。要试图找出某一特定的因果关系，就要有许多解释，说清楚什么引起了变化，把变化归因于任何一个因素或干预都是虚假的、无效的劳动。

人们可以用可能的联系的方法尝试着建立因果联系。用来描述这种联系的一种方法，它已经被用于能力建设培训干预，是波纹模式 I(James 2002):



在这一模式中，能力建设干预就象从天上掉下来的一个雨滴或是落在水面的一颗小石头——波纹向外扩散，引起了个人层面的变化。比如，很可能，一堂培训课程很可能让人们掌握更多的知识，学到新技巧和态度。如果参加能力培训的人能将学到的东西用于组织中，那就很可能在更广一点的组织层面发生变化。成员技巧和内部组织功能的改善也会提高伙伴对社区的服务质量，最后这会导致最终受益者生活上的变化。因此，雨滴的大小和方向既受运动周围环境的影响，反过来又对环境产生影响。

在这个模式中，能力建设干预只是带来变化的许多干预中的一种。然而，通过使用可能的关联，有可能判断出在某一层面的某种变化的确能引起更大层面上的变化。比如，INTRAC 曾就发生了什么变化和为什么得出可能关联采用半结构式的、开放式回答的访谈。它还以编年体的方式列出一个组织发生的主要事情，以此确定特定的能力建设干预是不是能起到有意义的效果。

人们不能将这种模式中从个人到组织的变化视为线性的，有的情况下，两个层面的变化相互强化，但是也有发生了显著的组织变化而个人行为没发生大的变化，或者反过来，个人发生显著变化，组织并没发生大的变化。还有，如果对当地受益者的生活产生影响，它会反过来引起个人和/或组织的变化，比如通过利益相关者的机制。

2.7 确定投资的层面

影响评估和其他表现测量系统失败的一个主要原因是不能或者不愿在设计、开发和实施中投入足够的资金和时间。实施如此复杂的系统既有直接成本也有间接成本。在人员、咨询时间和当地人和机构的时间上成本都很高。也有数据收集、分析和传播的直接成本。有与技能发展和加强对这领域的理解的相关的成本。间接成本包括对内部体系的影响、以及从更迫在眉睫的需要上分散注意力，在个人受挫、有抵制情绪和工作情绪低下给组织带来的成本。

在许多非政府组织中，这样的成本被置于日常费用中，而不是项目资金。因此，捐助者在一开始就给付与这一体系相关的涵盖直接和间接成本的足够资金这一条特别关键。一条清楚的结论是：整个影响评估过程——搜集数据，分析和传播它，以及它的广泛使用是个花费不菲的过程。它应该被看成是有所值的投资，而不仅是附加的体制成本。

2.8 组织学习：将评价与行动联系起来

在许多情况下，影响评估的重点放在了可问责的实现，然而，将得到的结果用于影响计划、调整和改善未来的行动以及开发新项目同样非常重要。为实现这一目标，仅仅搜集数据和信息，而不以有效的、可使用的、适当的方式分析和传播它是不够的。而对数据和信息的有效处理又依赖于投资于合适的基础设施用以搜集信息和/或叙述观点，以及用于开发分析技巧⁷和作出解释的人的洞察力。这一分析，反过来应该能有利于更好地理解 and 回应变化的环境，从而引起行动或实践的改善，

作为整合思想和行为的发展过程，在评价和改善行动间建立联系是一个学习的过程。没有支持学习的组织文化，很难有效建立这一联系。为了支持和保证一种学习的文化，组织需要表明⁸：

- 学习是正当的活动，就是说，学习被看作是每个人工作责任的有机部分，而不要是在个人自己的时间内做的事情。
- 学习要得到鼓励和支持，就是说，管理者要鼓励员工在组织实践和政策的发展中做出贡献，这要成为管理者的责任。
- 为学习提供适当的资源，就是说，有这样的认识，学习是要花时间的，它还可能需要其他资源，包括资金。
- 学习是有回报的。就是说，为组织演进做出贡献的个人，他们的努力得到承认。他们有机会向外界呈现组织的思想。
- 能克服内部的学习障碍。就是说，基于系统分析，设计出针对学习的内部障碍的战略，并让组织内的所有成员者清楚这一点。

3 评论

INTRAC's Praxis 项目的目标是支持公民社会的组织，通过下述方式有效地实现他们的使命，即不断产生、使用和交流组织能力建设领域里的有创见的、因地制宜的实践和研究。正如在这篇 Praxis 论文中所描述的：组织能力建设的一个关键挑战是：有效地评价一种干预是不是产生了长期的可维持的变化。为了加强当前的实践增加其他人研究的价值，这篇论文在结论部分将确认 Praxis 项目在未来的日子将重点关注的话题。

⁷ See the forthcoming Praxis Paper on analytical skills and adaptive capacities.

⁸ From the forthcoming Praxis Paper on organisational learning by Bruce Britton.

首先，有必要更清楚地了解组织能力建设的特点。为了更有效，是否有必要创造一些用于项目影响评估和组织能力干预的不同方法，需要进一步探索是否有可能，甚至是有意愿在原因和结果（即：在有效的组织和发展变化之间）中建立直接联系。虽然在这一领域取得了进展，人们仍需要发展和记载有创新精神的、适当的、可实现的影响评估方法，这些方法能够：

- 面对挑战, 测量重要的、而不是容易的事情
- 用简单、清楚和灵活的方法捕捉和评价组织变化中系统的、多维的、变化的特征
- 采用各种数量和质量的程序
- 让当地社区有更大程度的参与，并整合他们的故事和经验
- 平衡“核心”原则——比如：信任、平等、所有权、可信度、效度和正当性——与适应不同的环境的灵活性
- 认清并回应不同利益相关者的需要和议程
- 使用各种方法分析和强化不同来源的信息，这些信息要有一定程度的连续性和可比性

第二，当影响评估与组织环境内部的改善相联系时似乎更有效，在这种组织环境中，学习和经验受到重视。只有这样，搜集来的信息和分享的知识才能被建设性作用于，以产生深刻的反思和改良的实践。然而，当前缺乏学习和开放的动力。对权力的不平衡没有清楚的认知，许多参与其中的组织对暴露自己的局限和失败觉得不安全，这可能与他们能否得到资金相关联。近来重点从控制和可问责向学习的转变让人们更加乐观，但要克服这些局限还要做更多努力。

第三，提供必要的时间和资源上的投资非常重要，无论是人力还是财力上。所有事例表明成功的影响评估依赖于足够的投资。参与方及他们的捐助者要认识到这一点，否则这一过程实施得就不会好，产生很少有操作价值的信息，承受缺少可信度的痛苦。因此，效果评估应该被视为一种投资，一种能增加组织学习能力的投资，而不是额外的负担。

4 下面的步骤

为了应付本文中提到的挑战，Praxis 项目致力于与公民社会的支持者和研究者进行广泛合作，给他们提供相互间从经验中得到学习的机会。特别地，Praxis 相信在发展中国家和处于转型期国家的当地实践者在产生新理念和新方法中扮演着重要的角色。本论文旨在通过激发人们的论辩接受人们的反应来开放这一过程。然后，下面的步骤将是：

- 成立由有兴趣积极参与推动前面提到的话题的人组成的学习小组，主要用实时方式交流。
- 支持发展中国家和处于转型期国家的当地实践者以 PraxisNotes 的方式书写和传播他们的经验。
- 与当地研究者的组织合作进行案例研究及其深入研究。

- 在 INTRAC 网站的 Praxis 栏目创办一个互动的形式，以分享学习、交流信息和出版物。⁹
- 举办研讨会和学习讨论会，推动反思和分析。

如果你愿意参与这一过程或者有更多的问题需要解答，请访问 INTRAC 的网站或通过电子邮件及下面提供的地址和我们联系。请在你周围传播这篇论文，特别是给那些不容易接触到这类信息或是不容易上网的人。

电子邮箱: infopraxis@intrac.org

地址: Praxis Programme, INTRAC, PO Box 563, Oxford, OX2 6RZ United Kingdom

电话: +44 (0) 1865 201851 传真: +44 (0) 1865 201852

www.intrac.org

⁹ 新的 INTRAC 的网站目前重新设计，并将于 2005 年 3 月投入使用。

参考书目

- Bakewell, O. (2003) *完善发展过程:监测与评价实用指南*, 牛津: INTRAC.
- CDRA (2001) *测量发展:—把握无限*,年度报告,Cape Town: CDRA.
- Cracknell, B. (2002) *评价发展援助 – 议题, 问题与解决*, London: Sage.
- Davies, R. J. (1998) *促进组织学习的新途径:孟加拉基督教开发委员会的实验, 影响评估和项目评价 1 (1)*.
- ECDPM (2003) `评价能力发展', *ECDPM Newsletter*.
- Engelhardt, A. and Lynch, R. (2003) `向上流移动 –监测与评价的新挑战', *SPARCS Newsletter Autumn 2003*, Birmingham: The Parc.
- Fielding, N. and Fielding, J. (1986) *连结数据:社会研究中数量和质量研究方法的表达*, London: Sage.
- Fowler, A. (1997) *达成平衡:建设非政府组织在国际开发中的有效性指南*, London: Earthscan/INTRAC.
- James, R. (2001) *权力和伙伴:非政府组织能力建设经验*, Oxford: INTRAC.
- James, R. (2002) *人民与变化, 开发非洲 NGOs 的能力建设*, Oxford: INTRAC.
- James, R. (2002) *能力建设监测和评价实用指南*, Occasional Paper No. 36, Oxford: INTRAC.
- Lewis, D. (2001) *非政府开发组织管理导论*, London: Routledge.
- Mayne, J. (1999) *通过贡献分析评价归属:合理使用行为测量*,加拿大审议总署名办公室讨论论文.
- Morgan, P. (1999) *能力开放项目行为监测新论:我们正在学习什么?* Quebec: CIDA.
- Neely, A. (2002) *商业业绩测量—理论与实践*. Cambridge: Cambridge University Press.
- Oakley, P. (2001) *评价授权: 概念与实践评论*, Oxford: INTRAC.
- Oakley, P., Pratt, B. and Clayton, A. (1998) *结果与影响: 社会发展评价*, Oxford: INTRAC.

- O'Neill, O. (2002) 'Reith Lectures: 信任与社会, London: BBC.
- Paton, R. (2003) 社会企业的管理和测量, London: Sage.
- Riddell, R.C., Kruse, S-E., Kyollen, T., Ojanpera, S. and Vielajus, J-L. (1997) 寻求影响和方法:非政府组织评价综合研究',给 OECD/DAC 评价专家组的报告, n, 赫尔辛基:外交部国际合作司.
- Roche, C. (1999) 发展性机构的影响评价: 学会关注变化价值 (*Learning to*), Oxford: Oxfam.
- Smillie, I. and Hailey, J. (2001) 应变管理:南亚非政府组织的领导、战略和管理, London: Earthscan.
- Taylor, J. (2003) 发展地使用测量, Cape Town: CDRA.
- USAID (2000) 测量体制能力: 测量与评价的近期实践, *TIPS 15*, Washington: USAID.
- Wallace, T. and Kaplan, A. (2003) 行为援助乌干达获得的开发实践所致变化的经验,工作论文 , London: ActionAid.

缩写表

BBC	英国广播公司
CADECO	能力发展咨询处
CDRA	社区发展资源协会
CIDA	加拿大国际开发署
CSO	公民社会组织
DAC	开发援助委员会
ECDPM	欧洲发展政策与管理中心
IDRC	国际发展研究中心
INTRAC	国际非政府组织培训和研究中心
NGO	非政府组织
OECD	经济合作与发展组织
USAID	美国国际开发署

INTRAC

Postal address: PO Box 563, Oxford, OX2 6RZ

Registered and visiting address: 65 George Street, Oxford, OX1 2BE

Tel: +44 (0)1865 210851 Fax: +44 (0)1865 210852

Email: info@intrac.org Website: <http://www.intrac.org>

INTRAC is a company limited by guarantee No. 2663769

Registered in England A Registered Charity Charity No. 1016676